

Université de Montréal

**Éthique des populations :**  
**Une étude des fondements axiologiques propres aux grandes familles utilitaristes**

*Par*

Octave Arveiller

Département de philosophie, Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures et postdoctorales en vue de  
l'obtention du grade de M.A. en philosophie, option recherche

Juillet 2023

© Octave Arveiller, 2023

Université de Montréal  
Département de philosophie, Faculté des arts et des sciences

---

*Ce mémoire intitulé*

**Éthique des populations :  
Une étude des fondements axiologiques propres  
aux grandes familles utilitaristes**

*Présenté par*

**Octave Arveiller**

*A été évalué par un jury composé des personnes suivantes*

**Marc-Antoine Dilhac**

Président-rapporteur

**Jonathan Simon**

Directeur de recherche

**Christine Tappolet**

Membre du jury

## Résumé

Dans le contexte utilitariste de l'éthique des populations, deux positions s'opposent naturellement à propos de ce qui a de la valeur et de ce que nous devons faire afin de bénéficier au plus grand nombre. Là où certaines intuitions nous invitent à chercher la maximisation totale du bien-être, d'autres soulignent l'importance de faire de notre mieux pour parvenir à la plus haute moyenne possible. Ce document a pour vocation de traiter le débat théorique qui oppose ces deux avenues. Il s'agira donc de décrire et de préciser le fonctionnement et les conséquences de ces positions, afin de tenter d'apporter des éléments de réponses sur la solidité et la robustesse de ces deux axiologies. Dans ce but, nous évaluerons et répondrons aux objections qui leurs sont présentées, et ferons la lumière sur leurs lacunes respectives. Plus précisément, ce document s'intéressera principalement aux problèmes de la Conclusion Répugnante et de la Conclusion Sadique, qui ont fait couler beaucoup d'encre, mais également à d'autres objections couramment discutées dans la littérature. Nous concluons que contrairement à ce qu'ils peuvent laisser croire, ce ne sont pas des arguments décisifs à l'encontre de ces positions. Il s'agira de surcroît d'interroger et d'expliquer les intuitions qui fondent ces positions, les biais qui les menacent mais également les conséquences contre-intuitives qui en découlent.

*Mots-clés* : éthique des populations, utilitarisme, théorie des valeurs, conclusion répugnante, conclusion sadique, long-termisme, justice intergénérationnelle.

## **Abstract**

In the utilitarian context of population ethics, two positions are naturally opposed about what has value and what we ought to do to benefit the greatest number. Where some intuitions invite us to seek total well-being, others stress the importance of doing our best to achieve the highest possible average. This paper aims to address the theoretical debate between these two avenues. It will therefore be necessary to describe and clarify the functioning and consequences of these positions, in order to try to provide elements of answers about the solidity and robustness of these two axiologies. To this end, we will assess and respond to the objections presented to them, and shed light on their respective shortcomings. More precisely, this paper will focus on the problems of the Repugnant Conclusion and the Sadistic Conclusion, which have caused a lot of ink to flow, but also on other objections commonly discussed in the literature. We will conclude that contrary to what they may suggest, these are not decisive arguments against these positions. In addition, the document will question and explain the intuitions that underlie these positions, the biases that threaten them but also the counter-intuitive consequences that follow.

*Keywords* : population ethics, utilitarianism, value theory, repugnant conclusion, sadistic conclusion, long-termism, intergenerational justice.

# Table des matières

Résumé.....	3
Abstract.....	4
Table des matières.....	5
<b>Liste des tableaux.....</b>	<b>9</b>
<b>Liste des figures.....</b>	<b>11</b>
I. Introduction.....	15
A. Préambule à l'éthique des populations.....	15
B. Définition et terminologie utilisée.....	17
1. Les trois variables du débat.....	19
C. La prémisse utilitariste.....	21
II. Chapitre I – Fonctionnement des théories majeures de l'éthique des populations.....	23
A. La position totale.....	24
1. Deux intuitions vers la position totale.....	25
i. Le plus grand bien c'est le plus de bien-être possible.....	25
ii. Le présent n'occupe pas une place particulière.....	26
2. Le fonctionnement de la position totale.....	27
i. Arithmétique.....	27
ii. Géométrie.....	28
3. Les conséquences de la position totale : le long-termisme.....	30
i. Longévité.....	32
ii. Population.....	33
iii. Qualité de vie.....	35

B. La position moyenne .....	37
1. Des intuitions divergentes .....	37
i. L'échelle individuelle .....	40
ii. Ajouter de nouveaux individus : l'intuition de neutralité .....	40
C. Nouvelles approches des problèmes d'éthique des populations .....	42
1. Les théories à valeur variable .....	42
2. Les théories de seuil critique et de gamme critique.....	45
i. Les théories de seuil critique.....	45
ii. Les théories de gamme critique .....	47
3. Les théories non impersonnelles ( <i>person affecting views</i> ).....	49
i. Fonctionnement .....	49
Conclusion .....	50
III. Chapitre II – Critiques de la position totale.....	52
A. La Conclusion Répugnante .....	52
1. La problématique.....	52
2. Les solutions envisagées .....	57
i. Revoir les intuitions .....	58
a. Le biais égoïste .....	60
b. Le biais des extrêmes .....	60
c. Le biais de la sous-évaluation de la valeur des vies .....	61
ii. Tolérer et contre-attaquer .....	63
3. Les avenues alternatives .....	65
i. Les théories à valeur variable.....	65
ii. Les théories de seuil critique.....	66

iii. Les théories non impersonnelles.....	66
iv. La réponse de la position moyenne .....	67
v. La réponse du justicisme .....	70
a. Fonctionnement de la théorie.....	70
b. Rejet de la Conclusion Répugnante par le justicisme .....	74
B. Les conflits de temporalité .....	77
1. La problématique.....	77
2. La proposition de résolution .....	79
Conclusion .....	83
IV. Chapitre III – Critiques de la position moyenne.....	85
A. La Conclusion Sadique.....	85
1. La problématique .....	85
i. Le sadisme de la Conclusion Sadique.....	87
ii. À propos de la comparaison utilisée .....	88
B. Les situations dans lesquelles des actions n’ont pas d’effets sur certains individus .....	90
1. De « haut en bas » : la critique de Huemer .....	91
2. De « bas en haut » : la critique de Parfit.....	93
i. La population pertinente est intemporelle .....	94
ii. La population pertinente n’est que la population du présent .....	95
iii. La population pertinente n’est que la population de l’avenir.....	96
3. Une réponse plus générale à la détermination de la population pertinente.....	97
C. La Conclusion Infernale .....	100
D. Difficultés apparaissant lorsque la population grossit.....	102
1. Remise à plus tard de la souffrance.....	102

2. Le problème de la position moyenne lorsque la population est petite .....	104
Conclusion .....	110
V. Conclusion Générale .....	112
Références bibliographiques.....	115



## Liste des tableaux

Tableau 1. – Distributions pour trois populations .....	41
Tableau 2. – Comparaison des scénarios 1 et 2.....	54
Tableau 3. – Comparaison des scénarios 1, 2 et 3.....	54
Tableau 4. – Comparaison des scénarios 7 et 8.....	98



## Liste des figures

Figure 1. –	Monde A.....	28
Figure 2. –	Monde B.....	29
Figure 3. –	Monde C.....	38
Figure 4. –	Monde D.....	39
Figure 5. –	Monde E.....	44
Figure 6. –	Monde E et seuil critique $\alpha$ .....	46
Figure 7. –	Monde F.....	48
Figure 8. –	Scénario 1.....	55
Figure 9. –	Scénario 3.....	55
Figure 10. –	Graphique I.....	73
Figure 11. –	Graphique II.....	107



*Au sublime futur qui attend l'Humanité.*



# I. Introduction

## A. Préambule à l'éthique des populations

Le 15 novembre 2022, un nouveau cap fut franchi lorsque la population mondiale passa finalement la barre des huit milliards d'individus<sup>1</sup>. Afin de réaliser à quel point nous sommes aujourd'hui nombreux à coexister, des estimations récentes rapportent qu'un total de plus de 117 milliards d'humains<sup>2</sup> auraient foulé le sol de notre planète commune dans les 200 000 dernières années. Autrement dit, près de 7 % de tous les humains ayant un jour vécu sont encore vivants à l'heure actuelle ; et ce chiffre ne fait qu'augmenter tandis que les projections des Nations Unis sont au-delà des 8 % pour 2050, à l'approche du cap symbolique des 10 milliards d'individus vivant simultanément. Dans le même temps, et il va de soi qu'elles sont en grande partie responsables de cet accomplissement, les avancées scientifiques et technologiques se font de plus en plus rapides, se déployant massivement dans notre quotidien et transformant nos modes de vie et notre environnement plus rapidement et plus radicalement que jamais.

Alors que le niveau de vie de l'humanité est à son plus haut historique, il reste nécessaire de mitiger notre optimisme en précisant les fondations philosophiques permettant de guider de manière claire et concrète les actions de notre civilisation. Le pouvoir dont s'est doté l'espèce humaine en quelques siècles est tel que nous sommes désormais en mesure de redessiner drastiquement la portée et les limites de notre avenir, de sorte qu'il convient de s'interroger sur la direction à prendre pour laisser le meilleur monde possible à nos enfants. La formule est consacrée, l'intention louable, mais le tout reste inacceptablement vague. Que dois-on chercher à maximiser exactement ? Comment définir qu'une situation est meilleure qu'une autre ? Quelles sont nos priorités morales ? De Aristote à Bentham, en passant par Kant, le milieu savant n'a eu de cesse de chercher à résoudre le débat théorique de ces questions et de mettre en œuvre les politiques les plus adaptées à l'atteinte du plus grand bien. Pourtant, bien que la plupart des

---

<sup>1</sup> United Nations, «Day of Eight Billion, 15 November 2022 | Population Division», 2022.

<sup>2</sup> Toshiko Kaneda, Carl Haub, «How Many People Have Ever Lived on Earth?», 2022.

traités d'éthique aient étudié la question, à l'instar des fameux *Éthique à Nicomaque*<sup>3</sup> et *Éthique de la Raison Pratique*<sup>4</sup>, il ne semble pas que l'humanité se dirige vers les avenues les plus souhaitables. L'intérêt pour l'amélioration de la qualité de vie des générations présentes et futures est pourtant bien présente, et l'on assiste aujourd'hui à une nouvelle vague d'organisation citoyenne et académique cherchant à bâtir un futur meilleur. Cette situation s'explique notamment par les nombreuses réactions face aux nouvelles technologies capables de verrouiller de manière plus ou moins permanente l'avancement humain, et par la divergence d'opinions en ce qui concerne notre responsabilité de contribuer à la perdurance de la civilisation.

Les implications de ce débat semblent vertigineuses. D'une part, nos vies et nos actions sont en partie dictées par des considérations qui s'intéressent généralement aux conséquences prévisibles sur des échelles de temps allant de la seconde à au plus quelques décennies, très rarement au-delà. De l'autre, le passé est une goutte d'eau dans l'océan du futur et il existe un lien causal fort entre les décisions prises aujourd'hui et des systèmes plus larges comme la transmission des valeurs sur des millénaires, ou encore la vitalité des écosystèmes<sup>5</sup>. L'enjeu est si grand qu'il motive à lui seul la rédaction de ce document. Nous admettons donc ici notre sympathie pour les positions dont les intuitions sont les plus en phase avec ces préoccupations. Tout au long de ce mémoire, nous garderons en tête que les implications mentionnées sont d'une importance telle, et que la période actuelle pourrait être tellement décisive à leur propos, que les recherches d'éthique des populations sont une priorité morale clé de notre temps. La responsabilité des générations présentes devrait probablement être engagée, de sorte qu'il est de notre devoir de nous interroger sur ce domaine de la philosophie. Cela demande donc avant tout de comprendre l'état de l'art de la littérature, ce qui sera au cœur de notre propos.

Ce document se présente donc avant tout comme un état des lieux du débat historique entre les positions traditionnelles que sont l'utilitarisme total et l'utilitarisme moyen. Bien que nous soulignerons à maintes reprises l'importance du débat et que l'analyse sera présentée dans

---

<sup>3</sup> Aristote, 384-322 avant notre ère.

<sup>4</sup> Emmanuel Kant, 1788.

<sup>5</sup> Or, il peut paraître délirant de vouloir influencer le temps long, comme si l'avenir dépendait de forces hors de notre contrôle, de sorte que l'on pourrait être tenté de plaider pour un certain laisser-faire de l'ordre naturel des choses. Nous irons ici à l'encontre de cette idée en cherchant à défendre une philosophie de l'action et de la responsabilisation.



cet esprit, le problème sur lequel nous voulons nous concentrer n'est pas de sensibiliser le lecteur à l'éthique des populations mais bien de présenter les différents arguments en faveur et contre les positions traditionnelles, afin d'éclaircir et de mettre en évidence les lacunes persistantes. Pour ce faire, notre objectif sera dans un premier temps de parvenir à une clarification du fonctionnement des positions mentionnées. Afin d'en comprendre les subtilités, il sera pertinent de décrire également le fonctionnement de certaines théories connexes dans la mesure où celles-ci se sont formées en réponse au débat entre les deux premières. Elles apportent donc des clés de compréhension essentielles que nous pourrons exploiter par la suite. Sur cette base, nous chercherons alors à adresser les critiques qui froissent nos intuitions et empêchent tout consensus. Le troisième et dernier objectif sera de présenter les solutions proposées par les théoriciens. Le premier chapitre sera ainsi consacré au fonctionnement et aux représentations des positions totale, moyenne, à valeurs variable, des théories critiques et des positions non impersonnelles. Le second chapitre aura pour objet de préciser les difficultés majeures découlant de la position totale, et plus précisément les problèmes de la Conclusion Répugnante et des conflits de temporalités. Nous discuterons alors certaines réponses de la littérature. Enfin, le dernier chapitre aura une structure similaire et s'intéressera aux problèmes de la position moyenne. Les trois critiques abordées seront celles de la Conclusion Sadique, celles des situations dans lesquelles des actions n'ont pas d'effets sur certains individus, ainsi que celle de la Conclusion Infernale.

## **B. Définition et terminologie utilisée**

Avant de nous intéresser plus précisément à sa définition, comprenons bien que l'éthique des populations a pour objet les problèmes éthiques qui se posent lorsque nous cherchons à considérer et à comparer des groupes d'individus. Discipline relativement nouvelle dans le domaine de l'éthique, elle a toutefois des antécédents dans les domaines de la sociologie, de la politique, de l'économie ou encore de la philosophie politique. Elle pose donc des questions qui concernent les raisons morales en faveur ou contre la création de nouveaux individus, mais aussi les principes auxquels se référer pour déterminer la taille et la composition des populations

futures. Cette discussion s'inscrira plus précisément dans la lignée des arguments de Derek Parfit issus de son *Reasons and Persons* de 1984, de sorte que nous étudierons et préciserons les positions défendues par les partisans des positions utilitaristes dites totales et moyenne. Nous nous intéresserons par extension aux raisons morales de tenter d'assurer la survie indéfinie de l'humanité, et soulèverons des questions à propos des obligations que nous avons à l'égard de nos descendants, ce qui nous amènera inévitablement à parler des fondements de ces obligations<sup>6</sup>.

Afin de ne pas se méprendre sur la signification des termes relatifs à la discipline, nous sommes d'avis que les plus employés d'entre eux devraient être définis pour des raisons de précision, ainsi que pour permettre au lecteur de s'y référer au besoin. Cependant, il est habituel que certains auteurs aient leur propre définition de ces termes ; nous fournirons donc des mises en garde explicites en temps utile.

La discipline elle-même se définit selon Franz et Spears comme l'étude des propriétés axiomatiques des ordonnancements sociaux des vecteurs de bien-être pour différentes tailles de populations<sup>7</sup>. Dans ce cadre, une *population* est une entité qui représente l'ensemble de toutes les vies d'un monde donné. Une *sous-population* est une composante de cette entité, de sorte qu'une population est constituée d'un nombre fini de sous-populations<sup>8</sup>. Pour ce qui a trait au bien-être, le *niveau de bien-être d'une vie*, définie selon Gustafsson comme la valeur personnelle de cette vie pour l'individu qui la vit<sup>9</sup>. Nous présumons ici que la valeur personnelle est mesurable de manière interpersonnelle, de sorte qu'un niveau de bien-être dit bon, mauvais ou neutre est un niveau de bien-être tel qu'une vie à ce niveau serait elle-même bonne, mauvaise ou neutre, pour la personne qui la vit. Les valeurs d'utilité des individus sont ici normalisées de sorte qu'un niveau d'utilité de zéro correspond à une vie neutre pour l'individu en question, c'est-à-dire une vie qui est, du point de vue de celui qui la mène, ni meilleure ni pire qu'une vie

---

<sup>6</sup> Jefferson McMahan, «Problems of Population Theory», 1981, p. 96.

<sup>7</sup> Nathan Franz et Dean Spears, «Mere Addition is equivalent to avoiding the Sadistic Conclusion in all plausible variable-population social orderings», 2020, p. 1.

<sup>8</sup> Les trois définitions suivantes sont tirées de Johan E. Guðstafsson, «Population axiology and the possibility of a fourth category of absolute value», 2019, p. 81-82.

<sup>9</sup> Nous verrons par la suite que les niveaux sont définis différemment selon les auteurs, qui leurs attribuent des mesures et des échelles qui leur sont propres.

dépourvue d'expériences<sup>10</sup>. Dans ce contexte, une vie vaut la peine d'être endurée si et seulement si son niveau d'utilité est positif, raison pour laquelle nous utiliserons l'utilité individuelle comme indicateur de la qualité de vie. Enfin, rappelons qu'une *distribution* mathématique est une liste ordonnée de niveaux de bien-être, pouvant éventuellement contenir des répétitions, où chaque population finie détermine une distribution. Dans cette optique, la *taille d'une distribution* est simplement le nombre d'individus concernés, ce qui revient à parler de la longueur de la liste en question<sup>11</sup>.

Les termes plus anecdotiques seront quant à eux définis dans le texte ou dans les notes de bas de page, lorsque nécessaire.

## 1. Les trois variables du débat

Afin de mieux illustrer ce qui fonde une axiologie, attardons-nous un instant sur les trois variables centrales du débat, à savoir le bien-être, le nombre et l'identité des parties prenantes. Combinées, ces variables forment le cadre systématiquement utilisé par les théoriciens des populations, bien que des divergences nettes apparaissent vis-à-vis de l'importance accordée à chacune. Par exemple, le nombre de vies en jeu est d'une importance cruciale chez les totalistes comme MacAskill tandis qu'il revêt une importance secondaire pour les partisans de la position moyenne comme Pressman. La quantité de bien-être se situe toujours au centre de la discussion, mais plusieurs arguments ont été amenés pour justifier ce que l'on doit entendre par « maximisation du bien-être ». L'identité est quant à elle une valeur incontournable des auteurs qui prônent des positions non-impersonnelles. Dans cette section, nous détaillerons ces variables.

En premier lieu, l'éthique des populations s'intéresse à la notion de bien-être en cela que comparer les divers scénarios, c'est avant tout comparer des individus dont les vies peuvent être plus ou moins désirables<sup>12</sup>. Le bien-être individuel est donc un élément central de la prise de décision éthique, et une vie très épanouie et pleine de sens sera considérée intrinsèquement plus

---

<sup>10</sup> Walter Bossert, Susumu Cato et Kohei Kamaga, «Revisiting variable-value population principles», 2022, p. 4.

<sup>11</sup> Teruji Thomas, «Some Possibilities in Population Axiology», 2017, p. 811.

<sup>12</sup> Johan E. Guðafsson, «Population axiology and the possibility of a fourth category of absolute value», 2020, p. 83.

désirable qu'une vie de souffrance. Peu importe l'indicateur de niveau de vie utilisé, il semble raisonnable de comparer des vies en leur donnant respectivement un niveau de bien-être qui sert de base relative pour justifier qu'une vie est meilleure qu'une autre<sup>13</sup>, voire de base absolue pour justifier qu'une vie est strictement bonne ou mauvaise<sup>14</sup>. Ce document se concentrera donc en particulier sur la comparaison de populations vis-à-vis de la distribution du bien-être.

Toutefois, les populations ne sont pas fixes, et un deuxième élément valorisé ici par les tenants des positions impersonnelles est le nombre d'âmes qui vivent dans un monde particulier, car il est évident que les choix éthiques peuvent entraîner des conséquences différentes en fonction de la taille de la population concernée. Il convient donc aux différentes positions de proposer un arbitrage visant à rendre compte de cette réalité, notamment vis-à-vis de la quantification du bien-être<sup>15</sup>.

Troisièmement, certaines approches que nous appellerons ici non impersonnelles mettent l'accent sur l'identité des personnes qui vivent dans un monde particulier. Ces approches ont généralement comme prémisse qu'une chose ou une situation ne peut être bonne (ou mauvaise) que pour quelqu'un en particulier. De ce point de vue, ce serait un contresens d'essayer de comparer des mondes où les individus ne sont pas les mêmes dans la mesure où les seules comparaisons valides ne peuvent que provenir que de situations différentes pour un individu particulier. Cependant, gardons en tête que la majorité des théories étudiées ici seront anonymes dans le sens où elles tiennent pour acquis que la valeur d'une population ne dépend aucunement de l'identité des individus comparés, et nous justifierons en temps utiles les fondements de cette affirmation<sup>16</sup>. Pour l'heure, nous dirons simplement que cette hypothèse d'anonymité nous permettra de proposer des comparaisons évaluatives entre des distributions telles que définies précédemment.

---

<sup>13</sup> Michel Pressman, «A Defence of Average Utilitarianism», 2015.

<sup>14</sup> Johan E. Guðafsson, «Population axiology and the possibility of a fourth category of absolute value», 2020, p. 83.

<sup>15</sup> Comme chez Derek Parfit, «Reasons and Persons», 1986.

<sup>16</sup> En faisant notamment appel à Derek Parfit, «Reasons and Persons», 1986, p. 358.

## C. La prémisse utilitariste

« *The task of the benevolent is surely to diligently seek to promote the benefit of the world and eliminate harm to the world and to take this as a model throughout the world. Does it benefit people? Then do it. Does it not benefit people? Then stop.* »<sup>17</sup>

Lorsque vient le temps de se demander comment promouvoir le plus grand bien, ou plus simplement comment agir de manière éthique, plusieurs avenues sont envisageables qui cherchent à réconcilier nos intuitions morales tout en nous outillant pour développer une théorie qui nous dit précisément quoi faire, et pourquoi<sup>18</sup>. L'utilitarisme est l'une de ces théories qui œuvre au progrès social<sup>19</sup>, que l'on peut considérer comme l'une, sinon la plus intuitive des théories morales, et qui a historiquement occupé une place prépondérante dans les débats éthiques depuis le XVIIIe siècle. De ce fait, bien que sa nature nous empêche de prouver sa véracité, il nous semble que celle-ci présente plusieurs avantages intéressants qu'il convient de noter et qui plaident en sa faveur. Nous en listerons ici quatre, et bien qu'il ne soit pas dans les projets de ce mémoire de résoudre le débat qui oppose l'utilitarisme aux théories morales concurrentes, nous expliquerons tout de même les problèmes les plus flagrants de ces dernières.

Tout d'abord, suivant la méthodologie de l'équilibre réflexif<sup>20</sup> (*reflective equilibrium*) nous cherchons un juste milieu entre nos intuitions particulières et des considérations plus théoriques comme la plausibilité de principes généraux systémiques à propos de ce qui a de la valeur<sup>21</sup>. Bien que nous admettions que certaines expériences de pensées utilitaristes amènent des résultats contre-intuitifs, elles permettant toutefois de suivre une ligne de conduite claire et systématique qui paraît généralement juste. De plus, et parce que l'impartialité est une valeur qui semble juste,

---

<sup>17</sup> Chris Fraser, «The philosophy of the Mozi : The First Consequentialists», 2016, p. 138. Mozi 32/1-2.

<sup>18</sup> Richard Chappell, Darius Meissner, «An Introduction to Utilitarianism», 2023.

<sup>19</sup> Voir par exemple Jeremy Bentham, «An Introduction to the Principles of Morals and Legislation», 1789.

<sup>20</sup> Norman, Daniels, «Reflective Equilibrium», 2003.

<sup>21</sup> Richard Chappell, Darius Meissner, «Arguments for Utilitarianism», 2023.

nous acceptons de nous mettre derrière le voile d'ignorance<sup>22</sup> afin d'esquiver les biais personnels et pour mieux déterminer ce qui devrait être fait idéalement.

Les cadres théoriques concurrents que sont le déontologisme et l'éthique des vertus sont également envisageables dans le contexte de l'éthique des populations. Toutefois, afin de ne pas élargir davantage le cadre établi précédemment, nous limiterons notre attention aux points de vue liés à l'utilitarisme<sup>23</sup>.

---

<sup>22</sup> Originellement William Vickrey, «Measuring Marginal Utility by Reactions to Risk», 1945, p. 329, et John Harsanyi, «Cardinal Utility in Welfare Economics and in the Theory of Risk-taking», 1953, p. 434-435. Le voile d'ignorance est aujourd'hui associé à John Rawls, « A Theory of Justice », 1971.

<sup>23</sup> Richard Chappell, Darius Meissner, «Arguments for Utilitarianism», 2023.

## II. Chapitre I – Fonctionnement des théories majeures de l'éthique des populations

L'éthique des populations a donc pour objet les questions morales propres aux groupes qui découle de la manipulation des variables de nombre, de bien-être, voire d'identité. Toutefois, bien qu'elle s'intéresse à des questions variées telles que l'évaluation de la crise environnementale sur les populations et l'allocation intergénérationnelle des ressources, nous nous cantonnerons ici au débat théorique qui fonde les théories des valeurs. Nous n'aborderons donc pas de questions pratiques en tant que telles sur ce qu'il serait bien ou mal de faire, ni de prescrire de comportement à adopter. Dans la continuité de la question de Henri Sidgwick, nous cherchons à comprendre « Is it total or average happiness that we seek to make a maximum? »<sup>24</sup>.

Pour commencer, nous présenterons donc les intuitions qui les soutiennent respectivement et illustrerons la manière dont ces théories évaluent certaines situations. Nous aborderons donc plus largement dans ce chapitre les avenues qui trouvent leurs origines dans la position totale, comme le longtermisme et les théories critiques, et décrirons leurs composantes majeures. Nous ferons de même avec les alternatives principales sensibles aux arguments de la position moyenne comme les théories à valeur variable. Nous en profiterons par ailleurs pour faire la démonstration de l'incapacité des positions non impersonnelles à répondre aux défis relevés par les positions d'origine totale et moyenne. Les chapitres subséquents seront quant à eux dédiés à une analyse plus approfondie des arguments qui s'opposent à ces deux familles de positions rivales.

---

<sup>24</sup> Henry Sidgwick, «Methods of Ethics», 1874, p. xxxvi.

## A. La position totale

« The point up to which, on Utilitarian principles, population ought to be encouraged to increase, is not that at which average happiness is the greatest possible...but that at which the product formed by multiplying the number of persons living into the amount of average happiness reaches its maximum. »<sup>25</sup>

La position « totale » soutient que la désirabilité de la distribution d'une population est proportionnelle au bien-être total qu'elle contient<sup>26</sup>. Dans ce contexte, c'est avant tout l'augmentation du nombre d'individus ainsi que de leur qualité de vie qui contribue à améliorer le monde, car la valeur d'une distribution de population est vue comme la somme des valeurs des vies des membres de cette distribution. En d'autres termes, chaque vie supplémentaire est considérée désirable tant et aussi longtemps qu'elle contient du bien-être positif, et dans cette optique, l'identité de ceux qui accaparent le plus de bien-être n'est que secondaire<sup>27</sup>. Il s'ensuit que dans ce cadre d'évaluation éthique, un monde est considéré meilleur qu'un autre si la somme du nombre d'individus qu'il contient et de leur qualité de vie est plus importante, de sorte que la désirabilité d'une politique n'est jugée que sur sa capacité à augmenter le bien-être total, même si cela a pour conséquence collatérale d'entraîner une diminution du bien-être individuel de certains individus. Ainsi, par son caractère décisif vis-à-vis de l'évaluation de la désirabilité d'une distribution, l'approche peut sembler injuste. Comment dans ces conditions concilier le besoin de maximiser le bien-être total tout en préservant et en donnant de la valeur à des notions comme la justice ou l'équité ?

En outre, la position totale soulève des questions importantes quant aux limites de la population humaine. En effet, comme nous allons le montrer plus en détails lorsque nous aborderons les conséquences de soutenir une telle position, il est raisonnable de croire que le nombre d'individus amenés à exister pourrait être plusieurs ordres de grandeur supérieurs au nombre d'individus ayant vécu depuis le début de l'humanité. Ainsi est-il parfaitement

---

<sup>25</sup> Henry Sidgwick, «The Methods of Ethics», 1907, Livre 4, Chapitre 1, Section 2.

<sup>26</sup> Kieran Setiya, «The New Moral Mathematics», 2022.

<sup>27</sup> Thomas Teruji, «Some Possibilities in Population Axiology», 2018, p. 811.



envisageable que des milliers de milliards d'enfants de l'humanité naissent un jour<sup>28</sup>. Or, le futur de ces personnes pourrait être désirable si tant est que nous prenions des décisions stratégiques qui maximisent les chances que cela advienne, car nos actions peuvent en partie déterminer l'existence, la valeur et l'état de la planète sur laquelle vivra cette foule<sup>29</sup>. C'est-à-dire qu'une quantité faramineuse de bien-être existe probablement en puissance. Selon ses partisans, cette énorme responsabilité qui pèse sur nos épaules devrait dès lors suffire à nous pousser à l'action, car force est d'admettre qu'il existe comme une nécessité de contribuer à augmenter le bien-être total<sup>30</sup>. C'est devant cette constatation que la position totale soutient en pratique des politiques qui cherchent à mitiger les risques existentiels importants, tels que les menaces de pandémie, les impacts environnementaux, les dérives autoritaires, les dangers célestes ou encore les risques liés aux technologies. Elle nous invite ainsi à une réflexion profonde sur les limites et les enjeux de la croissance démographique et de la gestion des ressources collectives pour le bien-être de l'humanité.

## 1. Deux intuitions vers la position totale

En soutenant que la valeur morale d'une action doit être évaluée en fonction de sa capacité à participer au bien-être total, plusieurs avenues sont possibles. L'une d'entre elle est de valoriser la maximisation totale de la quantité de bien-être. Cette approche intuitive est motivée par deux raisons au moins, que nous allons décrire dans cette section.

### i. Le plus grand bien c'est le plus de bien-être possible

D'une part, la position totale suppose qu'en tant qu'axiologie elle se doit de théoriser un système hiérarchisant selon un principe universel les valeurs sociologiques et morales. Pour elle,

---

<sup>28</sup> Kieran Setiya, «The New Moral Mathematics», 2022.

<sup>29</sup> William MacAskill, «What We Owe the Future», 2022, p. 21.

<sup>30</sup> Kieran Setiya, «The New Moral Mathematics», 2022.

le bien est directement proportionnel à l'utilité, de sorte que le plus grand bien découle de la plus grande utilité totale, et non pas de la plus grande utilité moyenne<sup>31</sup>. La théorie préconise donc d'améliorer la vie des individus et d'amener à l'existence le plus de personnes heureuses possibles<sup>32</sup>.

## ii. Le présent n'occupe pas une place particulière

D'autre part, l'idée que « tout le monde compte » capture une partie de l'essence de cette position dans la mesure où elle reflète la croyance selon laquelle le présent n'a pas une place fondamentalement différente des autres périodes temporelles lorsque l'on cherche à découvrir comment évaluer le bien-être des populations. Les bénéfices intrinsèques sont donc égaux peu importe la temporalité dans laquelle les individus vivent<sup>33</sup>. Autrement dit, la temporalité n'est pas pertinente lorsque l'on s'intéresse aux conséquences de certaines actions<sup>34</sup>. Pour s'en convaincre, voyons plutôt l'expérience de pensée de Moorhouse de 2021<sup>35</sup> : enfouissez une bouteille cassée dans le sable puis marchez dessus pieds nus ; vous allez souffrir, ce qui est dire que vous allez réduire votre bien-être de Z. Par la suite, si personne ne s'occupe du sort de cette bouteille, il est envisageable que quelqu'un d'autre, disons un siècle plus tard, se promène à son tour pieds nus sur cette plage et finisse éventuellement par marcher sur ce même tesson. Toutes choses étant égales par ailleurs, nous conviendrons que sa souffrance n'a aucune raison à priori d'être différente de la nôtre un siècle plus tôt. Son bien-être sera également réduit de Z. C'est sur base

---

<sup>31</sup> Richard Chappell, Darius Meissner, «An Introduction to Utilitarianism», 2023.

<sup>32</sup> «A natural principle to invoke—referred to in this paper as the Principle of Maximization of the Numbers of Worthwhile Lives (Principle M)» dans Sreenivasan Subramanian, «The Repugnant, the Sadistic, and Two «Despotic» Conclusions in Population Ethics», 2023, p. 17.

<sup>33</sup> Il est vrai que des théories utilisant un taux d'escompte existent toutefois, mais cette approche est bien moins acceptée des philosophes qu'elle n'est présente dans les analyses économiques (Fin Moorhouse, «Longtermism: An Introduction», 2021). En effet, il semble que nous ayons de bonnes raisons d'escompter les bénéfices économiques, la raison principale étant qu'en recevant de l'argent plus tôt nous sommes capables de l'investir pour produire un dividende. Toutefois cette raison ne s'applique pas nécessairement à la qualité de vie, puisqu'il est impossible d'investir directement du bien-être aujourd'hui pour en récupérer davantage plus tard (Benjamin Todd, «Longtermism : the moral significance of future generations», 2017).

<sup>34</sup> Moorhouse note toutefois qu'il pourrait y avoir des raisons *pratiques* de préférer s'intéresser au court terme. Nous élaborerons sur cela au prochain chapitre.

<sup>35</sup> Fin Moorhouse, «Longtermism: An Introduction», 2021.

de cette intuition que la position totale reconnaît en réalité que le bien-être de ces individus en puissance, qui viendront à exister dans le futur, est aussi important que le nôtre. Elle reconnaît la valeur potentielle des générations futures et ne peut donc transiger sur l'importance de considérer leur nombre et leur qualité de vie dans le processus de décision.

En accord avec un principe d'égalité des individus, le simple fait de ne pas encore exister n'enlève rien et ne nous donne pas de privilèges particuliers. Les générations futures sont en effet bien incapables de descendre dans la rue pour manifester contre des actions qui leur nuiraient, incapables de s'exprimer dans des journaux ou à la radio et incapables de monter un parti politique qui défendrait leurs intérêts. La position totale prend le parti de ces individus vulnérables et se propose de leur donner une voix pour les sortir de la précarité de leur situation.

## 2. Le fonctionnement de la position totale

### i. Arithmétique

La position totale a une arithmétique extrêmement simple :

$Valeur_{PT} = N \times Q$  où N est le nombre d'individus et Q leur qualité de vie moyenne<sup>36</sup>

On observe que via cette fonction additive, un même résultat peut être obtenu en respectant certaines proportions pour les variables N et Q. Par exemple, la valeur  $V_{PT}$  est identique si 100 personnes ont une qualité de vie Q de 10 ou si 200 personnes ont une qualité de vie Q de 5. Dans les deux cas la valeur  $V_{PT}$  vaut 1000. En revanche, à bien-être constant,  $V_{PT}$  devient directement fonction du nombre d'individus de la population, c'est-à-dire que plus la population est grande et plus le monde est désirable puisque sa valeur augmente en proportion. Symétriquement, dans une population qui ne fluctue pas il est toujours préférable d'avoir un bien-être plus élevé que plus bas.

---

<sup>36</sup> Richard Chappell, Daniel Meissner et William MacAskill, «Population Ethics, An Introduction to Utilitarianism», 2022.

## ii. Géométrie

Les différents scénarios peuvent être illustrés sous forme géométrique, notamment via des graphiques en deux dimensions où l'abscisse détermine N et l'ordonnée représente Q. La valeur totale est alors le produit des deux variables, c'est-à-dire l'aire des rectangles ainsi formés. Toutefois, une population n'est pas nécessairement homogène, au contraire, elle est en règle générale composée d'individus ayant des qualités de vie bien différentes. Ainsi, nous pouvons représenter un monde composé de différentes populations (ou une population composée de différents individus) en séparant les populations selon des critères définis.

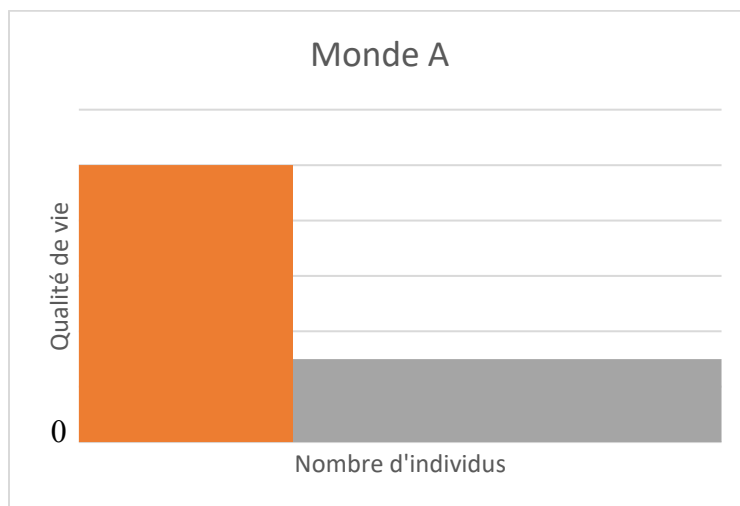


Figure 1. – Monde A

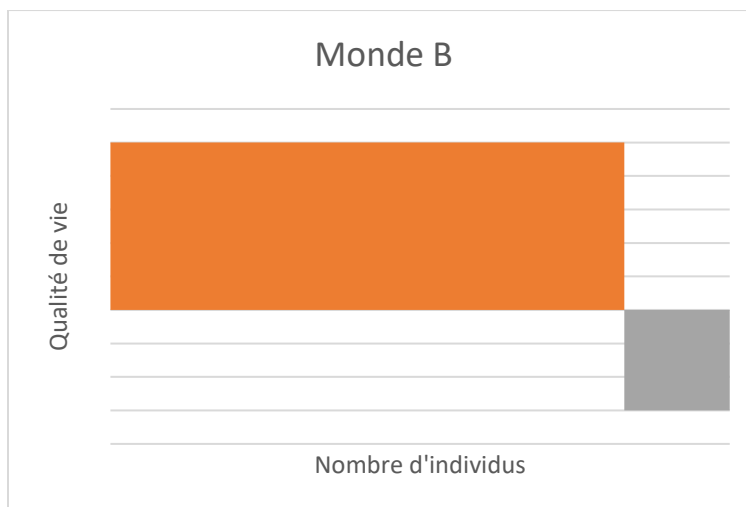


Figure 2. – Monde B

Dans le scénario du Monde A, le monde étudié est composé de deux populations. On observe d'une part que la population orange compte moins d'individus que la population grise, mais également que la qualité de vie moyenne de ses membres est significativement supérieure. L'unité de mesure étant la même, il est possible d'attribuer une valeur aux deux sous-populations, soit pour les comparer entre elles, soit pour calculer le bien-être total du monde A qui n'est que la somme des bien-être totaux des deux sous-populations. Dans cet exemple, les qualités de vie moyennes des populations qui habitent le monde A sont toutes positives, mais le phénomène est contingent.

Dans le scénario du monde B, on observe d'une part que la population minoritaire est aussi celle qui a la qualité de vie la moins élevée, et d'autre part que cette qualité de vie est strictement négative, ce qui veut dire qu'en moyenne, la vie de ces individus est misérable, elle ne vaut pas la peine d'être vécue car elle contient beaucoup plus de souffrances que de joie<sup>37</sup>.

<sup>37</sup> Ces deux scénarios proviennent à leurs tours de Derek Parfit, «Reasons and Persons», 1986, p. 388.

### 3. Les conséquences de la position totale : le long-termisme

La position totale défend donc des politiques qui maximisent le bien-être général absolu de l'ensemble de la population à l'étude. Or, une question demeure : si la saga de l'humanité était un roman, n'en serions-nous pas en réalité qu'à sa toute première page ?<sup>38</sup>

Comme nous l'avons mentionné, le futur est incroyablement vaste, et de cette constatation est née une position habituellement désignée par le terme long-termisme. Bien que le terme soit récent et ait été inventé par les philosophes MacAskill et Ord vers 2017<sup>39</sup>, tous deux figures majeures du mouvement Effective Altruism, la discussion autour de l'impact de nos actions sur l'avenir existe depuis bien longtemps. Cela étant, le long-termisme met l'accent sur l'importance immense de donner la priorité au bien-être et aux intérêts des générations futures, quand bien même cela impliquerait de faire un certain nombre de sacrifices dans le présent pour garantir ou à tout le moins pour augmenter la probabilité d'un futur radieux. Dans son livre *What We Owe the Future*, MacAskill la désigne comme « l'opinion selon laquelle influencer positivement l'avenir à long terme est la priorité morale clé de notre temps »<sup>40</sup>. Le présent forme le futur, et nos lignes politiques sont indubitablement amenées à avoir un impact significatif et réel sur la vie des générations futures. Ces considérations peuvent ainsi être appliquées à de nombreux domaines comme la politique, la technologie, l'environnement ou l'économie. À titre d'exemple, en ce qui concerne la question de la crise du climat et de la biodiversité, soutenir une position long-termiste implique de souhaiter sinon trouver des solutions techniques au problème, à minima limiter les émissions de gaz à effet de serre afin d'éviter aux générations futures de se retrouver confrontées à des conséquences catastrophiques qui menaceraient leurs survies et leurs développements de manière décisive. Dans cette optique, la priorité semble devoir être donnée au perfectionnement de l'avenir, et ce notamment par les avancées scientifiques et techniques qui permettront à l'humanité de réaliser son potentiel immense dans le respect des limites imposées par la physique. Or ces limites sont suffisamment larges pour permettre une

---

<sup>38</sup> Gaëtan Selle, The Flares, « La chronologie du FUTUR LOINTAIN : repenser les horizons temporels », 2023, 0:05 / 10:50.

<sup>39</sup> Wikipédia, entrée « Long-termisme », dernière révision 5 juillet 2023.

<sup>40</sup> William MacAskill, «What We Owe the Future», 2022, p. 12.

transformation radicale et encore insoupçonnée des vies concernées, ce qui explique pourquoi toute déviation de cette conduite est jugée moralement répréhensible dès lors qu'elles conduiraient éventuellement à l'une des pires catastrophes morales que l'humanité pourrait être amenée à connaître aux vues des gains incommensurables en jeu.

Cependant, il convient de noter que le mot Humanité a dans la bouche des long-termistes un sens un peu différent de celui que nous lui connaissons habituellement. L'humanité ce n'est pas ici le simple ensemble d'*Homo sapiens* qui forme la communauté humaine actuelle. En effet, de la même manière qu'*Homo sapiens* est le descendant d'une lignée d'ancêtres appartenant à des espèces différentes, les générations futures pourraient elles-mêmes ne pas se considérer comme des êtres humains et ne pas être biologiquement des *Homo sapiens*. Que ce soit via l'ingénierie génétique, le développement de la vie artificielle ou simplement par les mutations aléatoires naturelles qui forcent l'évolution darwinienne, de nombreuses possibilités d'altérer notre descendance existent. Il serait donc plus juste de parler d'humanité dans ce contexte en l'utilisant comme un terme qui exprime « civilisation d'origine humaine ». Pour ce qui est des intelligences artificielles notamment, leurs existences, leurs capacités ainsi qu'une partie de leurs connaissances seraient le résultat de l'évolution des mathématiques et de l'informatique, qui sont elles-mêmes des disciplines que l'on peut qualifier d'extensions de la créativité et de l'ingéniosité humaine. Si la vie numérique est possible, de tels algorithmes devraient alors en tout état de cause être considérés comme des agents de la civilisation d'origine humaine. Et nous voulons bien évidemment inclure dans le terme humanité ces êtres en tant qu'ils sont eux aussi des agents moraux liés à nous par une relation causale, puisqu'ils n'existeraient pas si nous venions à disparaître prématurément. Dans cette optique, l'on considèrera que la responsabilité morale des humains actuels s'étend plus largement à tous les êtres qui pourraient évoluer à partir notre descendance.

Or cette civilisation possède ce que l'on pourrait appeler un potentiel latent. Pour mieux comprendre les attentes du long-termisme vis-à-vis de ce potentiel et appréhender toute la force des enjeux qu'il soulève, attardons-nous donc un instant sur trois dimensions par lesquelles il

peut se définir. Selon Gaëtan Selle<sup>41</sup>, on parle ici de la longévité de la communauté à l'étude, du nombre d'individus en jeu ainsi que de la qualité de vie de ses individus. Rappelons d'ailleurs que la question de l'identité dont nous avons parlé plus tôt n'a pas sa place ici dans la mesure où la position totale, et par extension le long-termisme, adopte une position impersonnelle qui présuppose que les identités ne sont pas pertinentes et présument que les individus sont remplaçables les uns par les autres. Nous reviendrons plus tard dans ce chapitre sur ce concept lorsque nous ferons le tour des arguments qui supportent les positions non impersonnelles (*person affecting views*).

Poursuivons à présent en détaillant le rôle des trois autres dimensions dans la valorisation du potentiel de l'humanité.

## i. Longévité

Selon les dernières études en date, le consensus scientifique planche pour l'évolution d'*Homo sapiens* en Afrique il y a environ 300 000 ans<sup>42</sup>. Or, à en croire les estimations de la durée de vie moyenne des espèces mammaliennes, celles-ci vont de 0,6 million d'années<sup>43</sup> à 1,7 million d'années<sup>44</sup>. *Homo sapiens* est une espèce technologique qui se sert de son intelligence pour comprendre et modifier le monde à grande échelle et selon ses désirs. Par exemple, nous avons été la seule espèce capable de traverser les mers et les océans grâce aux outils que nous inventons plutôt que d'attendre des millions d'années que la lente évolution par voie de sélection naturelle ne nous équipe de nageoires et de branchies<sup>45</sup>. L'arrivée d'*Homo sapiens* en Australie par canoë il y a quelques dizaines de milliers d'années fait office d'exception parmi les grands animaux terrestres dans la très longue histoire de la vie sur Terre<sup>46</sup>. Le 21 juillet 1969 marque également le début d'une nouvelle ère pour notre civilisation, puisque nous sommes devenus, avec nos

---

<sup>41</sup> Gaëtan Selle, The Flares, « La chronologie du FUTUR LOINTAIN : repenser les horizons temporels », 2023, 1:40 / 10:50.

<sup>42</sup> Michael Greshko, « These Early Humans Lived 300,000 Years Ago—But Had Modern Faces », 2017.

<sup>43</sup> Barno Barnosky, « Has the Earth's sixth mass extinction already arrived? », 2011, p. 51–57.

<sup>44</sup> Mike Foote et David Raup, « Fossil Preservation and the Stratigraphic Ranges of Taxa », 1996, p. 128.

<sup>45</sup> Yuval Noah Harari, « Sapiens : A Brief History of Humankind », 2011, p. 71.

<sup>46</sup> « First Humans : Homo Sapiens & Early Human Migration (Article) | Khan Academy », consulté le 22 juillet 2023.



symbiotes, la première et unique espèce vivante à fouler le sol d'un autre astre. Il semble dès lors que l'humanité, forte de l'accumulation et de la transmission des savoirs au fil du temps, est une espèce qui ne cesse de développer sa capacité d'adaptabilité. Nous avons accumulé une telle avance sur le reste du vivant et un tel avantage stratégique et compétitif qu'il semble aujourd'hui envisageable que notre civilisation pourrait survivre sur Terre tant et aussi longtemps que celle-ci demeure habitable.

Autrement dit, sans même envisager les scénarios les plus optimistes de la conquête spatiale, et sans assigner de probabilités aux différents cas de figure, il est clairement concevable que notre civilisation perdure très longtemps. Aussi étonnant que cela puisse paraître, cela nous conduit donc à penser que nous ne sommes pas au début de l'Histoire, mais véritablement à ses balbutiements, c'est-à-dire à un stade si précoce que l'on pourrait le qualifier d'embryonnaire sans faire d'euphémisme. Or, notre futur est en réalité certainement moins déterminé par les limites de la physique que par nos actions, et il se pourrait bien que nous vivions une époque critique de la civilisation dans le sens où certains risques existentiels auxquels nous sommes confrontés peuvent entraîner, par définition, des conséquences de la plus haute importance sur l'ensemble de notre histoire, puisqu'ils ont la capacité de verrouiller une partie significative du potentiel de l'humanité prise dans son sens le plus large.

## ii. Population

Corollaire flagrant de la longévité attendue est le nombre faramineux de générations futures qui vont être amenées à l'existence. Alors que les estimations dépassent notre entendement, c'est toutefois la deuxième dimension à prendre en compte lorsque nous nous intéressons au potentiel de la civilisation d'origine humaine. Par déduction, la quasi-totalité de tous les humains de l'entièreté de l'histoire n'est pas encore née, et la plupart d'entre eux ne verront le jour que dans des temporalités très éloignées. Dès lors que l'on envisage que la vaste majorité des humains ayant foulé notre planète est une fraction minuscule de la part qui reste à naître, on prend du même coup conscience que ces milliards d'humains, à condition qu'ils aient à disposition temps, ressources, sagesse et expérience, pourraient coopérer ensemble pour

accomplir prouesses et prodiges aujourd'hui inimaginables. Moins de sept générations se sont succédées depuis la Révolution Industrielle, mais observons seulement comment l'humanité s'est métamorphosée ces deux derniers siècles, ouvrant la porte à plus encore de questions, de possibilités, découvrant toujours davantage de phénomènes naturels et institutionalisant de nouveaux champs d'études. Si l'accumulation des connaissances a permis à l'homme de dominer le monde<sup>47</sup>, comment appréhender la manière dont nos descendants vont comprendre les défis auxquels ils sont confrontés ? Selon cette position, l'importance de prendre des décisions non seulement éclairées mais également responsables devient on ne peut plus claire ; le présent semble avoir pour lui la responsabilité de l'avenir, de sorte que si nous avons de bonnes raisons de croire que nos décisions peuvent l'impacter, alors il semble impératif de créer des conditions favorables pour actualiser petit à petit cet avenir prospère. Ces considérations s'appuient évidemment sur un parti pris qui veut que l'épanouissement soit moralement très important dans l'univers, et il semble que le long-termisme tienne pour acquise une vision dans laquelle un univers rempli de descendants heureux est hautement désirable, ce qui n'est pas une évidence et sera discuté dans la prochaine section consacrée à l'utilitarisme moyen.

La longévité de la civilisation et le nombre de descendants potentiels nous empêchent donc d'écarter l'hypothèse d'un futur qui s'échelonne extrêmement loin dans le temps, et nous demande par là même de nous positionner vis-à-vis des implications morales que cela soulève. Face à cette constatation, la position totale est claire : puisqu'il existe des scénarios dans lesquels  $N$  tend vers un nombre extrêmement grand,  $V_{PT}$  tend donc elle aussi vers un nombre très grand et la bonne chose à faire serait de prendre les mesures plus adéquates et rationnelles possibles pour réaliser ce scénario très désirable. Mais la désirabilité de ce dernier possède une condition nécessaire, celle de la qualité de vie des individus concernés. Il convient donc de développer ce point pour mieux en saisir les subtilités.

---

<sup>47</sup> Yuval Noah Harari, «Sapiens : A Brief History of Humankind», 2011, p. 280.

### iii. Qualité de vie

La désirabilité de la continuité de la civilisation d'origine humaine sur de vastes échelles de temps suppose avant tout que la vie des individus concernés soit digne d'être vécue. Selon la position totale, l'atteinte de ce seuil se manifeste par la valeur 0. Lorsque la qualité de vie  $Q > 0$ , alors un individu, et par extension la moyenne d'une population, possède une vie dont on doit considérer qu'elle vaut la peine d'être vécue.

Or le monde contient des maux qui semblent presque infinis, à l'instar de la pauvreté, de la maladie ou de la douleur, maux que l'on peut concevoir comme autant de facteurs qui amoindrissent la qualité de vie des individus et des populations entières. Et il semble que l'on puisse envisager, au moins théoriquement, que nombre de ces expériences considérées négatives pourraient être amenées à disparaître sur des échelles de temps considérables<sup>48</sup> comme le souligne David Pearce lorsqu'il affirme que l'humanité est aujourd'hui « la seule espèce capable intellectuellement de reprogrammer la biosphère et créer un avenir glorieux d'une vie basée sur des gradients de bonheur »<sup>49</sup>. L'augmentation de la qualité de vie semble être un espoir partagé et est sans aucun doute la direction vers laquelle notre civilisation cherche à se diriger.

Cela peut se concrétiser via la dispense d'une meilleure éducation. Il est également envisageable de démocratiser la prévention des problèmes de santé et de systématiser de meilleures pratiques de prise en charge et d'accompagnement des souffrants. En réalité, tout reste à faire dans le domaine de l'amélioration des processus de soins. Plus évidente encore est l'amélioration que de nouveaux médicaments et des technologies médicales plus avancées pourraient apporter, allant jusqu'aux injections de nanorobots, à un Internet des Objets (*IoT*) plus développé voire à la reprogrammation génétique des niveaux de douleurs<sup>50</sup>. En d'autres termes,

---

<sup>48</sup> Voir le projet abolitionniste [dans le sens abolir la souffrance des êtres sentients] de David Pearce, une entreprise qu'il conçoit comme un impératif moral « sur 10 000 ans ».

<sup>49</sup> Gaëtan Selle, The Flares, « Augmenter le bien-être via les biotechnologies avec David Pearce », 2022, 0:01 / 1:50:48.

<sup>50</sup> Selon Pearce dans une entrevue donnée pour le podcast La Prospective, il existerait des centaines de gènes impliqués dans la souffrance, mais un gène en particulier semble crucial dans le processus, un gène qu'il appelle le fameux « bouton de volume de la douleur », le SCN9A qui a des dizaines d'allèles. Pearce propose donc en première instance de répandre les allèles bénins de ce gène dans la biosphère et de modifier peu à peu l'environnement par forçage génétique. Gaëtan Selle, The Flares, « Augmenter le bien-être via les biotechnologies avec David Pearce », 2022, 29:37 / 1:50:48.

les vies qui seront petit à petit amenées à être exemptées d'une part toujours plus importante des souffrances et des tragédies qu'un humain contemporain peu expérimenté seraient par définition des vies dont le facteur Q serait largement apprécié. Plus encore, les avancées en matière d'amélioration de la qualité de vie ne font en réalité que bonifier la borne inférieure de l'échelle qualitative des expériences. Si l'on considère que les individus ne peuvent connaître des expériences que dans une gamme bien particulière, il est aujourd'hui difficile de quantifier la borne supérieure potentielle qui limitera la qualité de vie des descendants. Les sentiments de joie, d'accomplissement et de félicité pourraient en tout état de cause être multipliés par des ordres de grandeur, comme le conçoit Pearce lorsqu'il précise le concept « d'espoir existentiel »<sup>51</sup> qui s'oppose à la notion de risque existentiel. Dans la mesure où nous sommes très loin de pouvoir imaginer jusqu'où l'expérience phénoménale des êtres sentients pourrait tendre, que ce soit vers le sublime ou l'horrible, il semble donc crucial selon la position totale d'empêcher les pires scénarios qui conduiraient à des vies abominables<sup>52</sup>.

Sur la base de l'étude de ces trois dimensions, on pourrait décrire formellement la position long-termiste par l'argument suivant :

---

<sup>51</sup> «Which reflect all of our dreams, aspirations, and wishes for a better future» dans Sam Barker, «Sam Barker and David Pearce on Art, Paradise Engineering, and Existential Hope (With Guest Mix)», 2020. On retrouve également cette notion plus tôt dans la littérature dans Owen Cotton-Barratt et Toby Ord, « Existential Risk and Existential Hope : Definitions », 2015.

<sup>52</sup> Il existe aussi un argument dit de « l'importance cosmique » que nous n'aborderons pas ici mais qui peut être mentionné. Selon l'argument, si nous sommes l'unique espèce douée d'un niveau de conscience élevé dans tout le cosmos, alors notre extinction verrait la fin de l'une des parties les plus rares et les plus précieuses de l'univers dans la mesure où nos expériences permettent littéralement à l'univers de s'observer et de se connaître. Or, dans l'éventualité où la Terre serait le seul endroit abritant autant de complexité, où des espèces pensent, ressentent et aiment, l'humanité serait la seule espèce capable de protéger la vie des catastrophes naturelles, et éventuellement de la faire prospérer au-delà de notre planète. Les catastrophes existentielles ne menacent en effet pas uniquement l'humanité mais également la faune et la flore. De ce point de vue, notre importance pourrait être cosmique. Gaëtan Selle, The Flares, « L'éthique du FUTUR LOINTAIN : les implications morales de nos décisions ? », 2023, 6:20 / 10:16.

P1. Tous les humains comptent de manière égale.

P2. Le futur est vaste.

P3. Tous les humains nés dans le passé ne représentent qu'une petite fraction de la civilisation d'origine humaine dans son ensemble.

P4. La qualité de vie des populations peut augmenter significativement.

---

C. Le potentiel de l'humanité est incommensurable et nous devons considérer l'immense perte qui pourrait résulter si nos actions réduisaient drastiquement la taille et la qualité de l'avenir.

Même si l'on s'abstient de toute spéculation futuriste et des considérations les plus extrêmes telles que le fait que l'humanité ensemence le cosmos ou que la conscience artificielle sera un jour atteinte, il n'en demeure pas moins que nous avons devant nous des millions d'années et des milliers de milliards de descendants à naître. En ce sens, il convient d'apprécier l'intuition selon laquelle le futur est précieux et pourquoi certaines positions de l'utilitarisme total proposent d'incorporer cette valeur dans leur axiologie. La question de la possibilité d'un large potentiel pose plusieurs problèmes et la position totale, tout comme le long-termisme, est sujette à de nombreuses critiques théoriques. Avant d'entamer le prochain chapitre, qui aura pour objet de détailler ces critiques traditionnelles et de mettre en lumière les réponses que la littérature leur a apportées, intéressons-nous justement à ces autres avenues et commençons par expliquer certains détails à propos de la plus célèbre d'entre elles, la position moyenne.

## **B. La position moyenne**

### **1. Des intuitions divergentes**

La position moyenne est une autre doctrine utilitariste qui cherche à répondre aux défis soulevés par l'éthique des populations. Cette dernière soutient avant tout que la désirabilité de

la distribution d'une population est strictement proportionnelle au bien-être *moyen* qu'elle contient<sup>53</sup>.

La fonction de valeur utilisée est donc simplement<sup>54</sup> :  $Valeur_{PM} = Q$

Cette manière de compter a plusieurs implications. D'une part, et contre ce qui fait la substance des propos précédents, la position moyenne considère que le fait de créer des vies supplémentaires n'est pas particulièrement désirable, à moins que ces dernières ne participent à augmenter la qualité de vie moyenne de la population. De même que pour la position totale, les individus sont ici traités comme des entités interchangeables possédant une valeur strictement identique, de sorte que les intérêts et le bien-être de chaque individu doivent être pris en compte.

S'il est possible de lui ajouter des considérations plus égalitariennes, par exemple dans le cas où deux populations ont le même bien-être moyen, le cadre original de la position moyenne traite exclusivement de la manière de comptabiliser l'utilité d'une population afin de proposer des comparaisons et des prescriptions fondées sur une base rationnelle<sup>55</sup>.

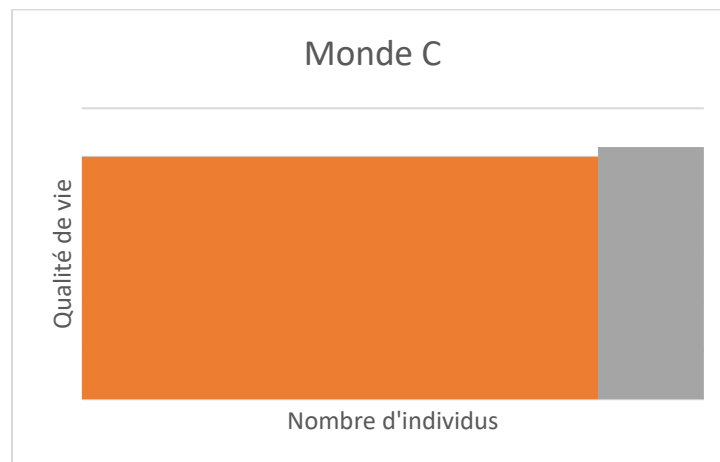


Figure 3. – Monde C

<sup>53</sup> Thomas Hurka, «Average Utilitarianisms», 1982, p. 65-69.

<sup>54</sup> Nous aborderons des variants à cette fonction au Chapitre III.

<sup>55</sup> Pour Grill « A high average is of course consistent with a very unequal distribution, and equality of well-being may matter morally, but the Total and the Average View are equally insensitive to equality ». Kalle Grill «The Sum of Averages: An Egyptology-Proof Average View», 2023, p. 104. En effet, les deux positions peuvent devenir des prioritarismes si elles cherchent à accommoder les enjeux égalitaires. Nous n'aborderons toutefois pas ces notions ici.

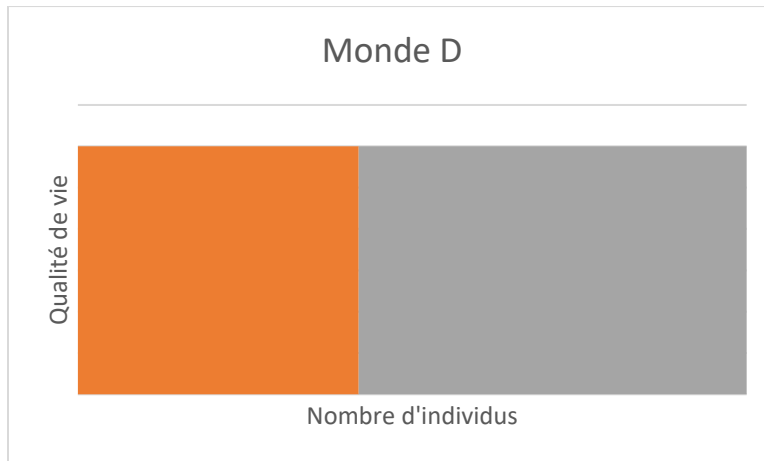


Figure 4. – Monde D

Dans le Monde C, la taille de population en gris a beau être plus petite, sa qualité de vie est en moyenne supérieure à celle de la population en orange, et sera donc considérée comme ayant une plus grande valeur. Pour ce qui est du Monde D, la position moyenne soutient simplement que les deux populations sont aussi bien loties ; elle ne se prononce pas sur la différence de taille de la population. Cette approche permet ainsi de comparer les mondes C et D sans se préoccuper de l'axe des abscisses, et il semble ici que C soit plus désirable que D.

Étant donné sa nature, la position moyenne a une approche un peu différente de la gestion des risques, mais les résultats qu'elle nous apporte sont dans de nombreux cas cohérents avec nos intuitions communes ainsi qu'avec les résultats de la position totale. En effet, tenons pour acquis que nous ayons de bonnes raisons de croire que nos descendants seront légion et que leur existence et leur bien-être dépendra en partie des choix que nous faisons, et la position moyenne arrivera à la conclusion que nous pouvons nous attendre à de larges gains en termes de bien-être moyen si tant est que des ressources suffisantes aient été investies dans la prévention et la mitigation des risques existentiels et que des politiques ambitieuses ont été mises en œuvre dans les délais impartis.

## i. L'échelle individuelle

Contrairement à la position totale, la quantité totale de bien-être n'a pas ici d'importance à l'échelle de populations mais uniquement à l'échelle individuelle. Davantage de vies heureuses n'est donc pas mieux que moins de vies heureuses, même s'il est largement plus désirable d'avoir une vie heureuse plus longue qu'une vie heureuse plus courte. Boonin-Vail l'exprime en ces termes « [we should aim] to produce more happiness for people, not to produce more people for happiness »<sup>56</sup>. Selon cette approche, le bonheur ne peut pas être considéré comme le sujet et il n'a certainement pas une valeur en soi. Au contraire, ce qui a de la valeur ce sont les individus, les patients moraux ; c'est donc sur eux que nous devons nous concentrer. Selon certains partisans de la position moyenne, la position totale a fait, entre autres choses, l'erreur de mettre sur un piédestal le bonheur au détriment des populations existantes et du bonheur individuel. La question qui se pose naturellement est donc la suivante. Si l'important est le bonheur individuel, que penser des fluctuations de populations ?

## ii. Ajouter de nouveaux individus : l'intuition de neutralité

Dès les années 1960, Jan Narveson parlait de l'intuition de neutralité de la manière suivante. « [T]here is no moral argument at issue here. How large a population you like is purely a matter of taste, except in cases where a larger population would, due to indirect effects, be happier than the first »<sup>57</sup>. Utiliser la taille de la population comme critère fondamental serait donc inexact dès lors que l'on comprend que cette considération est avant tout une question de goût. Autrement dit, le paramètre N devrait être considéré éthiquement neutre<sup>58</sup>. Si ce point de vue

---

<sup>56</sup> David Boonin-Vail, «Don't Stop Thinking About Tomorrow: Two Paradoxes about Duties to Future Generations», 1996, p. 268.

<sup>57</sup> Jan Narveson, «Utilitarianism and New Generations», 1967, p. 68.

<sup>58</sup> À ce propos, une certaine asymétrie peut toutefois être envisagée, qui a été largement défendue dans la littérature : l'intuition de neutralité ne s'appliquerait pas aux personnes dont la qualité de vie est strictement négative. Il ne serait donc pas neutre, mais bien délétère d'augmenter la taille d'une population malheureuse, ce qui est d'ailleurs la conviction de Narveson (Jan Narveson, «Utilitarianism and New Generations», 1967, p. 62-72). De même «Consider the view that, while the fact that a person's life would be worse than no life at all (or 'worth not living') constitutes a strong moral reason for not bringing him into existence, the fact that a person's life would



s'avère correct, il s'ensuit qu'un plus grand nombre de vies bonnes ne rendent pas le monde meilleur ni même plus désirable, en conséquence de quoi l'extinction de l'humanité n'a rien de particulièrement problématique en soi. Nous n'avons donc aucun impératif à coloniser l'espace ou même à investir massivement dans la défense contre un quelconque risque existentiel.

Pour Pressman cependant, bien que l'intuition de neutralité semble attirante, elle est en réalité incohérente et contradictoire. Sa démonstration s'appuie sur la situation suivante<sup>59</sup> :

Voici trois mondes A, B et C, dans lesquels vivent des individus dont les niveaux de bien-être sont les valeurs entre parenthèses. Le X représente le fait qu'aucun individu n'existe à cet emplacement.

$A = (1, 1, X)$	$B = (1, 1, 1)$	$C = (1, 1, 2)$
-----------------	-----------------	-----------------

Tableau 1. – Distributions pour trois populations

Selon l'intuition de neutralité, les mondes A et B ont la même valeur, de même que les mondes A et C. Toutefois, il est évident que C est plus désirable que B. Il s'ensuit une contradiction qui peut être résolue par l'introduction d'une notion conditionnelle voulant que la valeur du bien-être de la troisième personne dépende de son existence. Alors seulement la contradiction disparaîtrait puisque le monde A ne peut plus être comparé aux mondes B ou C. Bien que cette réponse permette peut-être d'éviter l'écueil théorique, elle rend la notion de même d'intuition de neutralité inutile dans la mesure où le propre de l'éthique des populations est de comparer des mondes dans lesquels la taille des populations varie largement.

Pour Pressman, cela s'explique par le fait que l'intuition qui voulait servir de fondation à la position moyenne est en réalité une composante des positions non impersonnelles, lesquelles sont très rarement défendues dans la littérature pour aborder les problèmes d'éthique des populations, comme nous le verrons dans la dernière section de ce chapitre qui leur sera dédiée.

---

be worth living provides no (or only a relatively weak) moral reason for bringing him into existence. This view which I shall refer to as 'the Asymmetry' is approved both by Narveson and by common sense» (Jefferson McMahan, «Problems of Population Theory», 1981, p. 100). Dans ce cas, la porte s'ouvre à d'autres intuitions comme celle de John Broome qui soutiendra en 2004 qu'il est également commun de trouver que si la vie d'un individu potentiel est suffisamment bonne, alors la créer n'est pas neutre, mais bien désirable (John Broome, «Weighing Lives», 2004).

<sup>59</sup> Michael Pressman, «A Defence of Average Utilitarianism», 2015, p. 394.

L'intuition de neutralité semble donc fautive, mais l'auteur n'y voit pas là une raison de remettre en cause son axiologie. En effet, tout en préservant l'intuition que le bonheur n'a de valeur que conditionnellement à l'existence d'un individu (laquelle ne *s'appuie pas* sur les positions non impersonnelles pour exister<sup>60</sup>), la position moyenne est de son côté capable de faire une distinction crédible entre A, B et C puisque son arithmétique lui permet de comparer une valeur commune aux trois mondes, la moyenne.

Avant d'approfondir les critiques soulevées par les théoriciens contre ces deux positions, nous ferons maintenant un court état des lieux de théories moins populaires mais qui participent aujourd'hui à l'élargissement des idées de l'éthique des populations, en commençant par présenter une position hybride qui cherche à mélanger les intuitions différentes sous-tendant les deux avenues discutées jusqu'alors. Ces présentations sont nécessaires pour comprendre les solutions proposées dans les Chapitres II et III.

## **C. Nouvelles approches des problèmes d'éthique des populations**

### **1. Les théories à valeur variable**

Alors que la position totale repose sur l'intuition selon laquelle le bonheur doit être accumulé de manière absolue, sa rivale propose une alternative consistant à ne s'intéresser qu'à la qualité de vie moyenne de la population. En effet, l'idée d'un impératif de multiplication effréné peut déranger. Certes, il faudrait conserver ou améliorer la qualité de vie des descendants, mais ce qui compte avant tout c'est la qualité, et non la quantité. Toutefois, une intuition extrêmement courante est celle qu'à qualité de vie similaire, un monde rempli d'humains est certainement préférable, de sorte qu'il paraît nécessaire d'accorder un certain poids à la quantité de personnes en jeu. Et il n'est pas vain de chercher à concilier les deux intuitions. Nous allons donc présenter

---

<sup>60</sup> Michael Pressman, «A Defence of Average Utilitarianism», 2015, p. 395.

cette nouvelle position comme une tentative de produire un compromis entre les deux avenues discutées précédemment.

Proposées en 1983 par Thomas Hurka<sup>61</sup>, les théories à valeur variable prennent en compte la qualité de vie totale en plus de la qualité de vie moyenne pour comparer la désirabilité des distributions de populations. L'idée derrière ces positions est que la valeur d'une vie digne d'être vécue varie en fonction du nombre d'individus qui existent déjà, de manière que cette vie ait plus de valeur lorsque la population est petite que lorsqu'elle est grande<sup>62</sup>. Pour ce faire, elles proposent de faire diminuer la valeur marginale des nouveaux individus, ce qui a pour conséquence mécanique que le résultat de la fonction ainsi produite se rapproche de la position totale lorsque la population est petite, attachant alors plus de valeur à la somme de bien-être collectif et de la position moyenne au fur et à mesure qu'elle augmente. Selon cette approche, la valeur des vies additionnelles décroît asymptotiquement, et l'on peut s'attendre à trouver une limite supérieure à la valeur totale qui peut exister au sein d'une population.

On peut traduire cette idée sous la forme suivante où la valeur devient le produit d'une variable et d'une fonction dans laquelle  $f$  est une fonction strictement croissante et strictement concave avec une asymptote horizontale. Autrement dit, à mesure que  $N$  augmente,  $f(N)$  augmente à un rythme décroissant, et ne dépasse jamais une certaine limite<sup>63</sup>.

$$\text{Valeur}_{TVV} = Q \times f(N)$$

---

<sup>61</sup> Thomas Hurka, «Value and Population Size», 1983, p. 496–507.

<sup>62</sup> «The repugnant conclusion (Stanford Encyclopedia of Philosophy)», 2017.

<sup>63</sup> Richard Chappell, Daniel Meissner et William MacAskill, «Population Ethics, An Introduction to Utilitarianism», 2022.

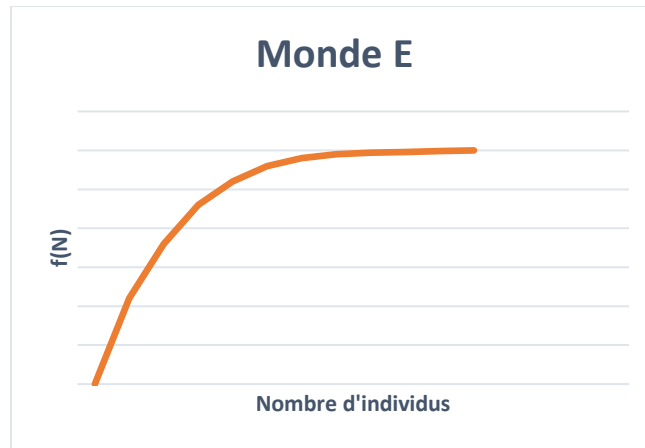


Figure 5. – Monde E

Si ces théories permettent d'éviter certains écueils de la position totale que nous verrons dans le prochain chapitre, il semble toutefois qu'elles ne parviennent pas à résoudre les problèmes que l'on retrouve dans la position moyenne<sup>64</sup> sans devenir incohérentes ni subir de contrecoups théoriques. En effet, la fonction d'utilité des théories variables montre des rendements marginaux décroissants, de sorte qu'elle a un comportement qui s'approche de la fonction moyenne lorsque la population est très large. Bossert, Cato et Kamaga se sont d'ailleurs intéressés à cette thématique et ont produit en 2022 un théorème qui prouve ce dernier point<sup>65</sup>. Le problème des théories variables est d'autant plus grave qu'il postule une diminution de la valeur marginale des vies uniquement lorsque celles-ci sont positives. Cette caractéristique entraîne un second problème, que l'on rencontre sous le nom de Conclusion Absurde<sup>66</sup>. Brièvement et à titre informatif, si l'on tient pour acquis que la valeur de chaque vie négative additionnelle ne diminue pas davantage en fonction de la taille de la population, contrairement à la valeur des vies positives, alors le simple fait d'agrandir une population en respectant les proportions peut transformer une population bonne en une population mauvaise. Si cette prémisse permet de formaliser l'idée selon laquelle plus de vies mauvaises est mauvais mais plus de vies bonnes n'est pas particulièrement bon, l'absurdité se manifeste on ne peut plus

<sup>64</sup> Voir Chapitre III, section Conclusion Sadique.

<sup>65</sup> Walter Bossert, Susumu Cato et Kohei Kamaga, «Revisiting variable-value population principles», 2022, p. 8-12. Voir le premier théorème « Let R be a generalized variable-value principle that satisfies the limit property. If R avoids the repugnant conclusion, then R implies the sadistic conclusion » et sa preuve dans les pages subséquentes.

<sup>66</sup> Derek Parfit, «Reasons and Persons», 1986, p. 391-418.

clairement via l'exemple suivant : une population  $P_1$  composée d'un million de vies A au bien-être très positif et d'une seule vie B au bien-être négatif semble intuitivement constitutive d'un monde désirable. Toutefois, puisqu'il y a une limite à la valorisation positive mais pas à la valorisation négative, une expansion proportionnelle, c'est-à-dire que  $\forall n$  où  $n \in \mathbb{N}$ , la suite  $P_n = nA + nB$  se dirige vers un état de plus en plus indésirable, ce qui semble faux puisqu'une population  $P_{10}$  composée de 10 millions de vies positives et de dix vies négatives n'est pas largement pire que  $P_1$ .

Les théories à valeur variable font donc face à plusieurs difficultés, mais le débat se poursuit et la question reste ouverte de savoir si elles peuvent être corrigées pour compenser de manière plus satisfaisante l'introduction de vies au bien-être négatif<sup>67</sup>.

## 2. Les théories de seuil critique et de gamme critique

### i. Les théories de seuil critique

Les théories de seuil critique sont des variations de la position totale. À ce titre, elles préféreront les distributions qui maximiseront le bien-être collectif total. L'idée ici est d'ajouter une contrainte à la position totale afin de ne compter que le bien-être positif qui se trouve au-dessus d'un certain seuil critique dénoté par la valeur  $\alpha$ .

$$\text{Valeur}_{TSC} = N \times (Q - \alpha)$$

L'implication principale de ces théories est le fait que créer de nouveaux individus n'est pas tout le temps désirable, puisque la valeur totale augmentera uniquement dans le cas où le terme de droite est positif, c'est-à-dire lorsqu'ils et elles possèdent une qualité de vie individuelle qui dépasse le seuil déterminé. De ce point de vue, la position totale devient donc un cas particulier de théorie de seuil critique, dans laquelle la valeur de  $\alpha$  est pile égale à zéro<sup>68</sup>, et alors le simple

---

<sup>67</sup> D'après Chappell et al (2023), une telle tentative a été proposée par Chappell, R.Y. dans «Parfit's Ethics», 2021, section 7.2.2.

<sup>68</sup> Gregory Kavka, «The Paradox of Future Individuals», 1982, p. 92-112.

fait d'avoir une qualité de vie positive suffit à participer à l'amélioration et à la désirabilité du monde. Mais l'on peut imaginer un seuil plus élevé, de sorte qu'il n'y ait virtuellement aucun doute sur le fait que l'ajout des vies ajoutées soit une bonne chose.

$$Valeur_{TSC} = N \times (Q - 0) = N \times Q = Valeur_{PT}$$

Si l'origine de ces théories réside dans le désir de pallier un défaut de la position totale, force est de constater qu'elles ont leurs propres problèmes, lesquels s'observent dans le scénario suivant<sup>69</sup>.

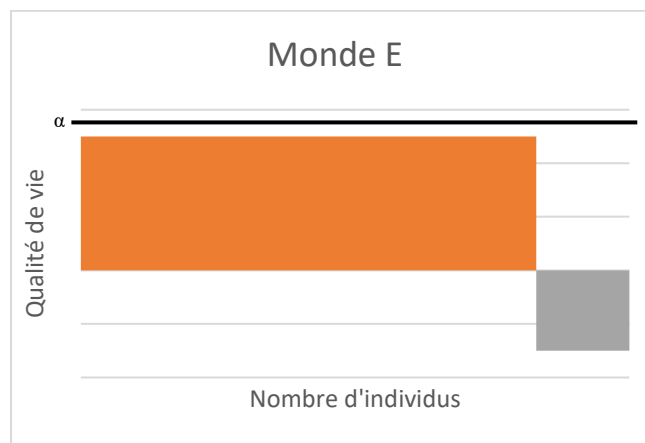


Figure 6. – Monde E et seuil critique  $\alpha$

On reconnaîtra ici que le Monde E n'est que le Monde B utilisé lors de la présentation de la position totale, auquel on a ajouté le facteur  $\alpha$ . Souvenons-nous que la position totale était en accord avec nos intuitions puisque dans ce scénario la distribution de la population orange était sans conteste plus désirable que celle de la population grise, du fait que l'aire qu'elle occupait au-dessus de zéro était plus grande. Ici, contre intuitivement, le fait que les populations soient toutes les deux sous le seuil  $\alpha$  signifie qu'elles sont toutes les deux indésirables. Or la population orange, dont le bien-être moyen est positif (et même plutôt proche du seuil) est bien plus grande que la population grise dont le bien-être moyen est négatif. Ce simple fait nous contraint à conclure que

<sup>69</sup> Richard Chappell, Daniel Meissner et William MacAskill, «Population Ethics, An Introduction to Utilitarianism», 2022.

la population la plus désirable des deux est la population grise, dans laquelle moins d'individus aux vies non désirables existent.

Ce résultat semble pourtant inacceptable. Heureusement, le problème peut être contourné via l'introduction d'une dimension supplémentaire à  $\alpha$ . Ce faisant,  $\alpha$  devient un intervalle et non plus un seuil, et la théorie de seuil critique peut évoluer en une théorie de gamme critique, aussi appelée intervalle critique, dont nous allons expliciter les implications.

## ii. Les théories de gamme critique

Les théories de gamme critique sont donc définies comme des positions selon lesquelles l'addition d'un individu est désirable si et seulement si sa qualité de vie excède la borne haute de la gamme critique considérée. À l'inverse, le résultat devient de moins en moins désirable à partir du moment où sa qualité de vie tombe sous la borne inférieure de cette gamme<sup>70</sup>. Les choses se compliquent lorsque le bien-être est compris dans les limites de la gamme, puisque les théories ne se prononcent pas sur la création de tels individus. Nous entendons par là qu'elles laissent le résultat indéterminé, de sorte que bien que la comparaison entre des vies ayant un bien-être différent au sein de la gamme reste possible, la Valeur<sub>TGC</sub> globale avant et après l'introduction de cette vie ne peut être dite avoir augmenté ou diminué<sup>71</sup>. En d'autres termes, pour être considérée bonne, une vie doit donc contenir significativement plus de bien-être qu'une vie mauvaise<sup>72</sup>, comme le présente le schéma ci-dessous adapté de Gustafsson (2020)<sup>73</sup>.

---

<sup>70</sup> Johan E. Guðafsson, «Population axiology and the possibility of a fourth category of absolute value», 2019, p. 86.

<sup>71</sup> Ruth Chang, «The possibility of parity», 2002, p. 659–688.

<sup>72</sup> La formule consacrée pour de tels cas est que ces vies sont « bof » (« meh lives »), ce qui reste un terme vague mais toutefois utile pour se représenter les cas qui ne sont ni bons ni mauvais mais qui n'ont pas une valeur nulle.

<sup>73</sup> Johan E. Guðafsson, «Population axiology and the possibility of a fourth category of absolute value», 2019, p. 92.

## Monde F

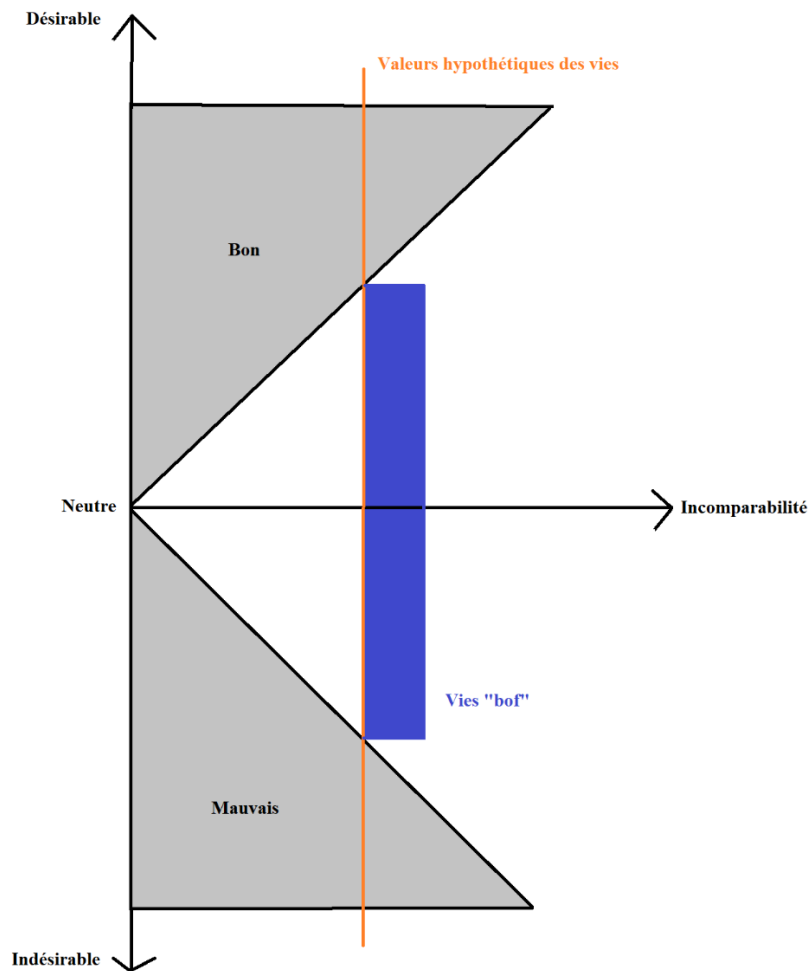


Figure 7. – Monde F

Des théories récentes et complexes comme les théories de gamme critique sont aujourd'hui à l'étude, qui méritent d'être analysées et débattues dans le cadre des études en éthique des populations. Si elles peuvent sembler prometteuses au lecteur, nous l'invitons à consulter Gustafsson (2020), *Population axiology and the possibility of a fourth category of absolute* pour plus d'informations.

Comme toutes celles que nous avons présentées jusqu'alors, les théories critiques sont impersonnelles. Il est toutefois possible d'envisager l'utilitarisme d'un point de vue plus personnel, via ce que nous appellerons ici la catégorie des théories « non impersonnelles ». Avant de nous attaquer plus en profondeur aux difficultés des positions totale et moyenne, arrêtons-



nous un instant pour prendre le temps de survoler cet autre type de théorie qui réajuste radicalement la manière de considérer l'utilité.

### 3. Les théories non impersonnelles (*person affecting views*)

#### i. Fonctionnement

Lorsque l'on se rappelle la fonction de la position totale, il peut être frappant de réaliser que rien ne précisait directement la nature de la valeur qui devrait être agrégée. Or, il est possible d'utiliser une notion du bien-être qui englobe le fait que l'on considère les individus pour ce qu'ils ont d'unique. Ce choix détermine si nous avons affaire à un utilitarisme impersonnel ou non impersonnel, et il devient alors nécessaire de préciser que la version développée dans la première section était partisane d'un utilitarisme impersonnel.

Les théories impersonnelles ont depuis longtemps été critiquées pour leur tendance à ne traiter les individus que comme des êtres interchangeables n'ayant d'une certaine manière pas de valeur intrinsèque. L'utilitarisme impersonnel aurait ce défaut qu'il ne souhaite pas que les individus aient une bonne qualité de vie *parce qu'*avoir une bonne qualité de vie est bon pour eux, mais davantage *parce que* cette caractéristique est bonne du point de vue du monde, c'est-à-dire parce que le résultat est impersonnellement meilleur<sup>74</sup>. L'idée derrière les positions non impersonnelles est de soutenir qu'un monde ne peut pas être meilleur qu'un autre à moins qu'il soit meilleur pour quelqu'un en particulier. Cette idée participe au rejet des conceptions selon lesquelles avoir davantage d'individus ayant une qualité de vie suffisamment bonne rendrait le monde meilleur. En effet, de ce point de vue les positions impersonnelles ne s'intéressent pas à proprement parler aux individus, et donc pas à ce qui est bon pour eux. Autrement dit, elles sont vues comme des positions qui ne s'occupent pas de ce qui a véritablement de la valeur<sup>75</sup>. En effet,

---

<sup>74</sup> Ralf Bader, «Person-affecting utilitarianism», 2022, p. 4.

<sup>75</sup> Des théories dualistes existent qui tentent de concilier les notions impersonnelles et non impersonnelles de bien à travers des fonctions de conversion.

les fondations de ces positions mènent ses partisans à devoir soutenir un principe d'incomparabilité existentielle qui postule que le fait qu'un individu particulier existe dans un monde mais pas dans un autre rend de fait impossible toute tentative de comparaison de son bien-être<sup>76</sup>. L'individualité est ici vue comme une caractéristique fondamentale. Elle est la troisième variable sans laquelle une axiologie solide est inenvisageable. On peut en illustrer les conséquences par le scénario du Monde A (Figure 1.), dans lequel les deux populations sont désormais incomparables, ce qui est révélateur d'une limitation de l'utilité de ces positions dans le cas présent.

Les théories non impersonnelles tentent de capturer l'intuition de neutralité dont nous avons parlé précédemment. Elles se heurtent donc aux critiques de Pressman (2015). Elles sont pourtant une avenue populaire souvent défendue pour des raisons d'intuition, possédant des avantages et des défis propres<sup>77</sup> qu'il serait trop ambitieux de chercher à résumer dans ce document. En effet, tandis que de nombreuses critiques existent dans la littérature, le but de notre entreprise ici est de comprendre la profondeur des lacunes de la position totale et de la position, et non d'opposer l'utilitarisme impersonnel et non-impersonnel.

## Conclusion

Tout le long de ce chapitre, nous avons cherché à expliquer les intuitions et le fonctionnement de plusieurs avenues en lien avec les positions totale et moyenne pour mieux être capable par la suite de comprendre la portée des critiques qui leurs sont faites. Alors que le débat a été traditionnellement dominé par la rivalité entre la position totale et la position moyenne, divers hybrides et évolutions ont émergé pour enrichir la discussion. Pourtant, aucune d'entre elles ne fait pleinement consensus parmi les philosophes contemporains. Dans le prochain chapitre, nous tâcherons d'expliquer plus en détail les problèmes que rencontre la position qui a eu le plus d'influence sur la discipline, et analyserons les diverses réponses que l'on

---

<sup>76</sup> David Heyd, «Procreation and Value: Can Ethics Deal With Futurity Problems?», 1988, p. 151–170.

<sup>77</sup> Dont le fameux « problème de non-identité » soulevé par Parfit dans son expérience de pensée « 14-years-old girl ». Voir Derek Parfit, «Reasons and Persons», 1986, p. 351-380.

peut trouver à ces critiques dans la littérature afin de mettre en évidence de manière accessible et synthétique les lacunes auxquelles elle fait face.

### III. Chapitre II – Critiques de la position totale

Comme nous l'avons vu précédemment, plusieurs intuitions mènent à vouloir maximiser le bien-être total d'une population plutôt que son bien-être moyen. Cependant, il semble que ce mode opératoire ne soit pas sans conséquences, de sorte que certains vont jusqu'à dire que l'entièreté du débat sur l'éthique des populations ne tourne en réalité qu'autour des termes de la résolution d'un problème particulier inhérent à cette position<sup>78</sup>. Ce problème est connu sous le nom de Conclusion Répugnante, et il sera le premier objet examiné dans ce chapitre. Nous recenserons alors certaines réponses plausibles qui ont été énumérées au fil du temps et mettrons en lumière leurs limites. Par la suite, nous aborderons une autre difficulté à laquelle fait face la position totale ; le conflit des temporalités. Cela nous permettra d'avoir une vue plus précise de la situation de la position totale avant d'aborder le cas de la position moyenne.

#### A. La Conclusion Répugnante

##### 1. La problématique

Parmi les critiques de la position totale, la plus importante et la plus largement discutée dans la littérature soutient que fonder une éthique des populations sur la somme total du bien-être entraînerait inmanquablement une « conclusion répugnante » que l'on ne saurait cautionner. Cette remarque, originellement soulevée par Derek Parfit dans son livre *Reasons and Persons*<sup>79</sup> (1984), a d'ailleurs fini par former aujourd'hui le cœur de la controverse et naturellement le principal argument de l'opposition. Considérons-là de la manière suivante :

---

<sup>78</sup> Michael Pressman, «A Defence of Average Utilitarianism», 2015, p. 389.

<sup>79</sup> «For any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better even though its members have lives that are barely worth living» Derek Parfit, «Reasons and Persons», 1984, p. 388.

P1. Le bien-être est quantifiable et varie entre les individus.

P2. Plus de bien-être individuel est plus désirable que moins de bien-être individuel.

P3. La valeur d'une vie équivaut à son niveau de bien-être net.

P4. La valeur d'un monde équivaut à la somme des valeurs des vies qui le peuplent.

---

C. Certains mondes dans lesquels la moyenne de bien-être est à peine positive sont plus désirables que certains mondes dans lesquels la moyenne de bien-être est extrêmement élevée.

Puisque dans le paradigme de la position totale les mondes sont comparés uniquement sur la base du produit des variables pertinentes ( $N \times Q$ ), la réduction de l'une de ces variables au profit de l'autre est une transformation qui sera considérée éthiquement neutre par le modèle tant et aussi longtemps que les proportions sont respectées. Autrement dit, dans une distribution donnée de population, une baisse de 50 % de  $N$  doit être compensée par une hausse de 50 % de  $Q$  pour que le résultat  $V_{PT}$  conserve exactement sa valeur. Or, ce fait qui paraît à première vue anodin est en fait lourd de conséquences, puisqu'il s'ensuit que nous devons admettre l'existence de mondes possiblement très désirables dans lesquels le bien-être moyen des individus est très bas, si bas qu'il est à peine positif, par exemple  $Q_1 = 0,000X$ . Si l'on admet que  $Q = 0$  est la limite à ne pas dépasser pour conserver une vie positive, alors un  $Q_1$  est une valeur qui touche tout juste le niveau où l'on considère que la vie vaut la peine d'être vécue. En effet, selon cette critique, la fonction utilisée nous pousse à admettre qu'un monde devient meilleur au fur et à mesure que sa population est démultipliée mais que la qualité de vie moyenne de ses habitants baisse. Si les premières itérations de ce type d'opération ne posent pas de problèmes, l'on peut rapidement arriver à des cas extrêmement contre-intuitifs. Comparons trois scénarios pour illustrer le propos.

Dans le premier ( $S_1$ ), 10 individus ont un bien-être moyen de 100. Dans le deuxième ( $S_2$ ) 1000 individus ont un bien-être moyen de 99. Il ne paraît pas incongru de dire que  $S_2$  peut être considéré bien meilleur que  $S_1$ , puisque leurs valeurs respectives  $V_{S_1}$  et  $V_{S_2}$  sont de 1000 et 99 000. Ici, la position totale déclare d'ailleurs que  $S_2$  est 99 fois meilleur que  $S_1$ .

Scénario 1	Scénario 2
$V_{S1} = 10 \times 100 = 10^3$	$V_{S2} = 10^3 \times 99 = 99 \times 10^3$

Tableau 2. – Comparaison des scénarios 1 et 2

En termes d'utilité totale  $S_2 > S_1$  semble plausible. Mais regardons ce qui se passe lorsqu'un troisième scénario ( $S_3$ ) prend place.  $S_3$  est un monde dans lequel un milliard d'individus existent, qui ont un bien-être moyen de 0,0001.

Scénario 1	Scénario 2	Scénario 3
$V_{S1} = 10 \times 100 = 10^3$	$V_{S2} = 10^3 \times 99 = 99 \times 10^3$	$V_{S3} = 10^9 \times 10^{-4} = 10^5$

Tableau 3. – Comparaison des scénarios 1, 2 et 3

En termes d'utilité totale,  $S_3 > S_2 > S_1$  et l'humanité semble avoir pour obligation morale de tout mettre en œuvre pour se prémunir des risques existentiels. L'itération de ce raisonnement est la clé de voûte de la Conclusion Répugnante. D'ailleurs  $V_{S3} = 100 \times V_{S1}$ , de sorte que  $S_3$  est 100 fois meilleur et plus désirable que  $S_1$ , ce qui peut indubitablement aller à l'encontre de nos intuitions les plus fondamentales.

Il peut sembler évident au lecteur que de tels mondes sont indésirables, car nous aurons naturellement tendance à valoriser davantage les mondes dans lesquels les individus sont très épanouis. À l'inverse, un monde M dans lequel  $Q_M$  est virtuellement identique à  $Q = 0$  équivaut intuitivement à un monde où l'existence ne vaut quasiment pas la peine d'être vécue et est de ce fait bien loin de l'utopie commune. Toutefois, il semble que ce soit bel et bien la conclusion de l'arithmétique utilisée par la position totale. Ces mondes très peuplés seraient fortement désirables dès l'instant où le nombre d'individus considérés est si large que la variable N compense strictement la baisse de bien-être Q si l'on considère le résultat du point de vue du niveau total de bien-être  $V_{PT}$ . Lorsqu'il s'agit de comparer des mondes, les plus désirables seront alors toujours ceux qui contiennent des populations largement plus importantes, constituées d'individus au bien-être strictement positif. Voyons plutôt cette situation sous forme géométrique.

$S_1$  pourrait être la représentation d'un scénario paradisiaque sur laquelle une communauté de dix individus vit en harmonie avec la nature et jouit de la vie, tandis que  $S_3$  serait plutôt un grand pays voire un continent, peuplé d'un milliard d'individus ayant en moyenne une vie qui mérite à peine d'être vécue. L'intuition commune, qui s'appuie sur l'image que l'on se fait de ces deux mondes, nous commande généralement de préférer  $S_1$  à  $S_3$ . Toutefois, l'arithmétique utilisée nous pousse à conclure que  $S_3$  est plusieurs ordres de grandeur plus désirable que  $S_1$ . En effet, comme le montrent les deux graphes ci-dessous (qui ne sont pas à l'échelle pour des raisons de clarté), l'aire du Scénario 3 est largement supérieure à celle du Scénario 1.

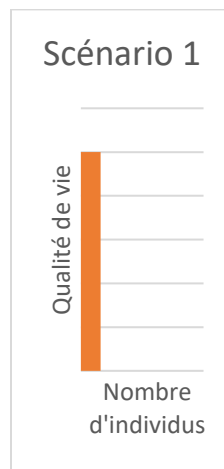


Figure 8. – Scénario 1

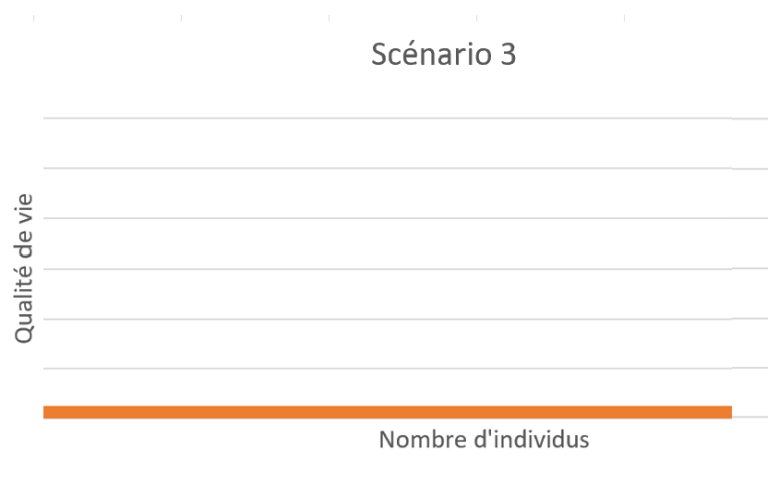


Figure 9. – Scénario 3

Il convient de noter que la situation n'est pas transposable. S'il est concevable qu'une population puisse augmenter virtuellement indéfiniment à condition d'avoir les ressources à disposition, ce n'est pas aussi évident dans le cas du bien-être. Les humains ont un potentiel de félicité limité, et passé un certain stade de confort et d'épanouissement, ces limites biologiques sont atteintes et l'amélioration de la qualité de vie n'est plus vraiment possible. Une fois les besoins et certains désirs comblés, l'excédent de ressources ne se transforme vraisemblablement jamais en bien-être supplémentaire. Prenons un cas bien connu, la nourriture. Lorsque l'on est repu, toutes calories excédentaires ingérées n'améliorent pas notre qualité de vie, bien au contraire. Le stockage ayant ses limites, Harari explique notamment que le surplus qui n'a pas vocation à finir dans un silo peut être distribué dans la communauté comme un investissement pour gagner d'autres formes de bien-être, tel que le pouvoir ou le statut<sup>80</sup>. Le caractère universel de l'argent permet en outre de transformer presque n'importe quelle marchandise en n'importe quelle autre par son intermédiaire, de sorte qu'après les échanges de bons procédés il reste possible de vendre sa nourriture excédentaire. Toutefois, même ainsi les frontières du bien-être ne sont que reculées, et celui qui a bien plus d'argent que de besoins se retrouvera en fin de compte en situation de ne plus pouvoir améliorer sa qualité de vie<sup>81</sup>. Il reste que la roue du désir pourrait continuer de tourner, mais alors on s'aperçoit rapidement de la futilité des désirs non nécessaires et non naturels se succédant sans fin. À cet instant, combler tous ses désirs est vain en cela que ce n'est plus un processus qui participe à améliorer notre qualité de vie. L'on peut également revenir sur des projets comme celui de David Pearce afin d'imaginer que dans une certaine mesure nous pourrions devenir capables de démultiplier les capacités humaines à éprouver du bien-être. Mais il paraît illusoire de croire que la technologie puisse nous permettre de transformer infiniment notre condition. Plusieurs ordres de grandeur pourraient certes être ainsi gagnés, mais il semble logique que le progrès technique se heurte in fine à des barrières de plus en plus infranchissables. De ce fait, il existe certainement comme une asymptote horizontale qu'il serait impossible de dépasser, ce qui nous pousse à croire qu'une telle opération possède elle aussi des rendements décroissants. Comment expliquer cela ? Notamment par le fait que la

---

<sup>80</sup> Yuval Noah Harari, «Sapiens : A Brief History of Humankind», 2011, p. 198.

<sup>81</sup> Daniel Kahneman et Angus Deaton, «High income improves evaluation of life but not emotional well-being», 2010, pour une étude concluant que l'argent fait le bonheur jusqu'à 75 000 dollars annuels.



plupart des biens sont instrumentaux. Or, ce type de bien a tendance à contribuer de moins en moins à ce qui a de la valeur intrinsèque lorsqu'on l'accumule. L'argent a une valeur marginale décroissante parce que plus on en possède, moins on peut extraire de bien-être à partir d'une quantité fixe à partir d'une certaine quantité<sup>82</sup>. Bien sûr, il serait en effet possible d'argumenter en utilisant l'expérience de pensée du monstre d'utilité créée en 1974 par Robert Nozick<sup>83</sup>. Toutefois celle-ci s'inscrit dans une critique plus générale de l'utilitarisme, que tenons ici pour vrai, de sorte que cette question relève d'un tout autre débat que celui que vise à mettre en lumière le présent mémoire. Au demeurant, cette complication ne résout pas le problème de la Conclusion Répugnante, dont la force réside dans l'observation de résultats contre-intuitifs dans l'arithmétique de la position totale. Lorsque l'on observe les nombreux cas où la qualité de vie moyenne chute, il semble toujours que l'on puisse compenser cette baisse par de nouvelles naissances, de sorte que le compromis entre les deux variables aille trop loin, et que la position soit intenable.

Cependant, les partisans de la position totale de rétorquer que cette dernière n'implique pas nécessairement que l'on devrait maximiser la taille de la population en pratique, mais plutôt que c'est une question empirique ouverte que de savoir comment promouvoir le bien-être en réalité<sup>84</sup>. Plusieurs avenues sont généralement considérées pour réfuter l'argument. Maintenant que nous avons une vision un peu plus claire du problème, l'objectif de la prochaine section sera de faire l'état des lieux de ces solutions.

## 2. Les solutions envisagées

En 2017, la *Stanford Encyclopedia of Philosophy* recensait plusieurs manières d'envisager la résolution du problème de la Conclusion Répugnante. Nous prendrons donc le temps d'explorer

---

<sup>82</sup> Michael Huemer, «In Defence of Repugnance», 2008, p. 925.

<sup>83</sup> «Utilitarian theory is embarrassed by the possibility of utility monsters who get enormously greater sums of utility from any sacrifice of others than these others lose ... the theory seems to require that we all be sacrificed in the monster's maw», Robert Nozick, «Anarchy, State, and Utopia», 1974, p. 41.

<sup>84</sup> Richard Chappell, Daniel Meissner et William MacAskill, «Population Ethics, An Introduction to Utilitarianism», 2022.

les avenues qui sont toujours d'actualité, et de les compléter avec la littérature plus récente au besoin. Cependant, nous ne nous intéresserons pas ici à deux d'entre-elles. Nous n'aborderons pas les réponses de philosophes non utilitaristes qui argumentent qu'elle peut être évitée via la réfutation de la transitivité de la proposition « mieux que »<sup>85</sup>. En effet, notre propos restera dans les limites de l'utilitarisme. Nous n'examinerons pas non plus les théories « Schopenhaueriennes » contre-intuitives qui suggèrent qu'il est toujours pire d'exister que de ne pas exister<sup>86</sup> car ce débat s'éloigne trop largement du reste des questions abordées ici. Nous examinerons plutôt dans un premier temps deux manières d'aborder le problème sous l'angle stratégique<sup>87</sup>, avant d'expliquer les mécanismes utilisés par les positions alternatives qui ont été présentées au chapitre précédent, ainsi qu'une autre avenue originale, le « justicisme ».

Le premier ensemble de solutions envisagé par MacAskill, Chappell et Meissner peut se déconstruire en trois concepts : revoir les intuitions, les tolérer ou attaquer les alternatives.

#### i. Revoir les intuitions

Les auteurs proposent dans un premier temps une approche dans laquelle la position totale n'a pas besoin d'impliquer que l'on cherche à maximiser la taille de la population en pratique. Pour ce faire, ils s'appuient sur la constatation que la Conclusion Répugnante fait appel à une intuition qui manque de sérieux dès lors que l'on observe qu'elle implique un nombre astronomiquement large de personnes pour arriver à un scénario répugnant. Dans la pratique, si l'on cherche à appliquer ce problème à l'humanité et que l'on considère qu'elle se compose aujourd'hui de huit milliards d'individus, le « scénario répugnant » exige que dix fois plus d'individus soient créés pour contrebalancer une division par dix du bien-être moyen. Or, comme

---

<sup>85</sup> L'axiome de transitivité voulant que si x est meilleur que y, et y est meilleur que z, alors x est meilleur que z. Thomas fait observer à ce propos que « Although indeed some, like Rachels (2004) and Temkin (2012), have taken the impossibility theorems as evidence of intransitivity, many others of us share Broome's sense that transitivity is simply 'an analytic feature' of comparatives (Broome, 2004, §4.1). » Thomas Teruji, «Some Possibilities in Population Axiology», 2018, p. 812.

<sup>86</sup> David Benatar, «Better Never to Have Been: The Harm of Coming Into Existence», 2008.

<sup>87</sup> Richard Chappell, Daniel Meissner et William MacAskill, «Population Ethics, An Introduction to Utilitarianism», 2022.

nous allons le voir, notre esprit est soumis à de nombreux biais l'empêchant de faire des comparaisons correctes de manière intuitive lorsque vient le temps de départager leurs désirabilités respectives. Cette intuition s'appuie sur une confusion à propos de la qualité des vies dont on parle lorsque l'on dit qu'elles sont « à peine dignes d'être vécues ». En effet, le bien-être des individus du Scénario 3 est, rappelons-le, strictement positif. Par ce simple fait, il est primordial de comprendre que le vocabulaire utilisé par la critique nous donne une image mentale erronée de la réalité de ces deux mondes, puisqu'il ne souligne pas suffisamment que ces vies *sont* bonnes. Au contraire, la critique met l'accent sur le fait que le bien-être est extrêmement proche de la valeur limite  $Q = 0$ . Mais n'oublions pas que dès lors que  $Q > 0$ , les individus qui vivent ces vies préfèrent catégoriquement exister, et préfèrent donc en toute bonne foi vivre leur vie que de ne pas exister. Il semble donc exagéré d'employer l'adjectif « répugnante » pour parler de cette conclusion lorsqu'on réalise à quel point dans le Scénario 3 les individus ont des vies véritablement et sincèrement dignes d'être vécues, du point de vue des concernés. Ce phénomène est à rapprocher de l'idée selon laquelle des arguments qui semblent de prime abord attirants peuvent se révéler erronés. En effet, il peut arriver que nous soyons incapables d'identifier la faute de raisonnement, d'où l'utilisation du terme de paradoxe<sup>88</sup>, ici qualifié de « paradoxe de la simple addition ». De la même manière que nous savons que le mouvement existe sans pour autant être capable de résoudre le fameux paradoxe de Zénon, il semble rationnel de parfois rejeter la conclusion d'un argument qui semble plausible. Plus encore, il semble rationnel de parfois rejeter une prémisse d'un argument sur la base de l'invraisemblance de sa conclusion. Mais ce n'est pas toujours le cas, et c'est précisément ce qui nous intéresse ici. Parfois, il ne semble pas rationnel de rejeter un argument sur la base de la contre-intuitivité de sa conclusion, dans la mesure où nos intuitions peuvent être elles aussi erronées<sup>89</sup>.

---

<sup>88</sup> Michael Huemer, «In Defence of Repugnance», 2008, p. 906.

<sup>89</sup> En particulier, Huemer prend l'initiative de fournir une liste de certains critères qui peuvent aider à déterminer si une intuition doit-être ou non revue. Il écrit, « (1) Does the intuition conflict with firm and widely-shared other intuitions? (2) Is there a plausible error theory, explaining what sort of biases or mistakes might generate the intuition? (3) Is there more than one independent line of argument against the intuition? (4) Is it a 'bare' intuition, lacking significant support from other intuitions, and lacking a satisfactory theoretical explanation? (5) Is there a natural theoretical explanation for a contrary view? A 'yes' answer to any of these questions counts in favour of revising an intuitive belief. None of these criteria are decisive, but each can contribute to the case for revision», Michael Huemer, «In Defence of Repugnance», 2008, p. 907.

Plusieurs auteurs soulignent que l'on peut suspecter l'intuition commune de reposer partiellement sur différents biais qui distordent notre jugement à propos de la Conclusion Répugnante. Examinons-en trois.

#### a. Le biais égoïste

Dans le cas à l'étude, le biais égoïste fonctionne de la manière suivante : puisque nous aspirons tous à une qualité de vie élevée, cela nous offense de nous imaginer à la place d'un individu du Scénario 3 plutôt que du Scénario 1. En réalité, nous nous demandons ce que nous préférons pour nous même entre  $S_1$  et  $S_3$ , et nous comparons un monde où nous serions extrêmement heureux avec un autre où nous le serions à peine. Cette façon de penser est tellement naturelle et instinctive que même si l'on parvient à s'en détacher, à prendre de la distance et à faire l'effort conscient de réaliser que ce n'est pas la question pertinente, il se peut tout de même qu'elle continue d'influencer nos évaluations. Pourtant, il est aisé de comprendre que notre préférence à vivre dans  $S_1$  ne montre aucunement que  $S_1 > S_3$ . Notre jugement ne s'appuie sur rien de plus que sur une comparaison entre la qualité de vie dont nous profiterions si nous vivions dans l'un ou l'autre monde. En réalité, à aucun moment dans notre évaluation intuitive ne prenons-nous en compte tous ses individus de  $S_3$  à qui a été donné le cadeau de l'existence dans un contexte de bien-être (faiblement) positif<sup>90</sup>.

#### b. Le biais des extrêmes

Nous avons déjà abordé une facette du biais des extrêmes en parlant du *large numbers bias*, cette déviation psychologique de la pensée rationnelle qui nous empêche d'apprécier à leur juste valeur les grands nombres. En effet, il semble que passé certains ordres de grandeur, toutes

---

<sup>90</sup> Et Tännsjö de spécifier que « We must guard ourselves against asking ourselves: in which world would I like to live? If we should solve the normative question by any consideration of this sort, then we should add that while it is probable that, if we opt for the Z world, we would come to live there, if we opt for the A world, it is highly unlikely that we will do so. When we realize this the intuition tends to go away. », Torbjörn Tännsjö, «Why Derek Parfit had reasons to accept the Repugnant Conclusion», 2020, p. 390.

les quantités frappent notre imagination avec la même force<sup>91</sup>. Ce principe existe pourtant également dans l'autre sens, et nous sommes plutôt mauvais à appréhender le résultat lorsque les quantités en jeu sont petites. Par exemple, une étude de la fin des années 1970 rapporte que les individus sont davantage tentés d'accrocher leur ceinture de sécurité lorsque leur sont présentées les statistiques des risques d'accidents de la route sur leur vie entière plutôt que par voyage<sup>92</sup>. Les résultats suggèrent que lorsque l'on présente uniquement le petit risque par voyage, les individus ont tendance à ne pas apprécier à sa juste valeur l'ampleur du risque composé sur toute une vie. Ils considèrent le risque par voyage comme négligeable et le négligent donc en oubliant à quel point un risque négligeable peut devenir extrêmement important lorsque composé de manière récurrente<sup>93</sup>.

### c. Le biais de la sous-évaluation de la valeur des vies

Enfin, en ce qui concerne plus précisément l'objet à l'étude, il semble que nous soyons victimes d'une tendance à sous-évaluer les vies de peu de valeur. Lorsque nous imaginons une vie de peu de valeur, et quand bien même nous créerions un personnage factuellement très détaillé, il n'est pas évident de déterminer précisément son niveau de vie. La vie peut être riche et variée, de sorte que lorsque nous cherchons à quantifier le bien-être d'une vie, notre tâche consiste en réalité à pondérer des combinaisons complexes d'objets très différents qui influençant le bien-être, et cela afin de déterminer une quantité unique représentative de la qualité de vie générale de cette vie. Il semble évident que la tâche est ardue. La difficulté est d'ailleurs telle que dans un scénario réaliste nous serions certainement incapables d'être suffisamment précis pour discriminer une vie à  $Q = -1$  d'une autre à  $Q = 1$ . Pour pallier ce problème, nous avons donc tendance, lorsqu'il nous incombe de comparer  $S_1$  et  $S_3$  par exemple, à imaginer des vies fictives simples comme si elles ne contenaient aucune expérience significative autre qu'un plaisir uniforme et monotone. De là surgissent d'autres intuitions qui sont associées

---

<sup>91</sup> John Broome, «Weighing Lives», 2004, p. 57-59.

<sup>92</sup> Paul Slovic, Baruch Fischhoff, Sarah Lichtenstein, Mark Layman et Barbara Combs, «Judged frequency of lethal events», 1978, p. 551-578.

<sup>93</sup> Michael Huemer, «In Defence of Repugnance», 2008, p. 909.

à la monotonie comme l'idée que cette vie monotone serait ennuyeuse ou qu'elle ne conviendrait qu'à un individu n'ayant pas d'autres besoins que le plaisir uniforme, autant de facteurs qui combinés nous incitent à avoir une mauvaise impression de cette situation même si cette vie est légèrement positive<sup>94</sup>.

De manière plus générale, ces trois réponses mettent l'accent sur le fait que les intuitions communes ne sont pas fiables ou tout au moins qu'il n'est pas clair que l'on puisse s'attendre à ce qu'elles distinguent efficacement les vies légèrement bonnes des vies neutres ou légèrement mauvaises. Dans cette situation, il semble alors que la prudence soit de mise lorsque l'on fait appel directement à l'intuition pour évaluer le problème de la Conclusion Répugnante. Plus encore, le vocabulaire utilisé pour décrédibiliser la position totale nous amène à fonder des intuitions déjà faillibles sur un malentendu plus subtil encore, alors que la seule controverse qui ait réellement lieu d'exister lorsque l'on questionne la valeur des vies des scénarios 1 et 3 est de savoir comment quantifier le bien-être de manière appropriée afin de déterminer la façon dont on comptabilise quelles vies valent la peine d'être vécues. Une fois ce point tranché, nous avons une valeur nulle que l'on le notera  $Q = 0$  et toutes vies individuelles ayant pour valeur  $Q > 0$  seront des vies intrinsèquement bonnes et désirables. À noter que le seuil  $Q = 0$  n'est pas le seuil minimum de survie des individus ; mais le seuil où la vie « vaut la peine d'être vécue ». On peut imaginer qu'une vie misérable parsemée de quelques instants de joie n'est pas suffisante pour être désirable et possède donc un bien-être négatif. Peut-être même que la majorité des humains du passé ont vécu de telles vies non désirables. Toutefois, la réalité historique n'a aucun lien avec l'argument philosophique et la définition du seuil de désirabilité. L'argument des auteurs appelle donc de ne pas se fourvoyer en se faisant des représentations trompeuses à propos de la qualité de vie dont on parle lorsque l'on dit  $Q = 0,000X$ . Ce faisant, ils nient l'utilisation de la Conclusion Répugnante comme raisonnement par l'absurde à l'encontre de la position totale.

Une critique peut toutefois être apportée à cet argument. S'il est exact que  $Q > 0$  implique une vie qui vaille la peine d'être vécue, nous pouvons tenter de nous mettre à la place de l'individu en question. Dans le cas  $Q = 0,000X$ , l'individu est proche de ne pas vouloir vivre, de sorte qu'il ou elle pourrait avoir des pensées comme « Si ma vie empire ne serait-ce qu'un tout

---

<sup>94</sup> Michael Huemer, «In Defence of Repugnance», 2008, p. 910.

petit peu, je mettrais fin à mes jours ». Face à ce genre de pensée, il semble difficile d'accepter que le problème vienne avant tout de nos intuitions qui seraient erronées, et nous accordons une certaine crédence au fait qu'un monde habité par une multitude d'âmes ayant cette qualité de vie et ces pensées, c'est-à-dire un monde peu résilient dans lequel les gens mettent régulièrement fin à leurs jours lorsque des incidents surviennent, est un monde qui devrait avoir une faible valeur.

## ii. Tolérer et contre-attaquer

La seconde manière de faire face à la Conclusion Répugnante est de la tolérer. Pour plus de littérature à ce sujet, le lecteur est invité à consulter les ouvrages suivants : Mackie (1985), Hare (1988), Tännsjö (2002), Huemer (2008), Spears and Budolfson (2021) et Gustafsson (2022)<sup>95</sup>. Cette réaction a du sens dès lors que l'on considère que, quand bien même cette remarque semble tout à fait surprenante et contre-intuitive, il est possible que le Scénario 3 soit véritablement plus désirable que le Scénario 1, et plus largement que des Scénarios Z ayant un très grand N et un petit Q soient plus désirables que des Scénarios A ayant un petit N et un grand Q, si tant est que  $V_{SZ} > V_{SA}$ .

Il importe alors de préciser que des implications contre-intuitives similaires à la Conclusion Répugnante ne sont pas propres à la position totale. Avec le développement de la logique, plusieurs théorèmes d'impossibilité ont été découverts par Arrhenius, lesquels prouvent qu'il est incohérent pour une théorie d'éthique des populations de satisfaire intuitivement tous les principes et axiomes désirables, dont fait justement partie le fait de ne pas entraîner une conclusion du style répugnante<sup>96</sup>.

Bien que l'écueil théorique soulevé par la Conclusion Répugnante doive être examiné avec attention dans toute sa particularité vis-à-vis de la position totale, Dean Spears et Mark Budolfson soutiennent que l'argument n'est pas en soi décisif dans un contexte d'élaboration de politiques,

---

<sup>95</sup> Walter Bossert, Susumu Cato et Kohei Kamaga, «Revisiting variable-value population principles», 2022, p. 1.

<sup>96</sup> « The impossibility theorems articulate various adequacy conditions for a population axiology, and state that these adequacy conditions are mutually incompatible », Thomas Teruji, «Some Possibilities in Population Axiology», 2018, p. 810.

dans la mesure où les critiques qui ressemblent à des conclusions répugnantes sont par nature inéluctables<sup>97</sup>. L'idée ici est de montrer qu'aucune position cherchant à s'établir comme base théorique crédible de l'éthique des populations ne peut fondamentalement éviter les résultats paradoxaux et de nature contre-intuitive. En d'autres termes, dans des contextes politiques et de prise de décision, il arrivera certainement que des résultats qui semblent inacceptables émergent. Toutefois, il convient de ne pas confondre le fait de tolérer la Conclusion Répugnante avec le fait d'accepter aveuglément des politiques réellement répugnantes. Le principe de tolérance souligne plutôt une observation qui nous contraint à reconnaître l'existence des dilemmes éthiques complexes et de ne pas négliger leur importance. Dans la réalité pratique, les décisions doivent toujours être modulées en tenant compte d'une panoplie de considérations, et l'argument de la Conclusion Répugnante doit être pris pour ce qu'il est, c'est-à-dire une manière de chercher une solution aux problèmes qui émergent lorsque nos actions affectent les futures générations<sup>98</sup>. Si cette implication gênante n'est pas propre à la position totale, elle doit quand même être adressée afin d'envisager sa résolution de manière créative. D'un autre côté, Chappell, s'inspirant du travail de Greaves<sup>99</sup>, reconnaît que l'on doit envisager le fait que chercher à éviter cet écueil n'est pas nécessairement rentable d'un point de vue théorique. En effet, pour ces derniers, « accepting the repugnant conclusion provides an easy response to the impossibility theorems since the total view is consistent with all the other axioms of these theorems. In light of this, a growing number of ethicists have come to accept and defend the repugnant conclusion and the total view »<sup>100</sup>.

---

<sup>97</sup> Dean Spears, Mark Budolfson, «Why the Repugnant Conclusion is Inescapable», 2018.

<sup>98</sup> Wikipédia, entrée «Population Ethics», dernière révision le 5 juillet 2023.

<sup>99</sup> Hilary Greaves, «Population Axiology», 2017.

<sup>100</sup> Richard Chappell, Daniel Meissner et William MacAskill, «Population Ethics, An Introduction to Utilitarianism», 2022.



### 3. Les avenues alternatives

Outre les stratégies d'évitement décrites précédemment, plusieurs théoriciens s'accordent à dire que la Conclusion Répugnante est une vraie difficulté dont on ne saurait faire fi, et qu'il semble impossible de colmater les brèches par le rapiècement de la théorie. Cela les amène à remettre en question le cadre théorique mis en place par Parfit<sup>101</sup>, et à proposer des avenues alternatives de comptabilisation du bien-être dont les fonctions n'introduisent pas ce type de résultat lorsqu'elles définissent la valeur d'une situation et prescrivent les actions à privilégier. En voici quelques-unes.

#### i. Les théories à valeur variable

Puisque les théories à valeur variable sont des hybrides des positions totales et moyennes, on pourrait s'attendre à ce qu'elles en aient conservé les défauts. Mais ce serait oublier que ces théories ont été pensées justement pour pallier cette lacune, de sorte qu'au fur et à mesure de l'augmentation de la taille de la population, le résultat ne devient presque plus déterminé que par la valeur du bien-être moyen. Or, la Conclusion Répugnante ne prend son sens que lorsque l'on tente de réduire le bien-être de populations très larges. Dans les conditions variables, cela devient impossible dès lors qu'une augmentation extrême de la variable N ne peut plus compenser la baisse de Q que l'on cherche à éviter, de sorte que la fonction variable se comporte comme celle de la position moyenne. Ce compromis a donc la propriété d'éviter la répugnance, mais à un coût théorique souvent considéré insupportable<sup>102</sup> dont nous expliquerons les tenants et les aboutissants au Chapitre III lorsque nous aborderons la principale difficulté de la position moyenne.

---

<sup>101</sup> « More generally, Parfit is searching for a general theory of beneficence- "Theory X" -that also will cover population comparisons. Theory X corresponds to Kenneth Arrow's notion of a social welfare function-both attempt to provide a generic formula or algorithm for ranking social outcomes on the basis of their characteristics », Tyler Cowen, «What Do We Learn from the Repugnant Conclusion?», 1996, p. 754.

<sup>102</sup> Gustaf Arrhenius, «Future Generations: A Challenge for Moral Theory», 2000.

## ii. Les théories de seuil critique

En ce qui concerne les principes de l'utilitarisme de seuil critique originellement théorisés et étudiés par Blackorby et Donaldson en 1984<sup>103</sup>, rappelons que la création d'un individu rend le résultat meilleur dès lors que son bien-être dépasse un certain seuil. La position totale n'est ici qu'un cas particulier dans lequel  $\alpha = 0$ . Mais la valeur de  $\alpha$  peut être modifiée de manière à éviter la Conclusion Répugnante. La condition nécessaire est de placer le seuil  $\alpha$  au-delà du niveau considéré comme valant « à peine » le coup d'être vécu. À ce stade, le zéro servirait toujours à délimiter le seuil au-dessus duquel le bien-être net de l'individu est positif, tout en prenant en considération que ce seuil n'est pas suffisant pour que la vie vaille *incontestablement* la peine d'être vécue. Toute création de vie en dessous de ce seuil serait mauvaise, ce qui empêche de fait les scénarios répugnants dans lesquels une large population vivrait dans des conditions sordides mais néanmoins vivables. L'augmentation de  $N$  reste une chose appréciable, mais se trouve alors limitée par la dégradation des conditions de vie à un niveau plus important que dans l'approche traditionnelle, de sorte que le fait d'avoir une large population proche ou exactement au niveau du seuil de bien-être  $\alpha$  ne puisse pas être considéré comme une situation répugnante. La stratégie semble donc gagnante, mais elle l'est en réalité uniquement sur ce tableau, car il semble là encore que le bénéfice n'ait été réalisé qu'au prix d'une concession profonde qui a pour conséquence une nouvelle difficulté que l'on nomme la Conclusion Sadique. Nous reviendrons sur ce point. Entre-temps, voyons comment la classe non impersonnelle des théories utilitaristes traite le problème de la Conclusion Répugnante.

## iii. Les théories non impersonnelles

Comme précédemment expliqué, nous ne nous étendrons pas sur ce point, mais par souci d'exhaustivité nous dirons tout de même quelques mots à leurs propos. Rappelons la position défendue ici ; il est raisonnable de croire que les théories non impersonnelles ne sont pas très

---

<sup>103</sup> Charles Blackorby, David Donaldson, «Social criteria for evaluating population change», 1984, p. 13-33.

bien armées pour lutter contre les problèmes d'éthique des populations dans la mesure où elles sont soumises à un principe d'incomparabilité postulant que l'identité est un facteur incontournable. En l'état, elles nient donc le fait que nous puissions avoir des raisons non instrumentales d'ajouter des vies positives au reste de l'humanité<sup>104</sup> puisque cela n'améliore pas la vie des individus existants. Plus encore, Kavka soutient un « paradoxe des individus futurs » selon lequel nous n'aurions aucune obligation morale de promouvoir le bien-être des générations futures<sup>105</sup>. Ces théories ont donc comme bénéfice, nous devons l'admettre, de tuer la Conclusion Répugnante directement dans son œuf et il n'est pas nécessaire d'aller plus loin pour montrer pourquoi elles la considèrent comme un problème nul et non avvenu. Nous ne disons toutefois pas que c'est une victoire en soi, car il semble que la solution qu'elles offrent contre le problème à l'étude ne permet nullement de trouver de nouvelles parades à la répugnance dans le cadre de l'impersonnalité que nous défendons. Les positions non impersonnelles s'empêchent d'ailleurs dans d'autres difficultés vis-à-vis de l'axiologie recherchée ici<sup>106</sup>, à l'instar du problème de non-identité illustré dans l'expérience de pensée « The 14-years-old girl »<sup>107</sup> où n'importe quel choix qui engendre une vie positive est bon peu importe le niveau de bien-être du nouvel individu. En outre, elles demeurent incapables d'expliquer de manière convaincante pourquoi certaines actions sont meilleures que d'autres sur le long terme, ce qui les rend d'autant moins crédibles. Nous n'y ferons donc plus allusion, et continuons maintenant cette section sur les avenues alternatives capables de déjouer la Conclusion Répugnante en explicitant la méthode employée par la position moyenne, à laquelle nous consacrerons le prochain chapitre.

#### iv. La réponse de la position moyenne

Dans la mesure où cette position s'intéresse exclusivement à la moyenne de bien-être par personne dans une distribution, la valeur du résultat ne dépend aucunement de la quantité

---

<sup>104</sup> Richard Chappell, Daniel Meissner et William MacAskill, «Population Ethics, An Introduction to Utilitarianism», 2022.

<sup>105</sup> Gregory Kavka, «The Paradox of Future Individuals», 1982, p. 93.

<sup>106</sup> Toby Ord, «The Precipice: Existential Risk and the Future of Humanity», 2020.

<sup>107</sup> Derek Parfit, «Reasons and Persons», 1986, p. 358.

d'individus d'une population. Il s'ensuit intuitivement qu'il est impossible de parvenir à une situation répugnante, puisque la réduction du bien-être moyen ne peut jamais être compensée par une quelconque création de nouveaux individus. Ces derniers ne sont donc désirables que du fait qu'ils participent à l'amélioration de la qualité de vie qui préexiste à leur naissance. Dans le cas contraire, la valeur diminue et le résultat est considéré comme pire.

Toutefois, certaines subtilités rentrent en jeu et l'on peut remarquer l'existence d'un analogue à la Conclusion Répugnante pour certains types de positions moyennes. En effet, si l'on a jusque-là passé sous silence leurs existences, il existe en réalité différentes manières de comptabiliser la moyenne d'une population<sup>108</sup>. L'utilitarisme moyen est donc sujet à une scission ; Hurka découpe la doctrine d'abord en six, puis en onze branches, qu'il appelle A1 à A11<sup>109</sup> et dont on retiendra qu'elles ne sont pas parfaitement équivalentes. Si elles comptabilisent le bien-être de la moyenne exactement de la même manière, elles diffèrent exclusivement par la manière dont elles calculent cette moyenne du bien-être. Considérons par exemple la différence entre les deux formes principales que sont A1 et A2<sup>110</sup>. Le théoricien a ici plusieurs choix. Il peut décider de commencer par calculer le bien-être total éprouvé par un individu au cours de sa vie en faisant la somme du bien-être éprouvé à chaque instant. En répétant cette opération pour chaque membre d'une distribution, il peut moyenniser le bien-être d'une population entière. C'est la méthode notée A1. En revanche, il pourrait tout autant faire valoir l'ordre inverse, à savoir commencer par recueillir le bien-être moyen de chaque individu aux instants  $t, t_{+1} \dots t_{+n}$ . En effet, cela lui permettrait d'avoir accès à une valeur de la moyenne de bien-être par personne à chaque instant, valeurs qu'il pourrait ensuite utiliser afin de moyenniser le bien-être total au cours de la période d'intérêt. Et nous voilà avec une variante A2 qui, selon les circonstances, ne donnera pas nécessairement les mêmes résultats.

Ce qui pose ici problème avec A1 et A2 est que certaines de leurs prescriptions sont assez déplaisantes, pour ne pas dire répugnantes. A2 par exemple nous incite à supprimer les individus qui éprouveront moins de bien-être que la moyenne, quand bien même leurs vies seraient des

---

<sup>108</sup> Contrairement aux sommes qui sont indifférentes à ce genre de priorité, d'où le fait que la position totale soit une et indivisible.

<sup>109</sup> Thomas Hurka, «More Average Utilitarianisms», 1982, p. 116.

<sup>110</sup> Thomas Hurka, «Average Utilitarianisms», 1982, p. 65-69.

plus exquises<sup>111</sup>. Sans rentrer dans les détails, qui dépasseraient notre propos, A1 ne nous demande pas d'en arriver à de telles extrémités mais peut arriver à la conclusion qu'il est moralement impératif d'arrêter de se reproduire dans des situations où les vies des générations futures seraient elles aussi extrêmement heureuses.

Certes, on pourrait argumenter que ces problèmes sont ce qu'ils sont mais qu'ils n'ont rien à voir avec le type de répugnance de la Conclusion Répugnante. Hurka objecte à cela que des variantes comme A2<sup>112</sup> peuvent pourtant entraîner une conclusion somme toute similaire dans laquelle un monde est meilleur qu'un autre tandis que ses habitants ont des vies légèrement positives. Il l'explique lui-même en ces termes : « Let us try to imagine an ideal (finite) course of human history, one in which the average happiness per person is very high for a very long (finite) period of time. Then if A2 is correct there is another vastly longer course of human history which is better than this one, even though the average in this second course is never more than barely positive, and the experiences of the people in it never more than barely worth having. The second course is better because, if it is long enough, the sum of the average happiness per person at each time is greater than it is in the first course<sup>113</sup> ».

Force est de constater que l'utilitarisme moyen, si tant est que l'on puisse le nommer au singulier, doit éviter ce type de variantes pour ne pas subir la même critique que sa rivale. Certaines versions existent qui parviennent donc à se défaire du joug de l'argument clé des opposants de l'utilitarisme totale. Toutefois, celles-ci ont également leurs faiblesses, que nous aborderons plus en détail au prochain chapitre.

Entre-temps, explorons une dernière alternative que nous n'avons pas évoquée jusqu'à présent, et qui introduit d'autres valeurs pour conditionner la création et la gestion du flux d'individus. Nous nous intéresserons donc dans la prochaine section à la position justiciste comme une dernière tentative du totalisme de combiner les visions traditionnelles à d'autres intuitions morales comme la valeur du mérite.

---

<sup>111</sup> On remarquera que dans certains cas, chaque assassinat a pour conséquence de reporter sur un autre la « faute » d'être le plus bas de la moyenne, et ce jusqu'à ce qu'il ne reste qu'un seul individu.

<sup>112</sup> Ainsi que, au moins, A5 et A6.

<sup>113</sup> Thomas Hurka, «Average Utilitarianisms», 1982, p. 66.

## v. La réponse du justicisme

La position justiciste trouve ses fondements dans des idées de Franz Brentano<sup>114</sup>, qui avait à l'origine pour but de proposer un socle à la théorie husserlienne de la valeur. Depuis ses débuts, l'utilitarisme doit en effet faire face à des difficultés concernant des problèmes de justice qu'elle ne semble pas traiter. Or pour Feldman, ces théories ont la particularité secondaire intéressante de produire une explication et même une résolution du paradoxe de la simple addition de Parfit. On notera tout d'abord que le justicisme est une forme d'utilitarisme total, et qu'à ce titre il présente des similarités notamment avec les théories de seuil et de gamme critique que nous avons précédemment discutées. Cela peut paraître à première vue déroutant dans la mesure où nous avons évoqué la Conclusion Répugnante justement comme un écueil inhérent de l'utilitarisme total. Il convient donc de rester vigilant quant aux propositions avancées par le justicisme. Suivant en partie la structure de l'article de Fred Feldman *Justice, Desert, and the Repugnant Conclusion* (1995), nous commencerons par décrire le fonctionnement de la doctrine avant d'expliquer comment celle-ci propose de résoudre notre problème d'intérêt.

### a. Fonctionnement de la théorie

Pour expliquer le fonctionnement de la théorie, il est nécessaire dans un premier temps de revenir sur les fondamentaux de la notion de valeur que nous avons utilisée jusqu'alors. Le cœur de la théorie justiciste se fonde plus précisément sur la décomposition de la valeur en deux éléments, ou plutôt sur l'existence de deux types de valeurs distinctes en ce qui concerne les questions morales. D'une part, il existe la valeur que l'on peut appeler valeur individuelle ( $V_{\text{indiv}}$ ), et qui a trait à la valeur d'un bien au sens large *pour* un individu spécifique. Cette dernière permet donc de comptabiliser à quel point une situation, une action ou encore un objet est bon pour l'individu en question. D'autre part, s'en distingue une valeur plus générale qui se rapporte au bien-être général, ou comme le dit Feldman, au bien pour l'ordre du monde ( $V_{\text{monde}}$ )<sup>115</sup>. Or ce

---

<sup>114</sup> Franz Brentano, «The Origin of Our Knowledge of Right and Wrong», 2009.

<sup>115</sup> Fred Feldman, «Justice, Desert, and the Repugnant Conclusion», 1995, p. 194.

second type de valeur, gardons-le bien à l'esprit, peut tout à fait être en contradiction avec le premier type. En effet, la notion de justice est intimement liée à celle de mérite, ce qui rend la rivalité évidente dès lors que l'on remarque que ce qui est bon pour un individu, le plaisir<sup>116</sup>, ou tout ce qui rend la vie bonne et digne d'être vécue, peut tout à fait rentrer en conflit avec ce que celui-ci mérite.

On a donc deux affaires à régler. L'une eu égard à la valorisation intrinsèque d'un bien pour un bénéficiaire A donné et l'autre vis-à-vis de la question de savoir dans quelle mesure la valeur du monde considéré évolue du fait de l'occurrence de l'évènement qui apporte du bien-être à A. Nous l'avons dit, une dose de bien-être non méritée peut être bonne pour un individu sans pour autant que ce fait ne rende le monde ni meilleur ni plus désirable<sup>117</sup>. Quoi qu'il en soit, ces deux auteurs affirment que lorsque quelque chose qui n'est pas mérité arrive, l'on s'en offusque et l'on crie à l'injustice<sup>118</sup>. Dans ce cas, la justice joue un rôle apparemment crucial puisqu'elle contribue à rendre la souffrance moins mauvaise pour le monde dans la mesure où elle semble restaurer une sorte d'équilibre équitable qui maintiendrait l'ordre social. On dira qu'elle a un effet positif sur  $V_{\text{monde}}$ . Soit dit en passant, la psychologie sociale montre d'ailleurs à quel point les animaux non humains sont également attachés à une certaine forme de justice, chose qui s'observe entre autres lors de la fameuse expérience comportementaliste de Dr Sarah Brosnan et Frans de Waal, qui met en lumière l'aversion spontanée à l'injustice de singes capucins<sup>119</sup>. Peut-être peut-on penser que ces expériences argumentent peut-être en faveur de la thèse de Feldman et de sa compréhension de l'ordre du monde, selon laquelle la valeur de nos actions est au moins en partie déterminée par l'ajustement en fonction du mérite (*desert-adjusted value*) du résultat<sup>120</sup>. Selon ce point de vue, les situations que l'on doit analyser en les

---

<sup>116</sup> Si l'on tient pour acquis un utilitarisme hédonique classique par exemple.

<sup>117</sup> Se pose alors la question du mérite tout court, de savoir si nous méritons quoi que ce soit, et même de savoir si le libre arbitre existe véritablement. Malheureusement, il est évident que nous n'avons pas la place ici d'élaborer sur l'épistémologie sous-jacente à ces concepts, qui mériterait que l'on s'y attarde bien plus longuement.

<sup>118</sup> Chercher à définir ou à circonscrire la notion de justice irait bien au-delà du propos de ce mémoire, disons donc simplement que l'on peut soupçonner Feldman d'entendre par là ce que l'on veut communément dire par justice distributive.

<sup>119</sup> L'expérience a depuis été reproduite sur des chiens, des oiseaux et même des chimpanzés, mais nous proposons au lecteur d'admirer par lui-même les résultats de l'expérience sur les capucins. <https://youtu.be/meiU6TxysCg/>.

<sup>120</sup> Fred Feldman, «Justice, Desert, and the Repugnant Conclusion», 1995, p. 195.

comprenant comme composées notamment d'une part de mérite ou de démerite affectent strictement la valeur réelle d'un monde<sup>121</sup>.

Avant de continuer notre démonstration, citons certaines causes plausibles et pertinentes du mérite qui nous amènent à croire que le justicisme semble intuitivement fondé. Dans son article, Feldman en recense au moins trois : le simple fait d'être une personne, l'effort conscient, ainsi que la valeur morale. Si l'effort compris comme la persévérance déployée à atteindre des objectifs ambitieux et la valeur morale dans le sens des qualités morales telles que l'honnêteté, la générosité, l'intégrité ou encore la compassion sont des avenues communes pour reconnaître les méritants, il n'en va pas de même pour la première cause qu'il évoque puisqu'il semble que celle-ci relève au contraire davantage de la chance ou à tout du moins de l'accidentel, des notions que l'on opposera volontiers à celle du mérite. Nous admettons tout de même de les tenir ici toutes les trois pour acquis comme Feldman, et proposons en outre de rajouter les réalisations comme la contribution sociale ou la production d'un accomplissement significatif pour évaluer le mérite d'un individu, tout en considérant que ce dernier peut vraisemblablement être influencé par des facteurs tels que les privilèges sociaux, les biais cognitifs voire les inégalités structurelles.

L'idée centrale de Brentano revient donc à dire que pour comptabiliser adéquatement la  $V_{\text{monde}}$  d'une situation ou d'un évènement au cours duquel un individu reçoit quelque chose qui augmente sa qualité de vie, il est vital d'examiner à quel point celui-ci le mérite. Pour ce faire, nous ferons appel à une reconstitution du graphique de la représentation de la théorie de Feldman<sup>122</sup>. Ce graphique concerne des actions isolées qui sont évaluées dans les abscisses par  $V_{\text{indiv}}$  et dans les ordonnées par  $V_{\text{monde}}$ , pour une valeur de mérite hypothétique  $m = 100$ .

---

<sup>121</sup> De même que nous n'avons pas pris la peine de définir la notion de justice, nous ne donnerons pas non plus de définition précise de ce que l'on entend par « a mérite b », mais postulons vaguement comme Feldman que ce qu'un individu mérite fait référence à ce qui serait approprié de manière distributive que ce dernier reçoive.

<sup>122</sup> Bien que l'auteur ne fournisse pas de prescriptions arithmétiques, le graphique illustre bien l'idée dont il est question. Extrait du *Graph I* dans Fred Feldman, «Justice, Desert, and the Repugnant Conclusion», 1995, p. 198.



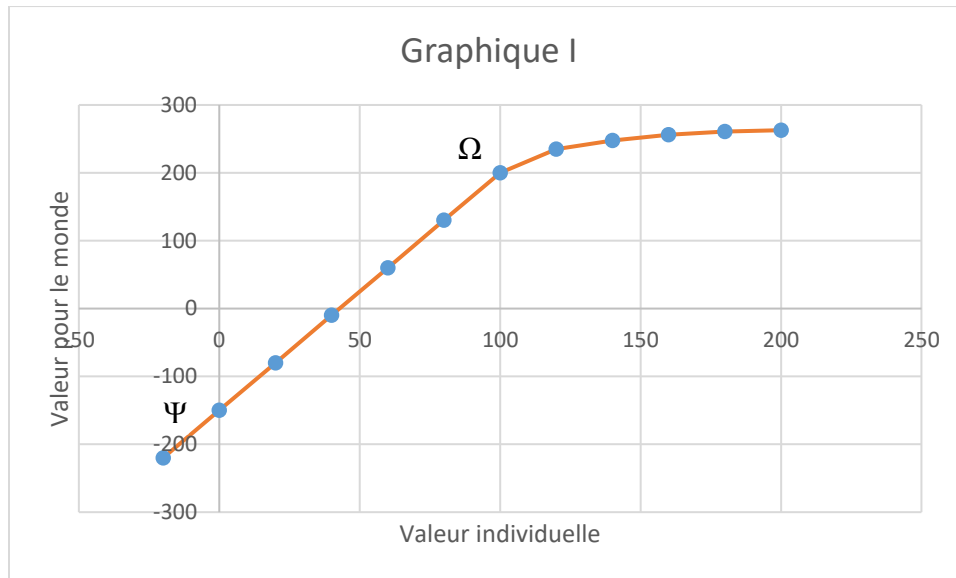


Figure 10. – Graphique I

Constatez que la courbe représente une fonction totale  $f(x)$  qui semble composite<sup>123</sup>. Deux motifs ressortent en particulier. D'une part,  $f(x)$  se comporte comme une fonction affine de la forme  $f(x) = ax + b$  où  $b < 0$  jusqu'au point  $\Omega$  qui correspond à  $x = 100$ . Nous appellerons cette partie de la fonction  $f_1(x)$ . À partir de  $\Omega$ , on observe que les rendements marginaux décroissent, de sorte que la seconde composante  $f_2(x)$  admet une asymptote horizontale d'équation  $y = L$  où  $L \in \mathbb{R}^+$  et dont on peut se douter qu'il n'est pas exagérément grand. Si  $L$  n'est pas spécifié par Feldman, il est toutefois possible de décrire plus précisément que  $\lim_{n \rightarrow \infty} f_2(x) = L$  et que par extension  $\lim_{n \rightarrow \infty} f(x) = L$ .

Ces caractéristiques ont plusieurs implications, que nous allons aborder sous peu, tandis que le principe fondamental de la doctrine justiciste préconise que la valeur totale d'un monde soit donnée par la somme des valeurs pour le monde des vies qui le peuple, tel que :

$$Valeur_J = \sum V_{monde} \quad \text{où } V_{monde} \text{ est la valeur pour le monde d'une situation particulière}$$

Dit autrement, la thèse centrale est donc que la valeur du monde qui compte véritablement est la somme de la valeur des vies qui le peuplent, lorsque celles-ci sont ajustées

<sup>123</sup> Ceci est évidemment une simplification. En réalité, il n'importe pas ici de deviner avec précision et élégance mathématiques la relation qui unit  $V_{monde}$  à  $V_{indiv}$ , et l'on notera que Feldman lui-même ne la fournit pas. L'objet d'intérêt ici est avant tout le *comportement* adopté par la courbe, de sorte que nous laissons aux soins du lecteur les calculs de cette autre considération.

au mérite. Une conséquence de cette arithmétique est qu'un monde peut être considéré bien meilleur qu'un autre alors même qu'il contiendrait moins de valeur individuelle totale. Le lecteur attentif aura bien sûr reconnu là une caractéristique qui nous intéresse tout particulièrement pour dévier la Conclusion Répugnante. Selon le justicisme, la valeur d'un monde ne dépend donc pas uniquement de l'ensemble de valeurs individuelles qu'il contient, mais également de détails à propos de la distribution de cette valeur, où comme le dit Feldman : « who gets to enjoy the goods? who gets to suffer the evils? to what extent do these recipients deserve these goods and evils when they receive them? »<sup>124</sup>. Les points saillants  $\Psi$  et  $\Omega$  correspondent quant à eux d'une part à la situation obtenue lorsqu'un individu ne reçoit aucun bien-être tandis qu'il appartient au niveau  $m = 100$ . L'évènement a alors une valeur nulle pour l'individu, c'est-à-dire  $V_{indiv} = 0$ , et on remarque de surcroît que la valeur pour le monde d'une telle situation est très négative<sup>125</sup>. D'autre part,  $\Omega$  est le point de bascule qui correspond à une situation où l'individu très méritant reçoit exactement son dû. On observe là encore que  $V_{monde} - V_{indiv} > 0$ , de sorte que le monde se porte d'autant mieux que justice a été rendue envers les méritants.

#### b. Rejet de la Conclusion Répugnante par le justicisme

Dès sa version de 1995, le justicisme se targue de pouvoir rejeter la Conclusion Répugnante. Cette section a pour but d'expliquer l'argument de Feldman, et nous ne manquerons pas d'illustrer notre propos par quelques calculs simples.

Dans un premier temps, rappelons et adaptions les exemples précédemment utilisés au contexte du justicisme. Le Scénario 3 représente un monde dans lequel la qualité de vie moyenne des individus est infinitésimalement positive. Posons arbitrairement que cette valeur correspond à 1, de sorte que la notation qui nous intéresse sera de la forme  $V_{indiv} = Q = 1$ . Rappelons également que dans la position totale classique,  $V_{PT} = N \times Q = 10^9 \times 1 = 10^9$ . Enfin, conserverons

<sup>124</sup> Fred Feldman, «Justice, Desert, and the Repugnant Conclusion», 1995, p. 201.

<sup>125</sup>  $V_{monde} \approx -150$  dans le cas présent, mais rappelons que ces chiffres ne sont donnés qu'à titre indicatif et ne constituent pas en eux-mêmes une démonstration. Le simple fait que  $V_{monde} < 0$  est suffisant en soi.

l'hypothèse selon laquelle la population d'intérêt est méritante au seuil de  $m = 100$ <sup>126</sup>. Le point  $\Psi$  du Graphique I (Figure 10.) montre la valeur intrinsèque d'un habitant type du Scénario 3 : si l'individu mérite 100 et ne reçoit que 1, alors on peut dire qu'il a été injustement spolié du bien-être qu'il méritait, de sorte à n'avoir reçu qu'une toute petite fraction de son dû. Cette situation peut à bien des égards être considérée comme inadmissible et ce fait est représenté par une valeur négative proche du point  $\Psi$  qui correspond à  $V_{monde} \approx -150$ . Or, si un milliard d'individus vivent ce genre de vie, la valeur totale pour le monde du Scénario 3 serait en l'occurrence  $V_J = 10^9 \times -150$ , soit un nombre négatif extrêmement grand et par conséquent extrêmement indésirable. Les Scénarios Z, en tant que généralisations du Scénario 3, sont donc des mondes atroces dans lesquels ajouter des vies dont la qualité de vie correspond à la qualité de vie moyenne ( $V_{indi} = 1$ ) est une très mauvaise idée. En effet, chaque vie supplémentaire ainsi créée conduit à une déplétion de  $V_J$ , un résultat qui redevient finalement cohérent avec nos intuitions premières. Là où le totalisme classique présente la situation en expliquant que chaque vie additionnelle rend le monde meilleur, le justicisme de rétorquer que ces vies-là rendent le monde significativement pire. Il s'ensuit qu'aucune Conclusion Répugnante n'émerge lorsque les populations croissent largement. Pour ce qui est du Scénario 1 en revanche, Feldman nous laisse constater que l'arithmétique utilisée permet de conserver une valeur largement positive. Ainsi, un monde où  $m = 100$  et  $V_{indiv} = 100$  est un monde dans lequel chaque vie participe pleinement à l'ordre du monde, de sorte que  $V_{monde} = 200$ , et donc que  $V_J = 10^9 \times 200$ .

Du point de vue justiciste, les mondes Z sont horribles et les mondes A sont merveilleux, ce qui peut tout à fait correspondre avec nos intuitions. Dès lors, on peut se demander si la position a quelque chose d'utile à nous apprendre sur la nature de la Conclusion Répugnante. Feldman estime que la distinction proposée par Brentano lui ait permis de localiser précisément le problème de raisonnement qui entraîne ladite répugnance. L'erreur résiderait en réalité dans la supposition selon laquelle « the value of the world as a whole is the sum of the values for the residents of the lives lived there »<sup>127</sup> tandis qu'en réalité les valeurs pertinentes à prendre en compte relèvent davantage de ce qui est bon pour le monde, et cette notion n'a de sens que

<sup>126</sup> Dans une note de bas de page p. 202, Feldman mentionne qu'en faisant cette hypothèse, il adopte un point de vue égalitariste et affirme que les individus de cette population méritent une certaine considération morale.

<sup>127</sup> Fred Feldman, «Justice, Desert, and the Repugnant Conclusion», 1995, p. 202.

lorsqu'on fait entrer la notion de justice via le mérite dans l'équation. Un monde Z, pour Feldman, c'est un monde criant d'injustice dans lequel des individus vivent des vies misérables où ils ne reçoivent qu'une minuscule fraction de la valorisation totale commune alors qu'ils méritent bien mieux que ça du fait d'être des humains. Et c'est selon lui cette caractéristique, l'injustice, à laquelle nous sommes sensibles et qui nous pousse instinctivement à refuser la valeur de tels mondes. Il est par ailleurs intéressant de noter que cette axiologie a la particularité de contrer également l'appel au « monstre d'utilité » que nous avons abordé précédemment, dans la mesure où les rendements décroissants, soit la partie  $f^2(x)$  de  $f(x)$ , ont la propriété de faire tendre la valeur pour le monde des expériences incommensurablement qualitative pour le monstre vers l'asymptote raisonnable  $y = L$ .

Le justicisme apporte donc une réponse originale et plausible au problème classique de l'éthique des populations. Son argument se différencie en s'appuyant notamment sur deux idées centrales que sont le mérite et l'adéquation entre mérite et qualité de vie d'un individu pour lui-même. Finalement, cette théorie peut être rapprochée des théories de gamme critique dont nous avons dit jusqu'alors qu'elles permettent également de dépasser le paradoxe classiquement établi. En effet, outre le fait qu'elles soient toutes deux des théories totalistes, elles se ressemblent également par la manière dont elles approchent le problème puisque la notion de mérite peut sembler relativement similaire à la notion de seuil qui définit une qualité de vie minimum que les individus devraient pouvoir réclamer. Bien plus pourrait être dit sur le rapprochement de ces théories et des objections pourraient également être soulevées<sup>128</sup>. Toutefois ce document n'a pas la prétention de couvrir l'entièreté du débat d'éthique des populations, et s'il nous semblait intéressant de présenter la thèse de Feldman pour souligner la variété des possibilités envisagées par les théoriciens, d'autres points méritent également qu'on s'y attarde.

Nous concluons cette section en disant que le raisonnement qui fonde la Conclusion Répugnante fait ressortir une difficulté majeure d'un utilitarisme classique souvent considéré comme le plus intuitif. Pourtant, alors que les lacunes sont encore bien présentes dans le débat

---

<sup>128</sup> Notamment, nous invitons le lecteur consciencieux à consulter Arrhenius (2003), lequel remet en cause l'interprétation par Feldman de la clause *ceteris paribus* de la Conclusion Répugnante.

contemporain, des voies alternatives concurrencent l'hégémonie de la position dominante et la question reste aujourd'hui ouverte de savoir si l'argument est rédhibitoire. Or, ce n'est pas la seule critique qui a été formulée à l'encontre de la position totale et il peut être utile de passer au travers d'une autre raison qui sert traditionnellement à remettre en question la dialectique classique. Nous continuerons donc notre examen en portant notre attention sur le fait que nombreux sont ceux qui pensent encore qu'il existe un conflit au sein de la position totale entre les actions à privilégier à court et à long terme.

## **B. Les conflits de temporalité**

### **1. La problématique**

Parce que la position totale valorise les futurs qui pourraient receler une immense quantité de bien-être total, nous avons vu au Chapitre I qu'elle a naturellement tendance à privilégier les actions et les politiques qui ont pour conséquence de maximiser l'utilité totale en réduisant le plus significativement possible les risques d'extinction de l'humanité. Cela amène parfois certaines confusions quant à la priorisation des décisions à prendre, et certaines peurs quant à leurs impacts pratiques. Il est donc commun de faire valoir que ces positions valorisent davantage la survie à long terme que les problèmes de court terme. Selon l'argument, cette situation inadmissible permet de réfuter que la position totale est une doctrine enviable. Après tout, il peut sembler tout aussi répréhensible de détourner les yeux des problèmes réels du monde actuel pour focaliser son attention sur le futur lointain.

Dans ces conditions, la position totale semblerait se résigner à accepter que soit commit au nom du plus grand bien ce que le sens commun prend pour abject et inacceptable, comme la prise de décisions unilatérales, la peine de mort ou la non-assistance à personne en danger. Dans son article de 2001 *Existential Risks : Analyzing Human Extinction Scenarios and Related Hazards*, Boström propose par exemple de réfléchir sur l'utilité de maintenir une capacité d'action préventive de dernier recours, afin d'être en mesure de sauvegarder l'humanité d'un risque existentiel trop important. Il écrit « Creating a broad-based consensus among the world's nation

states is time-consuming, difficult, and in many instances impossible. We must therefore recognize the possibility that cases may arise in which a powerful nation or a coalition of states needs to act unilaterally for its own and the common interest. Such unilateral action may infringe on the sovereignty of other nations and may need to be done preemptively »<sup>129</sup>. En d'autres termes, et suivant la formule consacrée, on ne ferait pas d'omelette sans casser des œufs. Or, force est de constater que l'humanité fait aujourd'hui face à des problèmes d'action collective complexes parfois anthropogéniques comme la crise environnementale, la course à l'intelligence artificielle<sup>130</sup> ou la sécurité nucléaire. Ces risques menacent gravement notre existence et notre prospérité, en conséquence de quoi il devient primordial de trouver un ensemble de solutions efficaces et acceptables. En effet, si le plan mis en œuvre pour parer ces menaces n'est pas suffisamment efficace, on peut alors s'attendre à un chaos décentralisé voire à la fin de l'humanité, tandis que s'il n'est pas suffisamment acceptable, c'est la dystopie centralisée et autoritaire qui guette. Récemment, Boström a, par exemple, défendu l'idée selon laquelle il risque de devenir nécessaire de construire et d'utiliser à son plein potentiel un système de surveillance globalisée qui monitorerait tout jusque dans l'intimité des foyers<sup>131</sup> si des mesures adéquates alternatives ne sont pas rapidement mises en place. Cette surveillance de masse à l'échelle planétaire serait décrétée faute d'alternative pour éviter tout comportement clandestin ou terroriste à même de nuire à l'humanité<sup>132</sup> et à la maximisation du bien-être total.

---

<sup>129</sup> Nick Boström, «Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards», 2002, p. 21.

<sup>130</sup> Sur ce point, le Center for AI Safety a publié pour la première fois en mars 2023 la déclaration suivante : « AI experts, journalists, policymakers, and the public are increasingly discussing a broad spectrum of important and urgent risks from AI. Even so, it can be difficult to voice concerns about some of advanced AI's most severe risks. The succinct statement below aims to overcome this obstacle and open up discussion. It is also meant to create common knowledge of the growing number of experts and public figures who also take some of advanced AI's most severe risks seriously. *Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war* », Center for AI Safety, «Statement on AI Risk, AI experts and public figures express their concern about AI risk».

La lettre ouverte a été, en date de rédaction de ce mémoire, signée par plus de 350 cadres, chercheurs et ingénieurs travaillant dans le domaine de l'intelligence artificielle. Le même mois, certains chefs de file d'OpenAI, de Google DeepMind et d'Anthropic ont d'ailleurs été reçus par le président Joe Biden pour parler des futures réglementations de l'IA. Si l'enjeu est d'améliorer la qualité de vie des humains, il en va de toute évidence d'un principe de justice intergénérationnelle.

<sup>131</sup> Nick Boström, «The Vulnerable World Hypothesis», 2019, p. 455.

<sup>132</sup> Cette hypothèse s'inscrit dans un contexte où le développement des technologies est vu comme un tirage de boules dans un sac opaque ; tirage qui peut nous être fatal si la boule tirée est noire. La vulnérabilité de l'humanité face à ce type de jeux de hasard est évidente, puisqu'il est impossible de revenir en arrière pour réessayer. On pourrait dire que la société est aujourd'hui technonaïve et qu'elle doit parvenir à maturité. Plus précisément peut-

La critique porte donc d'une part sur la notion de faire « le plus grand bien » et d'autre part sur la prémisse selon laquelle tous les êtres humains comptent de manière égale. Posons-nous une question pour illustrer le propos. D'après la position totale, pour atteindre le plus grand bien, devrait-on canaliser notre énergie afin d'aider un nombre modeste d'individus vivants qui souffrent ou plutôt déployer nos efforts envers le nombre astronomiquement large d'individus potentiels qui vivront dans le futur lointain ? Du fait des quantités en jeu, la réponse est de toute évidence que nous devrions nous intéresser davantage à ces derniers individus. Ils ne comptent pas moins que les premiers, de sorte que si nous avons pour précepte de maximiser l'efficacité de nos actions sur le bien-être total, le choix devrait être vite fait.

Cela découle de ce que nous avons déjà tenu pour hypothèse que le futur est incroyablement vaste. Benjamin Todd, fondateur de l'organisation 80,000 Hours, explique que « Since the future might be very long, there could be far more people in the future than in the present generation. This means that if you want to help people in general, your key concern shouldn't be to help the present generation, but to ensure that the future goes well in the long-term »<sup>133</sup>. Toutefois, le détracteur de la position totale n'est pas nécessairement de cet avis, et peut trouver répréhensible un tel point de vue. Nous allons donc tenter d'expliquer en quoi la préservation du futur n'est pas en contradiction avec l'amélioration des conditions de vie contemporaines.

## 2. La proposition de résolution

Par nature, un risque existentiel rendrait de façon permanente l'humanité incapable de se développer suffisamment pour créer autant de nouveaux individus qu'il est physiquement permis, un gâchis inacceptable pour les plus fervents partisans de l'utilitarisme total. Toutefois, cela n'a rien à voir avec le fait de négliger les problèmes actuels. Simplement, nous avons des raisons de croire que des milliers de milliards d'humains restent à naître et que certaines des

---

être, il s'agit en réalité pour l'humanité de combler l'écart entre sa puissance et sa capacité à la maîtriser, écart que l'informaticien Tristan Harris appelle le gouffre de sagesse.

<sup>133</sup> Benjamin Todd, «Longtermism : the moral significance of future generations», 2017.

actions que nous entreprendrons dans un futur proche auront un impact sur le cours de la civilisation dans son entièreté.

Malheureusement, il semble qu'une bonne partie de la société n'apprécit pas ce fait à sa juste valeur. Supposons un instant que ces derniers acceptent volontiers la prémisse « P1. Tous les humains comptent de manière égale »<sup>134</sup>. Dans ce cas, s'il est alors possible d'éviter une catastrophe morale dans le futur, par exemple en prenant des dispositions pour que la guerre russo-ukrainienne ne dégénère pas en une troisième guerre mondiale, le simple fait que la majorité des individus qui seraient victimes des conséquences de la guerre n'existent pas encore ne permet pas de justifier l'inaction. Les futures générations peuvent légitimement avoir une importance morale capitale pour des raisons numériques. La question qui se pose alors est de savoir par quels moyens nous pouvons aider ces futures générations. Une partie de la réponse a déjà été explicitée : nous devons simplement faire en sorte que ce futur voit le jour, et donc décourager quiconque mettrait en péril son existence. D'un autre côté, ce vaste futur peuplé d'un nombre incommensurable d'âmes est incertain, et il se pourrait tout aussi bien qu'il soit utopique que dystopique, ou plus vraisemblablement n'importe où entre les deux. Si notre priorité absolue est d'assurer la survie de la civilisation, une priorité non moins importante est de faire en sorte que la société humaine s'épanouisse. Or, l'épanouissement d'une société dépend souvent des valeurs qu'elle prône et qui la constitue, de sorte qu'améliorer les valeurs sur lesquelles se fondera le futur semble d'une grande pertinence<sup>135</sup> pour améliorer le bien-être total d'une population. Les récentes transformations civilisationnelles qui nous ont conduits à vivre dans des sociétés qui prônent des valeurs telles que la démocratie ou l'abolition de l'esclavage<sup>136</sup> sont des exemples réussis, et on l'espère pérenne, d'amélioration des valeurs morales. Notons toutefois

---

<sup>134</sup> Certaines « rank-discounted utilitarian axiologies » ont été défendues récemment, voir Geir Asheim et Stéphane Zuber, «Escaping the repugnant conclusion: Rank-discounted utilitarianism with variable population», 2014.

<sup>135</sup> On remarquera que ce problème des valeurs fait écho au problème de l'alignement des valeurs décrit dans Nick Bostrom, «Superintelligence, paths, dangers, strategies», 2014, p. 185-208. Bostrom y décrit l'importance de choisir avec soin et d'implémenter correctement les valeurs encapsulées par les futurs algorithmes d'intelligence artificielle générale. À l'instar du présent problème, si les valeurs ne sont pas judicieusement choisies, il pourrait en résulter le verrouillage irréversible de tout un pan des futurs très désirables. Il en va sensiblement de même pour les valeurs qui fondent les sociétés humaines, à la différence près que l'on peut tout de même penser qu'une amélioration positive pourrait tout de même se produire plus tard dans le futur.

<sup>136</sup> Et plus récemment encore l'antiracisme et le féminisme.



que rien ne peut être définitivement tenu pour acquis, et que la défense des libertés gagnées est un combat à part entière.

Cela étant dit, nous pouvons désormais considérer le problème sous une perspective nouvelle. La critique dit essentiellement que le présent serait oublié et déconsidéré. Mais ce serait oublier que le futur n'est rien de plus qu'une conséquence du présent. Les décisions que nous prenons aujourd'hui ont un impact direct et réel sur la vie de nos enfants, et il est plutôt intuitif de penser qu'une société saine est sur la bonne route si l'un de ses buts est de créer une société future prospère. De fait, la promotion des intérêts des générations futures ne va pas nécessairement de pair avec la négligence des individus actuellement vivants, et les actions qui permettront d'améliorer le futur ont toutes les raisons de nous bénéficier aussi. Prenons le cas de la récente pandémie de COVID-19. La mise en œuvre d'un plan efficace destiné à éviter de futures pandémies permettra de toute évidence de sauver de nombreuses vies tant aujourd'hui que demain. Telle est notamment la vocation du Pandemic Sciences Institute dont la mission est de « radically accelerate the development and implementation of interventions that benefit humanity »<sup>137</sup>. Il en va de même pour les organisations qui luttent contre la prolifération des armes biologiques, comme le Biological Weapons Convention, et on pourrait citer pléthore de nouveaux centres académiques ayant chacun des ambitions plus ou moins larges, à l'instar du Future of Humanity Institute, du Future of Life Institute, du Centre for the Study of Existential Risk ou encore du Machine Intelligence Research Institute. Chacun de ces organismes, qu'il promeuve l'innovation en faveur de la transition environnementale, l'irénologie, l'éducation ou la santé, nourrit de précieuses avancées qui bénéficieront largement aux populations actuelles et futures. Bien que la position totale priorise le long terme pour des raisons numériques, elle ne se désintéresse jamais du court terme. Plus spécifiquement, si l'on a des raisons de croire que les risques de souffrance et de décès à court terme influencent la pérennité de l'humanité, alors la position totale soutient avec force les politiques permettant de retourner à la paix et la prospérité. D'ailleurs, même s'il existe un plafond au sacrifice total acceptable en échange de la maximisation du potentiel humain, il paraît invraisemblable que la société actuelle soit proche de cette limite. L'idée selon laquelle nous finirions moralement obligés de faire des sacrifices trop

---

<sup>137</sup> Pandemic Sciences Institute, section « about », page consultée le 22 juillet 2023.

exigeants est en vérité un leurre dès lors que l'on présente quelques chiffres. Ord par exemple écrit que « we can state with confidence that humanity spends more on ice cream every year than on ensuring that the technologies we develop do not destroy us »<sup>138, 139</sup>. Selle relève quant à lui qu'aujourd'hui 0,001 % du PIB mondial est délibérément destiné à protéger l'avenir à long terme de l'humanité, un montant qui représente moins de 0,2 % des revenus de l'industrie des casinos aux États-Unis. Il est donc évident que nos priorités ne semblent pas alignées sur ce qui a le plus d'importance<sup>140</sup>.

Nous avons donc vu que la critique rate sa cible, et qu'au contraire, la position totale et le long-termisme ont d'excellentes raisons d'aider les générations présentes. Toutefois, il serait charitable d'exprimer une intention possible de ses détracteurs, et de dire que le problème pourrait résider dans l'instrumentalisation des vies présentes à des fins futures. Celles-ci seraient au service du bien-être, plutôt que l'inverse<sup>141</sup>. Dans ce cas, il semble que le totaliste ne chercherait à venir en aide à son contemporain que dans la mesure où cela est la meilleure manière de parvenir à ces fins, qui est véritablement le futur glorieux qu'il revendique. En cela, il serait monstrueux car il négligerait la valeur en soi des individus vivants. Or, ce n'est que pure contingence que le futur s'aligne sur le présent. Si d'aventure la meilleure manière de préserver le futur était d'annihiler le présent, le totaliste retournerait volontiers sa veste et n'hésiterait pas à déchaîner l'enfer sur Terre. Bien que l'argument soit plausible, il ne nous paraît toutefois pas des plus pertinents ici puisqu'il remet en question les fondements mêmes de l'utilitarisme, pour qui la fin justifie les moyens. Toutefois, cette remarque nous amène indirectement à poser la question du manque de lois universelles et du traitement des petites probabilités aux énormes effets. La position totale doit-elle donc soutenir que laisser mourir un grand nombre d'individus aujourd'hui est « rentable » si par là on peut s'assurer de diminuer un tant soit peu la probabilité d'extinction ? Puisque les larges populations sont favorisées par la position totale, alors la

---

<sup>138</sup> « Less than 5% the amount that is spent annually on ice cream in the U.S. » dans *Frequently Asked Questions*. <https://longtermism.com/faq/>.

<sup>139</sup> Toby Ord, «The Precipice: Existential Risk and the Future of Humanity», 2020, p.58.

<sup>140</sup> Gaëtan Selle, The Flares, « L'éthique du FUTUR LOINTAIN : les implications morales de nos décisions ? », 2023, 7:55 / 10:16.

<sup>141</sup> Kieran Setiya, «The New Moral Mathematics», 2022.

sauvegarde d'un futur lointain potentiellement très peuplé nous invite à questionner ces actions. Nous laisserons ici la question ouverte pour des recherches ultérieures<sup>142</sup>.

En revanche, nous ferons observer que la critique de la priorisation du long terme sur le court terme semble pouvoir être renversée, et l'on pourrait noircir des pages entières à propos des dangers du court-termisme et du biais du présent, direction actuelle de l'immense majorité des politiques humaines, et moteur incontestable de conséquences gravissimes. « 'now' commands so much more attention. [...] in politics the dominant time frame is a term of office, in fashion and culture it's a season, for corporations it's a quarter, on the internet it's minutes, and on the financial markets mere milliseconds »<sup>143</sup>. Cette préoccupation pour le présent et l'immédiateté est telle qu'il y a de quoi se questionner sur nos habiletés à avoir de l'empathie pour nos descendants lointains. Mais si l'on ne privilégie pas le long terme, quelle temporalité doit-on privilégier ? Et pour quelles raisons ? Nous avons vu que le court-termisme est une impasse, mais devrions-nous nous laisser guider par une position intermédiaire selon laquelle il conviendrait de s'occuper autant du court que du long terme ? Dans tous les cas, il semble qu'aujourd'hui l'humanité souffre indubitablement du fait de ne pas suffisamment penser et prévenir les conséquences à long terme, en atteste la crise de l'environnement et de la biodiversité, ainsi que la course à l'intelligence artificielle dont on se souviendra qu'elle a débuté quelque part dans les années 2020. Mais comprenons également que préserver un court terme enviable est essentiel, parce que des drames à court terme peuvent rendre tous nos plans à long terme caducs. Par ailleurs, il n'est pas inutile de rappeler que c'est toujours au nom de longs termes aussi hypothétiques qu'idéaux que des humains ont créés des courts termes infernaux<sup>144</sup>.

## Conclusion

Nous avons vu dans ce chapitre que la Conclusion Répugnante mettait en évidence une lacune sérieuse qui entachait dramatiquement la réputation de cette position par ses

---

<sup>142</sup> Pour plus de détails, le lecteur est invité à se référer à la thèse de Petra Kosonen, « Tiny Probabilities of Vaſt Value », 2022.

<sup>143</sup> Richard Fisher, « The Perils of Short-Termism: Civilisation's Greatest Threat », 2019.

<sup>144</sup> La Fabrique Sociale, « Superorganisme : la seule transition possible pour l'humanité ? », 2023, 14:15 / 27:57.

conséquences dérangeantes. Toutefois, nous avons fait observer que plusieurs solutions originales avaient été proposées, solutions qui sont au cœur du débat contemporain. Nous avons également ajouté que d'autres critiques ont été formulées à l'encontre de la position totale, telle que le problème de la priorisation, aussi appelé conflit des temporalités.

Conséquemment, le chapitre subséquent aura pour but de continuer à tracer les grandes lignes et les faits saillants de la position moyenne présentée au Chapitre I, afin de mieux avancer dans notre réflexion à propos de la meilleure alternative à considérer. Nous y aborderons ainsi plusieurs critiques qui lui sont faites.

## **IV. Chapitre III – Critiques de la position moyenne**

La position moyenne est la doctrine la plus documentée dans la littérature scientifique après la position totale. Bien que peu de philosophes la soutienne dans sa forme classique<sup>145</sup>, il est intéressant de noter qu'elle capture l'intuition très répandue selon laquelle le nombre d'individus d'un monde n'est pas un facteur pertinent à la détermination de la valeur de ce dernier. Nous avons vu au Chapitre I le fonctionnement de la position moyenne ainsi que les arguments historiquement utilisés en sa faveur. Nous avons également abordé le cas de la position moyenne comme une réponse à la Conclusion Répugnante dans le Chapitre II. Ce chapitre prendra quant à lui le chemin inverse et développera certaines critiques traditionnellement soulevées pour la discréditer. Plus spécifiquement, nous procéderons de la même manière qu'auparavant, en évaluant dans un premier temps la littérature portant sur l'argument principal utilisé par les détracteurs, avant de nous pencher sur une partie des autres lacunes adressées à la position moyenne. Chemin faisant, nous ferons observer que les théories critiques et à valeur variable sont elles aussi concernées par l'argument principal, mieux connu sous le nom de Conclusion Sadique. Nous passerons notamment en revue plusieurs arguments de Huemer, Parfit et Hurka.

### **A. La Conclusion Sadique**

#### **1. La problématique**

La Conclusion Sadique est un argument que l'on peut retrouver dans plusieurs articles d'Arrhenius<sup>146</sup>. Très vite, il s'est imposé comme l'argument principal servant habituellement à discréditer la position moyenne. En substance, il explique que cette dernière entraînerait

---

<sup>145</sup> On retrouve toutefois un soutien récent en la personne de Michel Pressman dans sa contribution «A Defence of Average Utilitarianism», 2015.

<sup>146</sup> Voir notamment Guðstaf Arrhenius, «An Impossibility Theorem for Welfarist Axiologies», 2000.

invariablement une conclusion pire que la Conclusion Répugnante, puisqu'il est possible d'imaginer un monde dans lequel une population A possède une certaine qualité de vie moyenne  $Q_A$ , et dans lequel la doctrine affirme qu'il serait meilleur de créer un petit nombre d'individus malheureux qu'un grand nombre d'individus heureux ayant une qualité de vie inférieure à la population d'origine<sup>147</sup>. Autrement dit, il existerait des circonstances dans lesquelles la position moyenne préconiserait de créer des vies avec un bien-être négatif plutôt que des vies avec un bien-être positif, depuis la même situation initiale. Cette constatation découle encore une fois du compromis entre le bien-être et le nombre d'individus en jeu. En effet, le résultat s'explique par le fait que, lorsque l'on prend comme situation initiale une distribution à haute utilité, c'est-à-dire ayant un  $Q$  élevé, ajouter un nombre important d'individus heureux *mais moins* heureux que la population d'origine peut impacter à la baisse la moyenne davantage que si l'on ajoutait que peu d'individus malheureux. Illustrons ce fait de la même manière que précédemment, à l'aide des trois scénarios proposés par Huemer<sup>148</sup>. Le scénario 4 est ici le scénario de référence, c'est la situation initiale à partir de laquelle vont diverger les scénarios 5 et 6.

Dans  $S_4$ , le bien-être moyen d'une population de  $N$  individus est élevé, de sorte que  $Q_{S4} = 100$ . Dans  $S_5$ , on ajoute à la population d'origine un individu malheureux, que l'on appellera la victime, dont le malheur est représenté par la quantité  $Q_{victime} = -\varepsilon$ . La qualité de vie totale de ce monde est donc de  $100N - \varepsilon$ , et la qualité de vie moyenne  $Q_{S5} = \frac{100N - \varepsilon}{N+1}$ .

Dans  $S_6$ , on ajoute à la population d'origine un nombre  $N^2$  d'individus heureux, dont le bonheur moyen est représenté par la quantité  $Q_{heureux} = \frac{\varepsilon}{N^2}$ . La qualité de vie totale de ce monde est donc de  $100N + \varepsilon$ , tandis que la qualité moyenne vaut  $Q_{S6} = \frac{100N + \varepsilon}{N^2+N}$ .

Plus précisément, on constate que  $Q_{S5} > Q_{S6}$  dans tous les scénarios où le couple de variables  $(N ; \varepsilon)$  résout l'inéquation  $\frac{100N - \varepsilon}{N+1} > \frac{100N + \varepsilon}{N^2+N}$ .

Cette conclusion est tenue pour acquise par de nombreux philosophes, mais certains comme Huemer la trouvent bien moins plausible que la Conclusion Répugnante dans la mesure

<sup>147</sup> Selon les mots d'Arrhenius « A principle implies the sadistic conclusion if there is an initial utility distribution such that adding only people with negative levels of wellbeing (below neutrality) to that distribution is better than adding only individuals with positive utilities (above neutrality) ». Gustaf Arrhenius, «An impossibility theorem for welfarist axiologies», 2000, p. 247-266.

<sup>148</sup> Michael Huemer, «In Defence of Repugnance», 2008, p. 911.

où contrairement à cette dernière qui stipule simplement qu'un certain nombre de bonnes vies peut être meilleur qu'un nombre significativement moins grand de vies merveilleuses, la position moyenne soutient de son côté qu'un certain nombre de bonnes vies peut être pire qu'un nombre significativement moins grand de vies atroces<sup>149</sup>, ce qui est inconcevable. Toutefois, on peut opposer cette vision à des articles plus récents qui remettent en cause le présent état de fait. En effet, certains auteurs<sup>150</sup> ne sont pas convaincus par deux caractéristiques de la Conclusion Sadique que sont d'une part son nom équivoque, voire trompeur, et d'autre part la nature de la comparaison sur laquelle s'appuie l'argument.

#### i. Le sadisme de la Conclusion Sadique

Dans le chapitre précédent, nous évoquons la suspicion de certains auteurs vis-à-vis de la connotation péjorative de l'adjectif répugnant pour désigner la conséquence la plus discutée de la position totale. Le contre-argument présenté ici s'appuie sur un raisonnement analogue, nous incitant à remettre en cause nos préjugés biaisés provenant de cette appellation provocatrice, et nous poussant à nous intéresser à la question de savoir si la Conclusion Sadique est véritablement sadique. Accordons-nous d'abord sur ce que pourrait vouloir dire sadisme dans ce contexte. On peut imaginer que puisqu'un sadique valorise la souffrance, un scénario sadique en serait un dans lequel la valeur d'un monde augmente à mesure que le bien-être négatif progresse. Or, il ressort de cela que ce n'est pas du tout ce que propose l'argument, et que le nom est peut-être assez mal choisi. Dans le scénario de la Conclusion Sadique, le fait d'amener à l'existence une ou quelques victimes n'est en réalité jamais vue comme une avancée positive améliorant la valeur du monde. Au contraire, il est trivial que  $Q_{S5} < Q_{S4}$  puisque  $-\varepsilon < 0$ . Armstrong (2014) fait remarquer que « Only if helping the victim is somehow impossible do you get the conclusion<sup>151</sup> », ce qui veut dire que plutôt que de proposer une apologie de la souffrance, l'argument expose le fait que, dans le cas où il serait impossible d'aider la victime, et seulement dans ce cas, ce qui

---

<sup>149</sup> Michael Huemer, «In Defence of Repugnance», 2008, p. 913.

<sup>150</sup> Voir Stuart Armstrong, «Embracing the “sadistic” conclusion», 2014.

<sup>151</sup> Stuart Armstrong, «Embracing the “sadistic” conclusion», 2014.

est par ailleurs une limite flagrante en pratique, alors on pourrait considérer que ce serait la moins mauvaise option que de créer une ou quelques victimes plutôt qu'une multitude d'individus heureux qui feraient davantage baisser la moyenne.

## ii. À propos de la comparaison utilisée

La forme de l'argument est donc bien différente de celle qu'on peut lui attribuer de prime abord si l'on se fie à son appellation suggestive, car elle repose sur un dilemme entre deux options mauvaises et se focalise plus particulièrement sur le contraste et la comparaison entre les options<sup>152</sup>. Par exemple, si une large population d'individus à la qualité de vie faible mais positive n'est pas une bonne chose, ce qui est sujet à débat comme nous l'avons précédemment expliqué, nous pouvons inventer une autre situation non désirable mais un tant soit peu moins mauvaise que la première, impliquant une victime unique. Dès lors, l'argument ne décrit pas la qualité de vie d'une population mais seulement que certaines circonstances sont pires que d'autres, de sorte que « the victim isn't playing a useful role here : they're just an example of a bad outcome better than (A), only linked to (A) through superficial similarity and rhetoric »<sup>153</sup>.

Dans le cas de la position moyenne, et cela est vrai pour la majorité des axiologies qui la subissent, la Conclusion Sadique est donc un argument qui peut être réduit à l'affirmation selon laquelle créer une large population de vies à peine positives est parfois mauvais. Or, c'est précisément ce qui est au cœur du débat de l'éthique des populations, de sorte que l'argument n'est pas aussi décisif qu'il semble l'être. Cette observation fait apparaître que l'argument est juste un cas particulier du principe plus général selon lequel s'il est mauvais d'ajouter une large population de vies à peine positive, alors il pourrait être mieux de faire autre chose de mauvais pour éviter ce premier problème. Ici, Arrhenius propose que cette autre chose soit la création d'une victime, ce qui nous pousse intuitivement à nous indigner. Mais, on peut utiliser d'autres exemples pour comprendre à quel point celui-ci en particulier est trompeur. En effet, l'on pourrait

---

<sup>152</sup> Voir la réponse de Ghatanathoah à Armstrong, 2014.

<sup>153</sup> Stuart Armstrong, «Embracing the “sadistic” conclusion», 2014. Voir aussi Sreenivasan Subramanian, «The Repugnant, the Sadistic, and Two “Despotic” Conclusions in Population Ethics», 2023, p. 7.



imaginer une situation analogue dans laquelle des individus s'auto-infligent une certaine dose de déplaisir pour éviter d'avoir des enfants, ou même simplement utilisent une partie de leurs ressources pour éviter d'avoir des enfants plutôt que de les investir autrement pour améliorer leur qualité de vie. Or, cela est monnaie courante dans beaucoup de pays, parfois même pour des raisons morales<sup>154</sup>. La création d'une victime qui catalyserait toute la souffrance nous répulse, mais est-il vraiment improbable qu'il existe des cas dans lesquels il est raisonnable d'accepter de la souffrance pour le bien commun ? C'est le problème ancestral du dilemme entre les faux positifs et les faux négatifs, et l'on accepte pourtant volontiers que quelques accidents de la route se produisent, ou que quelques innocents soient condamnés, pour que la majorité bénéficie de transports ou d'une justice optimisée, c'est-à-dire d'une meilleure qualité de vie.

---

<sup>154</sup> Beaucoup d'individus pensent notamment qu'il n'est pas éthique d'avoir des enfants si les conditions de leur épanouissement ne sont pas réunies.

## B. Les situations dans lesquelles des actions n'ont pas d'effets sur certains individus

Si l'on met de côté la Conclusion Sadique, la question se pose de savoir pourquoi la position moyenne, dont on a dit au Chapitre I qu'elle n'avait pas nécessairement besoin de reposer sur l'intuition de neutralité, n'est pourtant pas une position très populaire au sein de la communauté scientifique. On remarquera que les critiques abordent en réalité une thématique bien précise, naturellement soulevée par les utilitaristes, qui concerne les situations dans lesquelles des actions n'ont pas d'effets sur certains individus. Prenons l'action de mettre au monde un individu supplémentaire ayant un niveau de vie décent mais sous la moyenne. Si l'on présume que cette naissance n'affecte en rien la qualité de vie du reste de la population, pourquoi devrait-on dans ce cas faire valoir que le bien-fondé de cette action dépend de quelque chose qui serait extrinsèque aux conséquences pour l'individu concerné ?<sup>155</sup> C'est pourtant bien ce que semble dicter la position moyenne lorsqu'elle affirme que la naissance d'un individu sous la moyenne est une mauvaise chose, que c'est une action indésirable étant donné qu'elle entraîne une réduction de la moyenne globale.

La position moyenne se retrouve donc embarrassée par un ensemble de critiques que Pressman divise en deux catégories dans son article *A Defence for Average Utilitarianism* (2015)<sup>156</sup>. Cette section aura donc pour objectif de présenter et de tenter d'apporter des éléments de réponse à ces critiques.

---

<sup>155</sup> Michael Pressman, «A Defence of Average Utilitarianism», 2015, p. 397.

<sup>156</sup> Michael Pressman, «A Defence of Average Utilitarianism», 2015, p. 398.

## 1. De « haut en bas » : la critique de Huemer

L'un des arguments se manifeste tout d'abord sous la forme d'une objection voulant que la position moyenne échouerait car elle ne chercherait en réalité pas, contrairement à ces prétentions, à maximiser ce qui a une valeur intrinsèque, soit le bien-être, et cela généralement aux dépens justement de la maximisation du bien-être moyen. Pour Huemer<sup>157</sup>, cet argument d'équivalence exprime que si l'on pense qu'il est bon de maximiser la somme de bien-être d'une vie, alors on est engagé à soutenir la position totale et non la position moyenne. En effet, pourquoi considérer la moyenne en tant que telle si l'axiome fondamental prescrit de rendre les individus les plus heureux possibles ? Illustrons l'argument par une comparaison qui s'étend sur la durée.

D'une part, imaginons que l'existence entière de Charline est à un niveau bien-être  $Q_C$  pendant une durée  $t_C$ . D'autre part, dans un scénario différent l'existence entière de Charline a le même bien-être mais pendant une durée équivalente à la moitié de la durée  $t_C$ , soit  $\frac{1}{2}t_C$ , tandis qu'un autre individu, Hubert, prend le relais à l'instant où Charline disparaît et vit une vie épanouissante au niveau  $Q_C$  pendant une autre durée  $\frac{1}{2}t_C$ . En définitive, les deux situations ont duré  $t_C$ , une durée qui peut être décomposée en deux fois  $\frac{1}{2}t_C$ , et le bien-être de Charline et Hubert est dans les deux scénarios de  $Q_C$ . Raisonnons maintenant par tranche de temps. Dans la première situation, Charline est au niveau  $Q_C$  pendant la première moitié de  $t_C$ , tandis que dans la seconde situation Charline est également au niveau  $Q_C$  pendant la première moitié de  $t_C$ . L'histoire se déroule et dans la première situation, Charline continue de vivre au niveau  $Q_C$  pendant la seconde moitié de  $t_C$ , tandis que dans la seconde situation c'est Hubert qui est au niveau  $Q_C$  pendant la seconde moitié de  $t_C$ . On peut ici raisonnablement conclure que les deux moitiés de la première situation sont également désirables, puisque chacune est constituée du même bien-être pendant la même durée. De même, la position moyenne est une doctrine non impersonnelle voulant que l'identité de Charline et Hubert n'intervienne pas dans la question de la désirabilité de la seconde situation, de sorte que l'on doit également conclure que les deux moitiés de la seconde situation

---

<sup>157</sup> Michael Huemer, « In Defence of Repugnance », 2008, p. 918.

sont également désirables pour les mêmes raisons, à savoir que chacune est constituée du même bien-être pendant la même durée. Cela nous mène directement à la conclusion selon laquelle les deux situations sont également désirables<sup>158</sup>.

De toute évidence, la thèse huemérienne plaide largement en faveur de la position totale, ne laissant aucune place à la position moyenne. Toutefois, on remarquera peut-être que le raisonnement repose en grande partie sur une définition particulière de la nature de la valeur intrinsèque. Si l'argument semble à première vue séduisant, il ne résiste en réalité pas à un examen plus approfondi de ce concept par Pressman<sup>159</sup>. Plus précisément, l'argument rate sa cible si l'on regarde les choses autrement, car le problème peut ne pas être dû à un échec de la position moyenne, mais plutôt à la croyance populaire erronée de ce qui constitue une valeur intrinsèque. Ce qui nous induit en erreur, c'est peut-être l'intuition qu'un instant de bien-être peut avoir une valeur intrinsèque propre. Pourtant, force est de constater qu'une autre position raisonnable existe stipulant que rien n'a de valeur intrinsèque inconditionnellement. Notamment, et pour reprendre le thème de la condition évoqué précédemment, le bien-être d'un individu peut être vu comme n'ayant de valeur que conditionnellement à l'existence de l'individu. Il existerait alors deux cas de figure. Si l'individu existe, alors il remplit de facto cette condition. Mais dans le cas où l'on s'interroge sur l'impératif de créer de nouveaux individus, ces derniers ne remplissent pas la condition d'existence, et l'on doit concéder que ce n'est pas un problème moral que de savoir si les créer est une bonne chose ou non<sup>160</sup>. L'argument de Huemer serait alors rejeté dès lors que l'on comprend qu'il soutient en réalité la thèse selon laquelle si une quantité

---

<sup>158</sup> Huemer va plus loin encore et explique en outre pourquoi il est cohérent de présumer que plus d'individus heureux est plus désirable. Il observe notamment deux symétries pertinentes. L'une est spatio-temporelle ; de sorte qu'en s'appuyant sur l'intuition voulant que la survie de l'humanité possède une valeur qui n'est pas sujette aux rendements décroissants, il tient pour acquis que le bien-être d'une vie positive contribue au bien-être total de manière indépendante de la temporalité où se situe cette vie. Par symétrie, on doit donc conclure que l'ajout d'individus dans l'espace doit être considéré tout aussi désirable que l'ajout d'individus dans le temps, or l'ajout d'individus dans l'espace c'est la définition même de l'augmentation de la population. L'autre est basée sur l'axe plaisir/souffrance et reprend en substance le dilemme de l'intuition de neutralité, s'appuyant plutôt sur l'intuition commune voulant que si les vies au bien-être négatif, soit les vies qui ne valent pas la peine d'être vécues, ne doivent jamais être créées, et qu'il serait immoral de les amener à l'existence, alors par symétrie encore, il semble que les vies positives aient une bonne raison d'être amenées à l'existence.

<sup>159</sup> Michael Pressman, «A Defence of Average Utilitarianism», 2015, p. 398.

<sup>160</sup> Plutôt, ces questions sont purement esthétiques, ou matière de goût et de préférences.

a une valeur intrapersonnelle, alors elle doit nécessairement en avoir une interpersonnelle, ce qui n'est pas démontré.

Cet élément de réponse n'est donc pas décisif dans le sens où il ne permet pas de soutenir que la position moyenne devrait prévaloir contre le totalisme. Plutôt, il fait partie de ces arguments qui dévient l'attaque, et qui exposent une situation dans laquelle les positions rivales défendent chacune leur vision de ce qu'est la valeur intrinsèque. À la lumière de ces possibilités, il semble que la position moyenne s'en sorte indemne.

Interrogeons-nous à présent sur une autre manière d'aborder l'argument à propos des populations non concernées. Cette dernière, popularisée par Parfit, est souvent mieux connue sous l'appellation de l'argument de l'Égyptologie.

## **2. De « bas en haut » : la critique de Parfit**

Hormis le fait de savoir si une action qui n'affecte personne doit être jugée selon d'autres critères, une autre question cruciale se pose dans le contexte de la position moyenne, celle de déterminer si les individus non affectés par une action doivent être pris en compte lors de la prise de décision. Autrement dit, lorsque aucun effet réel n'est induit sur un groupe d'individus, doit-on identifier ce groupe comme pertinent ? Cette question demande donc plus largement quelles sont les populations pertinentes à prendre en compte lorsque l'on calcule la qualité de vie moyenne. Si la réponse semble évidente car il nous apparaît qu'elle présente un contresens flagrant, nous allons voir que les réponses qui nous viennent naturellement en tête ne sont pas des plus convaincantes. Voyons plutôt les possibilités qui s'offrent à nous<sup>161</sup>.

---

<sup>161</sup> Une liste des possibilités envisagées par Parfit est dressée par Michael Pressman, «A Defence of Average Utilitarianism», 2015.

### i. La population pertinente est intemporelle

Cette idée a pour conséquence que les décisions que nous devons prendre dépendent largement de la qualité de vie de toute l'humanité jusqu'à la décision, et notamment de celle des habitants de l'Égypte Antique, ce qui semble incongru. Cet argument, que l'on retrouve habituellement sous le nom d'Égyptologie, implique notamment que la prescription pour un couple de faire un enfant dépend de faits à propos des vies d'individus très éloignés spatio-temporellement de ce couple. Plus spécifiquement, il semble que « On this view, whether or not a person should have a child now depends in part on how well off -and how numerous-people were in the Stone Age, and it is difficult to see why this sort of consideration should be relevant »<sup>162</sup>. En effet, si l'on a des raisons de croire que l'enfant aura une vie heureuse, disons que  $Q_{enfant} = 90$ , tandis que l'égyptologie comme champs d'étude archéologique nous fournit des preuves que cette population vivait dans des conditions qui lui permettait d'approcher une moyenne vers les 50, alors amener à l'existence l'enfant permettrait de rehausser la moyenne. En revanche, si les Égyptiens étaient extrêmement heureux, plus heureux que l'enfant à naître, alors sa naissance constituerait une faute morale. Mais cette considération semble pourtant tout à fait hors de propos dès lors que l'on observe que la vie de l'enfant est, dans les deux, cas strictement identique. En somme, si la position moyenne s'avérait exacte, il s'ensuivrait que la désirabilité d'une vie positive serait soit bonne soit mauvaise en fonction des circonstances et des nouvelles découvertes paléontologiques en Égypte. Pourtant, nos intuitions sont généralement formelles sur le fait que la prospérité de civilisations disparues n'ayant jamais eu d'interactions avec nous n'est pas et ne devrait pas être un facteur pertinent à la détermination de la valeur d'une vie, ni une composante normative d'une éthique des populations raisonnable.

Une réponse récente consiste à changer les modalités de la position moyenne classique pour adopter ce que Grill appelle la position de la Somme des Moyennes. Cette modification a pour caractéristique principale d'isoler les différentes temporalités futures afin de permettre une évaluation indépendante de la manière dont ces futurs se comparent au passé et au présent. Elle attribue en outre une valeur fondamentale à la longévité de la population, contre la position

---

<sup>162</sup> Jefferson McMahan, «Problems of Population Theory», 1981, p. 115.

classique<sup>163</sup>. Selon l'auteur, cette position évite l'Égyptologie dans la mesure où elle additionne les moyennes de la population locale indexées dans le temps dans un contexte particulier<sup>164</sup>.

ii. La population pertinente n'est que la population du présent

Une alternative envisageable serait d'adopter une version temporellement dirigée vers le présent, et qui engloberait tous les individus encore vivants à l'instant où la décision est prise. On la retrouve parfois sous le nom de l'argument de la Mort Prématuurée (*Early Death*), car ses conséquences sont des plus sombres. En effet, maximiser la moyenne du bien-être dans ces conditions, c'est réduire le nombre de vies peu désirables au minimum. Enfin c'est ce que l'on pourrait croire si l'on s'arrête à la surface de l'argument. En réalité, celui-ci ayant été très discuté<sup>165</sup>, il mérite que nous nous y attardions pour expliquer comment cette conclusion peut être évitée, et parce que cette idée motive une préférence théorique qu'il convient de mettre en lumière.

L'argument se présente traditionnellement sous la forme suivante<sup>166</sup>.

P1. Lorsque l'on cherche à maximiser une moyenne, le plus simple reste souvent d'éliminer les entrées les plus petites de la distribution.

P2. Or, la position moyenne fait justement la moyenne d'entrées qui représentent le bien-être.

---

C. Il est désirable d'éliminer les individus sous la moyenne.

De toute évidence, ce résultat est extrêmement contre-intuitif pour la majorité d'entre nous, en particulier lorsque la qualité de vie moyenne de la population est positive, et plus encore

---

<sup>163</sup> Kalle Grill «The Sum of Averages: An Egyptology-Proof Average View», 2023, p. 105.

<sup>164</sup> Puisque nous n'allons pas vraiment en profondeur ici, le lecteur curieux est invité à consulter plus en détail les pages 106 et 107 de l'article «The Sum of Averages: An Egyptology-Proof Average View», Kalle Grill, 2023 qui propose une description formelle de l'argument.

<sup>165</sup> Kalle Grill «The Sum of Averages: An Egyptology-Proof Average View», 2023, p. 105-106.

<sup>166</sup> On le retrouve également chez Hurka (1982, p. 66) pour qui la position moyenne « requires us to kill people who will feel less than the average amount of happiness from now on into the future », et chez Greaves (2019, p. 51) « it would constitute an improvement if anyone whose instantaneous well-being is consistently below average died sooner rather than later ».

lorsqu'elle est très haute. Dans une société où des millions d'individus vivent des vies extraordinaires, il semble plus que douteux de vouloir supprimer ceux qui ont des vies un tout petit peu moins extraordinaires que les autres.

Fort heureusement, la position moyenne a de quoi se soustraire plutôt facilement à cette attaque, notamment en faisant valoir un principe axiologique centré sur la valeur tout au long de la vie plutôt que simplement à un moment déterminé. Effectivement, comment distribuerait-on le bien-être lorsque l'on parle d'une relation ou encore d'un accomplissement si ce n'est en se basant sur la somme de bien-être tout au cours de la vie ? On ne saurait vraiment donner une certaine quantité de bien-être à un instant particulier pour de tels phénomènes. Cette vision semble d'autant plus naturelle que « it is doubtful whether momentary well-being is a meaningful notion at all, since what seems to matter is only lifetime well-being, including of course momentary experiences as contributors (Bramble 2018) »<sup>167</sup>. Une position moyenne qui suivrait ce principe pourrait donc sans problème échapper à l'argument de la Mort Prématinée. Il s'ensuit que la question de savoir si la population pertinente devrait être temporellement dirigée vers le présent reste ouverte.

iii. La population pertinente n'est que la population de l'avenir

La dernière version qui peut nous venir rapidement en tête lorsque l'on cherche à se demander quelle est la population pertinente à utiliser est quant à elle temporellement dirigée vers le futur, de sorte que nos actions dépendraient de la qualité de vie des individus vivant dans le futur. Cette version est souvent nommée l'Égyptologie Renversée, et il semble difficile de justifier une telle version pour les mêmes raisons que pour l'Égyptologie.

---

<sup>167</sup> Kalle Grill «The Sum of Averages: An Egyptology-Proof Average View», 2023, p. 105.



### 3. Une réponse plus générale à la détermination de la population pertinente

Une manière plus complète de répondre à la question de la pertinence peut être trouvée dans Pressman (2015)<sup>168</sup>, qui répond qu'une position impersonnelle se doit d'être temporellement neutre et d'inclure tous les individus, qu'ils soient ou non affectés par l'action. Sa démonstration s'appuie sur une redéfinition des termes de la question à poser. Pour lui, la question de la détermination de la population pertinente à prendre en compte n'est pas tant de savoir quels groupes temporels d'individus devraient être inclus, mais plutôt pour quels groupes temporels d'individus non affectés doivent être inclus en plus de ceux qui sont affectés dans tous les groupes temporels. Son point se décompose de la manière suivante. D'une part, il existe des individus non affectés et des individus affectés par une action. D'autre part, il semble que le fait de penser que la population pertinente est intemporelle soit souvent associé à l'idée que les individus affectés doivent être inclus dans la population pertinente peu importe le groupe temporel auquel ils appartiennent, c'est-à-dire peu importe *quand* ils existent. C'est exactement le constat que fait Égyptologie ou encore Égyptologie Renversée. Mais Égyptologie, si l'on ne prend que cet exemple pour simplifier le propos, ne s'arrête pas là puisqu'il propose d'inclure en plus tous les individus non affectés, peu importe leur groupe temporel<sup>169</sup>. Or, c'est là que le bât blesse et que nos intuitions sont contrariées ; ce qui nous dérange véritablement dans Égyptologie c'est que l'on inclue dans la population pertinente à la maximisation du bien-être moyen des individus non affectés par nos actions, ici les Anciens Égyptiens. D'ailleurs, peu importe que ces individus non affectés vivent dans le passé, le présent ou le futur, l'écueil provient du fait qu'ils ne sont pas affectés tout court. La meilleure question à poser est donc de se demander si la population pertinente devrait inclure seulement les individus affectés ou bien les individus affectés plus les individus non affectés. Si la meilleure réponse est la première, alors il est trivial que tous les individus non affectés doivent être exclus, peu importe leur groupe temporel.

---

<sup>168</sup> Michael Pressman, «A Defence of Average Utilitarianism», 2015, p. 405.

<sup>169</sup> Égyptologie Renversée inclut pour sa part les individus non affectés du présent et du futur.

Bien que nous nous soyons accommodés du terme « individus non affectés », Pressman va à ce point du raisonnement exposer une nuance ignorée jusqu’alors, en scindant ce concept en deux et en postulant une distinction nette entre les individus non affectés par une action et les valeurs non affectées par une action. Les individus non affectés sont des personnes, par exemple Charline dans les deux moitiés de  $t_c$ , tandis que les valeurs non affectées sont les quantités de bien-être non affectées lorsque l’on regarde la situation d’un point de vue impersonnel, par exemple Charline puis Hubert dans les deux moitiés de  $t_c$ . On rappellera que dans les deux cas, le bien-être moyen est exactement le même, et ce qui change est juste l’identité des personnes qui reçoivent ce bien-être, un fait jugé non pertinent pour toutes les positions impersonnelles, dont fait partie la position moyenne. Dans ce contexte, il semble que si l’on choisit d’inclure dans la population pertinente deux personnes ayant le même bien-être, l’on doit accepter la proposition selon laquelle une seule personne qui garde le même bien-être doit aussi être incluse. Intéressons-nous maintenant à la question de déterminer si les valeurs non affectées devraient être prises en compte, et examinons les scénarios suivants tirés de Pressman (2015).

Scénario 7	Scénario 8
5, 5, 5, 5, 5, X	5, 5, 5, 5, 6, 2

Tableau 4. – Comparaison des scénarios 7 et 8

Lorsque l’on cherche à comparer les distributions de  $S_7$  et  $S_8$ , il semble que les 5 ne sont pas affectés, et que leur prise en compte ne change pas le résultat. Dès lors, seules semblent pertinentes les valeurs X et 2, ainsi que le dernier 5 de  $S_7$  et le 6 de  $S_8$ .

Mais ajoutons un Scénario 9, tel que  $S_9 : (3, 7, 3, 7, 8, 1)$  et cherchons à comparer  $S_7$ ,  $S_8$  et  $S_9$ .

On observe qu’aucune valeur de  $S_9$  n’est présente dans  $S_7$  ni  $S_8$ , de sorte que les 5 qui étaient auparavant des valeurs non affectées deviennent tout à coup extrêmement affectées et il semble dans ce cas pertinent de les prendre en compte pour effectuer la comparaison. Cette expérience met en lumière qu’il n’y a en réalité pas de valeurs qui soient fondamentalement affectées ou non affectées ; mais plutôt que savoir si une donnée est affectée ou non affectée dépend plutôt d’un jugement qui dérive de quelque chose de plus fondamental, à savoir les

données brutes complètes des distributions comparées<sup>170</sup>. Le point important ici est que la conséquence d'une action qui demande de faire un choix c'est l'ensemble total du résultat, de sorte que ce sont ces ensembles totaux qui doivent être comparés entre eux, et non uniquement certaines de leurs parties comme on aurait tendance à le croire en regardant simplement les différences entre  $S_7$  et  $S_8$ . Nous avons donc de bonnes raisons de croire que l'on devrait inclure les valeurs non affectées dans toutes les comparaisons puisqu'elles n'existent pas en tant que telles mais sont relatives aux comparaisons. Et donc, par voie de conséquence, l'on devrait également inclure les individus non affectés. Pour ces raisons, Pressman conclut que les positions non impersonnelles doivent donc adopter une vision temporellement neutre vis-à-vis de la question de la population pertinente à préférer. En outre, il propose l'explication selon laquelle les exemples cités par Parfit, comme l'Égyptologie, semblent persuasifs parce qu'ils sont en réalité fondés sur des intuitions de positions non impersonnelles<sup>171</sup>.

Bien que la doxa semble exclure de fait les individus non affectés, on comprend désormais mieux pourquoi la manière de choisir la population pertinente n'est en réalité pas aussi évidente qu'on pourrait le croire de prime abord. Dans ce contexte, puisqu'elle possède des justifications plausibles et non arbitraires pour appuyer ses affirmations, il semble que la position moyenne ne puisse être écartée sur la simple base des arguments de Huemer et de Parfit.

D'ailleurs, en guise de conclusion de cette section il importe de faire remarquer que les difficultés de déterminer la population pertinente à prendre en compte ne sont pas propres à la position moyenne. La position totale aussi est vouée à devoir faire ce choix, de sorte que l'on peut réfuter le fait que l'argument discréditerait la position moyenne, si elle était incapable de procéder à cette détermination, au profit de la position totale. En réalité, cette question de détermination précède le débat entre les positions moyennes et totales, et doit donc être adressée sans être influencée par une vision partisane pour l'une ou l'autre<sup>172</sup>. La question de la pertinence est entièrement distincte de la question de ce qu'il est meilleur de maximiser car ce sont des concepts indépendants.

---

<sup>170</sup> Michael Pressman, «A Defence of Average Utilitarianism», 2015, p. 409.

<sup>171</sup> Nous avons déjà fait observer cette tendance au Chapitre I lorsque nous cherchions à comprendre le problème de l'Intuition de Neutralité.

<sup>172</sup> Michael Pressman, «A Defence of Average Utilitarianism», 2015, Note de bas de page 39, p. 406.

Une autre objection à la position moyenne que nous allons désormais analyser concerne le traitement par cette dernière de la souffrance. Plus précisément, nous allons préciser notre pensée sur le célèbre argument dit des Deux Enfers<sup>173</sup>, et nous ferons observer que ce dernier s'appuie sur des intuitions proches de celles qui portaient Égyptologie.

### C. La Conclusion Infernale

Considérons la situation suivante et imaginons deux mondes infernaux. Dans E1, 10 individus souffrent atrocement. La qualité de vie moyenne est de  $Q_{E1} = -100$ . Dans E2, 10 millions d'individus souffrent atrocement, mais leur souffrance est légèrement moindre que dans E1, de sorte que l'on puisse dire que  $Q_{E2} = -99$ . Si l'on suit l'axiologie de la position moyenne, alors on doit considérer que  $E2 > E1$ . Une éthique appliquée suivant les normes de la position moyenne nous incitera donc à faire advenir E2 plutôt qu'E1, ce qui va clairement à l'encontre de la position totale<sup>174</sup>.

Explicitons un peu plus l'idée derrière l'argument des Deux Enfers que l'on retrouve parfois sous le nom de Conclusion Infernale pour faire écho aux autres « Conclusions » de l'éthique des populations. La conclusion montre en substance qu'un principe qui valorise le fait de moyenniser la qualité de vie des individus peut stipuler qu'ajouter des vies négatives peut être une excellente chose dans certaines situations où la qualité de vie moyenne est déjà négative. En effet, si cette action permet de relever la moyenne, alors c'est une action indéniablement positive, et donc souhaitable. Plus encore, le fait d'amener à l'existence des individus aux vies atroces peut être considéré comme une action bien meilleure que l'action de n'amener personne à l'existence si les vies ainsi créées sont ne serait-ce qu'infinitésimalement moins atroces que la moyenne des vies de la situation d'origine. L'argument diffère donc substantiellement de la Conclusion Sadique pour qui créer de telles vies reste répréhensible, mais simplement moins répréhensible que de créer une large population de vies positives qui ferait baisser davantage la moyenne. Il nous semble toutefois évident qu'E2 est bien pire qu'E1, en conséquence de quoi nous sommes

---

<sup>173</sup> Derek Parfit, « Reasons and Persons », 1986, p. 393.

<sup>174</sup> Pour qui E2 est 990 000 fois pire que E1.

naturellement tentés de soutenir une version asymétrique de l'intuition de neutralité, afin de déterminer un principe qui ne s'appliquerait pas à aux individus dont la vie est composée de bien-être strictement négatif. Nous voulons en réalité penser qu'il n'est pas neutre mais mal de créer des vies misérables.

Rappelons maintenant que nous avons défini l'intuition de neutralité dans le Chapitre I comme la croyance selon laquelle le fait d'ajouter des individus devrait être considéré éthiquement neutre. Rappelons également que nous avons montré via l'expérience de pensée des trois mondes A, B et C pourquoi cette intuition est incohérente et contradictoire. Si l'argument contre l'intuition de neutralité est correct, alors il s'ensuit logiquement que l'erreur se trouve dans nos intuitions davantage que dans la Conclusion Infernale, et que par conséquent la position moyenne juge de manière contre-intuitive, mais pourtant correcte, le cas des Deux Enfers.

En somme, il semble que contrairement à la Conclusion Sadique qui reste plutôt controversée, la Conclusion Infernale peut être catégoriquement rejetée comme un argument ne remettant pas en cause l'intégrité de la position moyenne, car elle puise sa force dans des intuitions non impersonnelles incapables de bâtir une axiologie solide et pertinente au regard des problèmes de l'éthique des populations. Mais cet argument, quoique suffisant pour réfuter l'objection, peut être doublé d'un autre argument positif servant à expliquer nos intuitions pour la situation à l'étude<sup>175</sup>. Imaginons un cas symétrique aux Deux Enfers, soit les Deux Paradis dans lesquels  $N_{P1} = 10$  et  $Q_{P1} = 100$ , et  $N_{P2} = 10\,000\,000$  et  $Q_{P2} = 99$ . Nous avons dit que l'intuition selon laquelle la taille est une question de goût dans les Paradis et non dans les Enfers s'explique par des intuitions non impersonnelles. Mais en réalité, lorsque nous considérons différemment ces deux scénarios symétriques, il apparaît que nous faisons clairement une distinction entre causer un tort et ne pas prodiguer un bienfait. Or, cette distinction est d'ordre déontologique, et ne saurait s'appliquer dans un contexte conséquentialiste comme celui-ci<sup>176</sup>.

Les Deux Enfers semblent donc définitivement incapables de rejeter la position moyenne. Un dernier mot toutefois à ce propos, qui concerne la manière dont il conviendrait de réagir face

---

<sup>175</sup> Michael Pressman, «A Defence of Average Utilitarianism», 2015, p. 418-419.

<sup>176</sup> Michael Pressman, «A Defence of Average Utilitarianism», 2015, p. 422.

à ce cas d'étude. Une autre intuition généralement admise pour les Deux Enfers est que l'on devrait mettre fin aux jours de tous ces individus torturés. Or, le lecteur aura peut-être songé au fait que la position moyenne semble dire que la mort d'un individu dont la qualité de vie individuelle est de même valeur que la qualité de vie moyenne du monde qu'il habite est un évènement éthiquement neutre. La position moyenne semblerait à ce titre incapable de coïncider avec nos intuitions. Mais rappelons-nous que nous avons évoqué au Chapitre I l'existence de différences entre plusieurs catégories de positions moyennes. En particulier, la catégorie 1 n'implique pas ce résultat puisqu'elle prend comme valeur pertinente la somme des valeurs de tous les moments vécus. De ce point de vue, la mort d'individus souffrant atrocement est vue comme une bonne chose puisqu'elle augmente la qualité de vie moyenne de la population<sup>177</sup>. On peut néanmoins se demander si un tel raisonnement est compatible avec notre précédente explication consistant à prendre comme valeur pertinente la valeur tout au long de la vie pour échapper à l'argument de la Mort Préaturée.

Mis à part ce dernier point que nous n'aborderons pas davantage, il nous reste au moins deux critiques à faire valoir contre la position moyenne. Ces critiques sont présentées en dernier car elles ne sont pas en soi suffisantes pour rejeter la position, mais participeraient, si elles s'avéraient exactes, à un affaiblissement des raisons de la préférer. En outre, nous les traiterons dans une partie unique car elles concernent toutes les deux le même thème, à savoir des difficultés en lien avec la taille des populations, soit un aspect généralement peu évoqué par cette position.

## **D. Difficultés apparaissant lorsque la population grossit**

### **1. Remise à plus tard de la souffrance**

Depuis le début de ce document, nous estimons que la position moyenne présume que  $V_{PM} = Q_{PM}$ . Or, comme nous venons de le rappeler,  $Q_{PM}$  peut être calculée de différentes

---

<sup>177</sup> Michael Pressman, «A Defence of Average Utilitarianism», 2015, p. 423.

manières, ce qui donne lieu notamment aux trois catégories explorées au Chapitre I. Revenons un instant sur ces différences qui vont s'avérer extrêmement pertinentes pour souligner le problème qui nous intéresse, à savoir le fait que certaines positions moyennes justifient la remise à plus tard de la souffrance.

Si  $Q_{PM}$  est le rapport à un instant  $t$  entre l'utilité totale des individus vivants et le poids moral total des individus vivants<sup>178</sup>, alors le premier problème qui se pose provient du fait que l'utilité est dépendante du temps, de sorte que si la taille de la population grossit à  $t_{+1}$ , par exemple en devenant dix fois plus grosse, alors les plaisirs et les déplaisirs de chaque individu de  $t_{+1}$  comptent dix fois moins que ceux de chaque individu de  $t$ .

Ce phénomène justifie donc directement la remise à un moment ultérieur des déplaisirs, et inversement incite à multiplier les plaisirs immédiats. Cette situation ressemble fortement au problème économique de la tragédie des communs, situation dans laquelle chaque individu est poussé à catalyser au maximum les bienfaits et à externaliser au maximum les coûts. La souffrance est la même à  $t$  et à  $t_{+1}$ , mais dès lors qu'elle arrive plus tardivement dans un monde plus peuplé, alors cette même souffrance a un poids moindre dans l'équation, c'est-à-dire qu'elle compte moins. De manière plus générale, on perçoit le danger de laisser une population grossir en laissant libres les membres de se délester de toute la souffrance qu'ils peuvent externaliser sur les générations futures. Cela rappelle également d'une certaine manière le problème du conflit de temporalité de la position totale. La remise à plus tard de la lutte contre la crise environnementale est un exemple approprié particulièrement préoccupant et actuel, de même que toutes les luttes qui adressent des risques existentiels.

Face à cet argument, nous dirons qu'il est peut-être possible de résoudre le problème en définissant plutôt  $Q_{PM}$  comme le rapport entre l'utilité totale toutes époques confondues et le poids moral des patients moraux toutes époques confondues. Ce faisant, toutes les temporalités seraient alors prises en compte, ce qui aurait pour conséquence d'amener une sorte d'indépendance de l'utilité vis-à-vis du temps, c'est-à-dire de rendre la notion temporellement neutre. Au lieu de pondérer à la baisse la valeur d'utilité d'un individu en fonction de la quantité

---

<sup>178</sup> En effet, il peut être envisageable sous certaines conditions d'étendre l'agentivité des individus à l'ensemble plus large du règne animal voire à l'ensemble des sentients, tout en pondérant les poids moraux de manière non arbitraire.

d'individus qui existeraient actuellement, cela permettrait de pondérer cette valeur en fonction de *tous* les individus qui existeraient, toutes époques confondues. Cela a en outre l'avantage de permettre d'éviter la Conclusion Répugnante à l'échelle globale. En revanche, il semble que cette définition de  $Q_{PM}$  nous enjoint à voir la position moyenne sous un angle différent. Si la variable est temporellement neutre, alors l'objectif est qu'elle soit complètement maximisée au terme de la dernière vie incluse dans nos considérations morales. Jusqu'à la dernière vie nous avons l'opportunité d'augmenter cette moyenne. Or, s'il nous arrivait un jour d'avoir toutes les raisons de croire que la qualité de vie moyenne a atteint son apogée, alors la question reste ouverte de savoir s'il conviendrait d'arrêter de perpétuer l'humanité pour minimiser la future baisse de la moyenne, quand bien même cette dernière resterait largement positive. À ce point, il est possible que nous retombions également dans l'argument de l'Égyptologie.

Bien que nous ne soyons pas en mesure de résoudre cette difficulté, l'écueil méritait certainement d'être soulevé afin de permettre plus amples discussions à ce sujet. Cela étant fait, nous terminerons cette section par la mise en lumière d'une autre difficulté apparaissant lorsque la population grossit. Plus précisément, il s'agira de critiquer la position moyenne dans son incapacité à expliquer l'intuition largement répandue selon laquelle l'importance des individus dépend de la taille de la population. Dans un souci de cohérence avec ce qui a déjà été dit sur la fiabilité des intuitions communes, nous ne prétendrons donc pas que cette critique est décisive, mais mettrons simplement l'emphase sur cette incapacité flagrante.

## 2. Le problème de la position moyenne lorsque la population est petite

La dernière critique de la position moyenne à laquelle nous nous intéresserons et à laquelle nous tenterons d'apporter une réponse originale concerne les situations où la taille de la population est faible.

On retrouve cette critique dans *Value et Population Size* de Thomas Hurka (1983). S'appuyant notamment sur des extraits de Thomas d'Aquin<sup>179</sup> et de Leibniz<sup>180</sup> ayant trait à la

---

<sup>179</sup> Thomas D'Aquin, «Sommes contre les Gentils» III, 71, 1258-1265.

<sup>180</sup> Gottfried Wilhelm Leibniz, «Théodicée», section 118, 1710.



valeur des populations animales, Hurka soutient avec eux la position d'une valeur toute particulière donnée à l'existence en soi des espèces ainsi qu'à leur variété. Il fait en outre observer à quel point une espèce n'est que la somme de ces individus, de sorte que nous avons des raisons de croire que ladite valeur doit dépendre d'une fonction liée à la valeur personnelle des individus<sup>181</sup>. Or, ces auteurs tiennent pour acquis que la contribution de chaque individu à la valeur du monde n'est pas une constante mais plutôt une variable dépendant de la taille de la population à l'étude. Lorsque le nombre d'individus d'une population est faible, nous avons tendance à attribuer plus de valeur à chaque individu par rapport aux situations dans lesquelles la population foisonne. Cette intuition se manifeste par exemple au travers des divers statuts de conservation définis par la Liste Rouge Mondiale des Espèces Menacées<sup>182</sup> décrivant précisément les conditions périlleuses auxquelles sont exposées les espèces vivantes. Prenons une espèce en « danger critique » d'extinction<sup>183</sup> comme le Rhinocéros de Java qui ne subsiste plus que dans le parc national d'Ujung Kulon et dont la population est inférieure à une centaine d'individus. Ces animaux sont tellement peu nombreux que chaque individu bénéficie des plus grands soins, même lorsque ces derniers sont extrêmement onéreux. Les ressources mobilisées à chaque intervention pour protéger les rhinocéros du braconnage, pour leur prodiguer des soins et pour les aider à se multiplier pourraient être utilisées pour sauver de nombreux humains, de sorte qu'il n'est pas absurde de concéder que la valeur d'un rhinocéros de Java augmente selon une fonction qui dépend de la taille de sa population. Dans ce cas, on observe même qu'il existe un point au-delà duquel la valeur d'un rhinocéros se compare, voire dépasse largement la valeur d'un être humain.

Dans ce contexte, une intuition partagée voudrait que l'augmentation de la taille d'une petite population ait valeur supérieure à l'augmentation de la taille d'une population déjà large. Or, il semble que la position moyenne soit incapable de capturer cette intuition, d'une part car elle refuse d'impliquer la taille de la population dans son axiologie, et d'autre part car elle prescrit qu'amener à l'existence n'est pas une bonne chose si les individus créés ne participent pas à l'augmentation de la moyenne. Une question se pose toutefois, qui semble largement ignorée

---

<sup>181</sup> Thomas Hurka, «Value and Population Size», 1983, p. 496.

<sup>182</sup> Une initiative de l'Union Internationale pour la Conservation de la Nature.

<sup>183</sup> WWF France, «Le rhinocéros, un animal en danger», consulté le 22 juillet 2023.

dans la littérature, est celle de l'échéance à laquelle il doit être possible de dire que les individus créés participent à l'augmentation de la moyenne. Puisqu'à notre connaissance cette question n'a pas été discutée et qu'elle est en lien avec notre discussion précédente sur la définition de  $Q_{PM}$ , nous dirons quelques mots dessus afin de stimuler le débat.

Nous proposons notamment de faire remarquer le phénomène suivant, qui semble représentatif de l'Histoire de notre civilisation. Dans une population donnée, et selon une position moyenne à échéance longue, il semble envisageable de faire baisser la qualité de vie moyenne dans un premier temps en créant des individus sous la moyenne, afin que ceux-ci participent à l'augmentation de la taille de la population. Cette augmentation amènerait au bout du compte une augmentation du niveau de vie général suffisamment importante permettant d'augmenter la moyenne jusqu'à un niveau inatteignable autrement.

Graphiquement, ce scénario pourrait être illustré par une courbe ayant trois parties. La première est croissante et désirable, illustrant le fait qu'en grossissant la population acquière un meilleur niveau de vie moyen. Ce peut être dû à une amélioration sensible de la sécurité du groupe et du développement de sa capacité à s'organiser pour trouver de la nourriture par exemple. Cette augmentation est toutefois d'une forme qui laisse présager une asymptote horizontale après plusieurs générations. La deuxième partie n'est cependant pas aussi réjouissante, puisqu'à partir d'une certaine taille de population, on observe que la qualité de vie décroît significativement. Il est probable que ce point de bascule représente le moment où les ressources naturelles disponibles à proximité deviennent insuffisantes pour nourrir tout le monde, de sorte qu'il faille travailler proportionnellement beaucoup plus pour obtenir moins par personne. On pourrait associer cette phase à l'émergence d'une certaine forme de compétition qui tirerait les salaires vers le bas. Enfin, si la population continue de croître dans ces conditions, amenant à l'existence chaque fois plus d'individus moins bien lotis et participant d'autant plus à faire baisser la moyenne, arrive un moment critique qui change drastiquement le paradigme de cette société, propulsant la qualité de vie en quelques siècles vers des sommets que les ancêtres auraient été vraisemblablement incapables d'imaginer. La découverte d'une nouvelle source d'énergie pourrait être l'une de ces découvertes. On notera que le fait que la population était déjà large ou non au moment de la découverte est d'une importance capitale puisque c'est le

facteur limitant de l'extraction des gisements. Si la population est déjà large, l'exploitation peut être plus rapide et peut entraîner plus rapidement la production de biens et services au service de la population que si la population était petite.

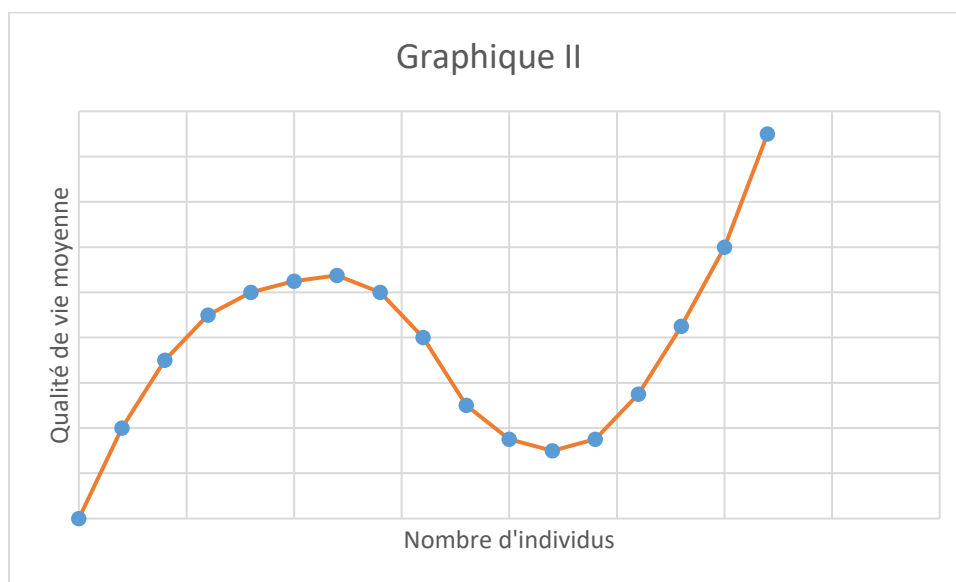


Figure 11. – Graphique II

Si l'on s'intéresse au régime alimentaire, il semble que nous ayant de bonnes raisons de croire que les chasseurs-cueilleurs avaient un niveau de vie tout à fait décent, qui n'a cessé de s'améliorer jusqu'au début de l'Holocène<sup>184</sup>. Peu après est advenue la sédentarisation, et avec elle le début de l'agriculture<sup>185</sup>, une invention que l'on qualifie aujourd'hui de fabuleuse mais qui cache en réalité un coût faramineux ayant probablement contribué à réduire la qualité de vie moyenne durant un temps. En effet, non seulement nos ancêtres ont commencé à travailler de plus en plus longtemps dans des conditions difficiles non adaptées à leur morphologie de primates, mais en plus sont-ils devenus littéralement dépendants de leurs champs, c'est-à-dire directement moins résilients face aux épidémies ou aux phénomènes météorologiques qui pouvaient toucher leurs précieuses céréales. Petit à petit, leur alimentation devenait moins variée, de sorte qu'on peut présumer de multiples carences là où le régime du chasseur-cueilleur

<sup>184</sup> La période interglaciaire actuelle ayant débuté il y a environ 12 000 ans.

<sup>185</sup> Yuval Noah Harari, «Sapiens : A Brief History of Humankind», 2011, p. 94.

se compose d'une multitude d'aliments différents très adaptés à ses besoins physiologiques<sup>186</sup>. Dans le même temps, le nombre de calories produites a été largement optimisé, et sa valeur absolue a été multipliée rapidement par plusieurs ordres de grandeur, permettant de nourrir, moins bien certes mais de manière viable tout de même, bien plus d'individus, ce qui participa à la croissance de la taille de la population. Cette augmentation permettant de résoudre des problèmes plus complexes et des défis collectifs plus importants, on assiste depuis plusieurs siècles à une augmentation spectaculaire de la qualité de vie des individus, augmentation qui contrebalance de beaucoup la diminution de la qualité de vie des premiers agriculteurs. En d'autres termes, si ce qui importe est de maximiser  $Q_{PM}$ , alors il peut être prescrit dans les faits de réduire temporairement  $Q_{PM}$  sur une certaine période dans le cas où l'on aurait de bonnes raisons de croire que cela entraînerait une augmentation plus importante par la suite<sup>187</sup>.

Nous appellerons donc cet argument original la Conclusion Totalitaire, pour faire écho aux autres arguments discutés jusqu'alors. Dans ce cas précis, il conviendrait certainement de chercher à rendre  $Q_{PM}$  temporellement indépendante, c'est-à-dire qu'il semble logique de préférer fonder la qualité de vie moyenne sur le rapport entre le total de bien-être toutes époques confondues et le nombre d'individus toutes époques confondues plutôt que sur des mesures à un instant  $t$  pour déterminer la désirabilité de la création d'une vie particulière au bien-être défini à  $t_{+1}$ .

Une limite à cette première définition dépend cependant des actions qui peuvent être prises à chaque instant  $t$ , de sorte que le fait qu'il est en pratique impossible de déterminer la somme du bien-être toutes époques confondues provient en partie du fait qu'il est impossible de savoir comment les différents futurs affectent le bien-être total du futur. On retrouve là une des critiques faites au à la position dont nous avons discuté au Chapitre I. Dans certaines situations toutefois, il semble largement dans nos cordes d'estimer qu'une action est préférable à une autre, ce qui résout alors le dilemme. Résoudre le changement climatique semble dans la majorité des scénarios imaginable permettre l'épanouissement de l'humanité et bénéficier bien davantage à

---

<sup>186</sup> Yuval Noah Harari, «Sapiens : A Brief History of Humankind», 2011, p. 94.

<sup>187</sup> Le raisonnement fonctionne toutefois également dans le sens inverse, de sorte que la création de certains individus au bien-être supérieur à la moyenne peut être une mauvaise chose si l'on considère la pression sur les ressources et la baisse future de la qualité de vie moyenne, qui pourrait être plus importante que le gain temporaire.

son développement que si elle était amenée à subir une crise environnementale violente. Cela s'explique notamment par le fait que la non-préservation des conditions d'habitabilité ainsi que la destruction des environnements fragiles sont des phénomènes qui participent en général à verrouiller certains futurs désirables, à fermer des opportunités et à entraver certaines conditions nécessaires à la prospérité. S'il existe en théorie des scénarios dans lesquels la crise environnementale est intéressante d'un point de vue civilisationnel parce qu'elle permettrait, disons, de forcer l'humanité à mettre en œuvre des solutions originales pour sauver sa pérennité, efforts qu'elle n'aurait pas réalisés autrement si tout allait bien. Étymologiquement, le terme crise vient du grec *κρίσις* (*krisis*) au sens de faculté de choisir, action de séparer. Lorsqu'elle survient, la crise est l'instant propice à l'examen des conditions contextuelles, elle nous offre l'opportunité de prendre une décision, de juger la situation pour effectuer des changements profonds et faire un tri nécessaire qui idéalement résulterait en une prodigieuse avancée technologique ou philosophique améliorant par la suite la qualité de vie moyenne de manière spectaculaire, de sorte qu'elle incarne l'instant où il semblerait incohérent et irrationnel de foncer tête baissée dans la gueule du loup, ou pour filer la métaphore du chapitre précédent, de mettre tous ses œufs dans le même panier.

De cette section nous retiendrons donc deux choses. D'une part la position moyenne semble en proie à un argument selon lequel les individus seraient moralement incités à catalyser au maximum le bien-être présent tout en remettant à plus tard le maximum de souffrance. Nous avons proposé de pallier cette difficulté en modifiant la définition de la qualité de vie moyenne, mais nous concédons humblement que cette proposition limitée ne parvient pas totalement à éviter l'écueil, de sorte que plus de recherches seraient nécessaires à ce propos. D'autre part, nous avons fait observer que la position moyenne ne parvenait pas à capturer l'intuition ayant donné naissance aux théories à valeur variable. Ce problème avait déjà été largement documenté dans la littérature, et si nous avons tenté d'en faire succinctement l'état des lieux, nous avons choisi de nous concentrer sur la problématique méconnue du choix de l'échéance pertinente. Nous avons à ce titre fait valoir que contrairement aux affirmations de la doctrine moyenne classique, il existerait des conditions sous lesquelles amener à l'existence des individus participants à faire baisser le niveau de vie moyen pourrait être considéré comme désirable.

## Conclusion

Pour conclure, nous avons vu dans ce chapitre que la Conclusion Sadique était le principal argument discuté dans la littérature et utilisé par les détracteurs de la position moyenne. En règle générale, nombreux sont les auteurs qui considèrent que la Conclusion Sadique est une faiblesse tellement plus importante que la Conclusion Répugnante qu'elle conduit sans autre forme de procès à la décrédibilisation de la position moyenne. Toutefois, un examen plus approfondi nous a permis de mettre en évidence que diverses solutions plausibles ont été apportées à ce problème. Notamment, si certaines réponses portaient sur la remise en cause des intuitions et la pertinence de l'adjectif qui la qualifie, d'autres ont choisi d'attaquer directement la comparaison utilisée comme socle de l'argument. Nous avons fait observer que la Conclusion Sadique n'est cependant pas la seule critique adressée à la position moyenne. On retrouve par exemple la critique de Huemer remettant en cause la nature de l'objet valorisé par certaines positions moyennes. Nous avons répondu à l'argument en invoquant une compréhension erronée de ce qui constitue la valeur intrinsèque d'une chose. Parfit a également fait valoir des points comme la Mort Prématuurée et l'Égyptologie. Mais ces arguments ont été mis à mal par Pressman dans ce que nous avons appelé une réponse plus générale à la détermination de la population pertinente. Le dernier argument contre la position moyenne auquel nous avons été capables de donner une réponse plausible s'est avéré être la Conclusion Infernale. Dans ce cas précis, l'argument a été méthodiquement démonté en faisant appel à l'incohérence de l'intuition de neutralité, et doublé d'un argument positif cherchant à rendre compte de la faillibilité de nos intuitions fondamentales.

Enfin, la dernière section s'est avérée moins fructueuse en cela que nous ne sommes pas parvenus à proposer de solutions satisfaisantes et complètes en réponse aux arguments de la remise à plus tard de la souffrance ainsi que de l'incapacité de la position moyenne à capturer l'intuition de la variabilité de la valeur des individus en fonction de la taille de la population. Toutefois, nous avons soulevé une idée qui à notre connaissance n'avait pas été discutée sous cette forme dans la littérature, et qui a trait à l'application réelle de la position moyenne à travers la question de la détermination de l'échéance pertinente à la prise de décision. Nous avons en outre conclu qu'il existait des conditions sous lesquelles amener à l'existence des individus

participants à faire baisser le niveau de vie moyen pourrait être considéré désirable, contre les prescriptions traditionnelles de la position moyenne, de sorte que nous avons nommé cet argument la Conclusion Totalitaire.

## V. Conclusion Générale

Tout au long de ce mémoire, nous avons tenté de chercher à savoir qui de la position totale ou moyenne semblait la plus apte à fournir une axiologie convenable répondant aux défis posés par l'éthique des populations. Il semble en effet possible de dire qu'une population doit être considérée meilleure qu'une autre eu égard à la qualité de vie de ses membres, que ce soit via la qualité de vie totale ou la qualité de vie moyenne. Ces deux axiologies théorisent différemment la nature et la classification de la valeur, définissent ce qui en possède et codifient les comparaisons entre les populations. Nous avons ainsi exploré les raisons morales, sous certaines conditions, d'amener ou non des personnes à l'existence. À première vue il semblait plausible de ne poser comme principe que la règle selon laquelle les vies heureuses ont plus de valeur que les vies misérables. Toutefois les populations réelles sont en perpétuel changement et leur taille varie à tout instant, ce qui complique largement les choses. Nous avons également cherché à comprendre les principes pertinents à la détermination de la taille et de la distribution des populations futures. Pour ce faire, nous nous sommes intéressés à plusieurs avenues discutées dans la littérature. Au fond, il semble que l'importance de ce document réside finalement davantage dans le travail de mise en lumière des différents problèmes philosophiques que dans l'apport d'une résolution à la question de Sidgwick, et il convient de remarquer que plusieurs arguments ou contre-arguments présentés ici sont suffisamment incertains pour que nous estimions probable d'avoir manqué une ou plusieurs considérations cruciales à ce tableau. Nous avons notamment évoqué des raisons et des intuitions qui pourraient nous pousser à désirer garantir la survie à long terme de l'humanité, et présumant du fait que de futures générations nous succéderont, nous avons essayé de circonscrire les fondations ainsi que la portée de nos obligations morales. L'éthique des populations est un champ d'études foncièrement ancré dans le quotidien, dont la résolution pourrait permettre de justifier des pratiques de la vie de tous les jours et guider la prise de décisions à une multitude de niveaux. Pour notre part, nous avons utilisé presque sans distinction des termes axiologiques et normatifs pour rendre compte de cette



réalité, et les avons appliqués presque exclusivement à l'avenir de l'Histoire humaine qui semble grandement dépendre de nos choix.

Il est également entendu que le conflit entre les positions que nous avons abordé met en lumière une véritable tension entre deux objectifs que sont d'une part le contournement de la Conclusion Répugnante et de l'autre l'esquive de la Conclusion Sadique. Ce conflit a été largement discuté dans la littérature, et il nous a paru pertinent de le soulever dans le contexte de plusieurs théories spécifiques.

Une limite évidente de ce document a trait à l'incertitude épistémique dans laquelle nous nous trouvons. Dès lors que nous envisageons que nous sommes extrêmement ignorants des effets à long terme de nos actions, il devient difficile de prétendre le prédire et encore moins de l'influencer. À ce titre, et nous avons exposé nos arguments dans une section dédiée, il peut être tentant de préférer s'attaquer uniquement aux problèmes actuels. Pourtant, nous sommes d'avis que les détracteurs des positions les plus tournées vers le futur devraient à minima concéder qu'en l'absence de certitudes nous avons au moins de bonnes raisons de continuer nos efforts en direction d'une meilleure compréhension des principes pertinents, dans le but de devenir progressivement moins incertains des décisions à prendre. « L'avenir est une nuit noire dans laquelle nous ne pouvons pas nous permettre d'avancer aveuglément » disait régulièrement Gaëtan Selle, il est donc de la responsabilité des éthiciens de devenir les fusées éclairantes de l'humanité.

Sur une note plus personnelle, nous tenions également à exprimer que les positions moyennes et totales semblent en vérité toutes les deux en accord avec nos intuitions axiologiques, même si elles sont incohérentes. Sur ce point, plusieurs choses sont à souligner. D'abord, il nous semble que l'on puisse en vouloir à la simplification excessive de ces théories, à leur essentialisation de la réalité et de ce qui fait le bien-être. En effet, ce type de représentation nous encourage largement à supputer que le bien-être est une quantité scalaire, c'est-à-dire une quantité qui peut être représentée par un nombre réel unique, tandis que « well-being is not like milk: it has many dimensions and is not, as Sen (1980) puts it, a "homogeneous magnitude" »<sup>188</sup>.

---

<sup>188</sup> Jacob Nebel, «Totalism without Repugnance», p. 203.

Pour Nebel, la Conclusion Répugnante est peut-être répugnante parce qu'elle simplifie à outrance ce qui rend la vie digne d'être vécue. Nos intuitions sont complexes et largement influencées par notre culture, c'est-à-dire par des fictions transmises qui nous poussent presque malgré nous à penser de certaines façons, mais également par de nombreux biais cognitifs insidieux, de sorte qu'il reste délicat, alors que peut-être indépassable, de s'en remettre à elles pour fonder de véritables justifications. Dans cette optique, nous avons choisi d'exposer la position justiciste avec l'idée de rappeler au lecteur que la prise en compte de l'existence de valeurs supplémentaires est une avenue à ne pas négliger lorsque vient le temps de bâtir une nouvelle éthique normative. Une autre limite évidente est celle des prémisses que nous avons utilisées tout au long de ce document, et plus particulièrement la prémisse utilitariste. En effet, s'il s'avère que cette doctrine est capable de capturer certaines de nos intuitions, il n'est pas ici question d'affirmer qu'elle est la seule voie disponible. Le problème d'incohérence auquel nous faisons face a donc probablement un lien avec ce choix, de sorte que si l'on s'accorde sur le fait qu'il semble difficile de ne jamais se soucier des conséquences de nos actions, on concédera également qu'il semble parfois déraisonnable d'être utilitariste jusqu'au bout<sup>189</sup>.

Enfin, nous terminerons par dire qu'il n'est peut-être pas si grave d'être incohérent dans une certaine mesure. Nos croyances, nos normes et nos valeurs sont en perpétuelle évolution. Or cette évolution se nourrit justement des incohérences individuelles, car les cultures n'ont de cesse de tenter de concilier ces contradictions. Bien des valeurs auxquelles nous attachons de l'importance sont incompatibles entre elles, mais peut-être ne devrions-nous pas nous en alarmer outre mesure, car c'est justement cette évolution qui donne son élan à la philosophie et à la créativité humaine tout entière, nous obligeant toujours davantage à penser, critiquer et réévaluer.

---

<sup>189</sup> Pour plus de détails à ce sujet, voir Shelly Kagan, «Does Consequentialism Demand Too Much? Recent Work on the Limits of Obligation», 1984.

## Références bibliographiques

Adams, Fred, Laughlin, Greg, Silk, Joseph. 2000, «The Five Ages of the Universe: Inside the Physics of Eternity». *Physics Today* 1 February 2000; 53 (2). <https://doi.org/10.1063/1.882973/>

Armstrong, Stuart. 2014, «Embracing the “Sadistic” Conclusion | Practical Ethics». <http://blog.practicaethics.ox.ac.uk/2014/02/embracing-the-sadistic-conclusion/>

Arrhenius, Gustaf. 2000, «An Impossibility Theorem for Welfarist Axiologies» *Economics & Philosophy* 16, no. 2 : 247–66. <https://doi:10.1017/S0266267100000249/>

Arrhenius, Gustaf. 2000a, «Future Generations: A Challenge for Moral Theory». F.D.-Diss. Uppsala: University Printers.

Asheim, Geir, Zuber, Stéphane. 2014, «Escaping the repugnant conclusion: Rank-discounted utilitarianism with variable population». *Theoretical Economics*, 9: 629-650. <https://doi.org/10.3982/TE1338/>

Bader, Ralf. 2022, «Person-affecting utilitarianism». In Gustaf Arrhenius, Krister Bykvist, Tim Campbell & Elizabeth Finneron-Burns (eds.), *The Oxford Handbook of Population Ethics*. Oxford University Press.

Barker, Sam, Pearce, David. 2020, «Sam Barker and David Pearce on Art, Paradise Engineering, and Existential Hope (With Guest Mix)», <https://futureoflife.org/podcast/sam-barker-and-david-pearce-on-art-paradise-engineering-and-existential-hope-featuring-a-guest-mix/>

Barnosky, Anthony, Matzke, Nicholas, Tomiya, Susumu et al. 2011, «Has the Earth’s sixth mass extinction already arrived?». *Nature* 471. <https://doi.org/10.1038/nature09678>

Beauchamp, Tom. 1997, «Justifying Physician-Assisted Deaths». In LaFollette, H. (ed.), *Ethics in Practice: An Anthology* (5th ed.). <http://hdl.handle.net/10822/899723>

Benatar, David. 2008, «Better Never to Have Been: The Harm of Coming Into Existence», Oxford University Press.

Bentham, Jeremy. 1789, «An Introduction to the Principles of Morals and Legislation». Bennett, J. (ed.).

Bennett, Jonathan. 1995, «The act itself». New York: Oxford University Press. <https://doi.org/10.1093/019823791X.001.0001.002.003>

Blackorby, Charles, Donaldson, David. 1984, «Social criteria for evaluating population change», *Journal of Public Economics*, Volume 25, Issues 1–2, ISSN 0047-2727. [https://doi.org/10.1016/0047-2727\(84\)90042-2](https://doi.org/10.1016/0047-2727(84)90042-2)

Boonin-Vail, David. 1996, «Don't Stop Thinking About Tomorrow: Two Paradoxes about 'Duties to Future Generations'», *Philosophy and Public Affairs* 25.

Bossert, Walter, Susumu Cato, and Kohei Kamaga. 2022, «Revisiting Variable-Value Population Principles» *Economics & Philosophy*. <https://doi:10.1017/S0266267122000268/>

Boström, Nick. 2019, «The Vulnerable World Hypothesis». *Glob Policy*, 10. <https://doi.org/10.1111/1758-5899.12718>

Boström, Nick. 2014, «Superintelligence: Paths, dangers, strategies». Oxford University Press.

Boström, Nick. 2002, «Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards» *Journal of Evolution and Technology* 9.

Bovee, Warren. 1991, «The End Can Justify the Means--But Rarely», *Journal of Mass Media Ethics*, 6:3, [https://doi:10.1207/s15327728jmme0603\\_1/](https://doi:10.1207/s15327728jmme0603_1/)

Brentano, Franz. 2009, «The Origin of Our Knowledge of Right and Wrong» (Routledge Revivals) (1st ed.). Routledge. <https://doi.org/10.4324/9780203092248/>

Broome, John. 2004, «Weighing Lives» (Oxford, 2004; online edn, Oxford Academic, 1 Oct. 2005). <https://doi.org/10.1093/019924376X.001.0001/>

Chang, Ruth. 2002, «The Possibility of Parity» *Ethics* 112, no. 4. <https://doi.org/10.1086/339673/>

Chappell, Richard, Yetter, Meissner, Daniel, et MacAskill, William. 2022, «Population Ethics». In R.Y. Chappell, D. Meissner, and W. MacAskill (eds.), *An Introduction to Utilitarianism*, <https://www.utilitarianism.net/population-ethics/>, Page consultée le 22 juillet 2023.

Contributeurs de Wikipédia, «Long-termisme», Wikipédia, l'encyclopédie libre, <https://fr.wikipedia.org/w/index.php?title=Long-termisme&oldid=205750953/>, Page consultée le 22 juillet 2023.

Contributeurs de Wikipédia, «Future of Earth», Wikipédia, l'encyclopédie libre, [https://en.wikipedia.org/wiki/Future\\_of\\_Earth/](https://en.wikipedia.org/wiki/Future_of_Earth/), Page consultée le 22 juillet 2023.

Contributeurs de Wikipédia, «Population Ethics», Wikipédia, l'encyclopédie libre, [https://en.wikipedia.org/wiki/Population\\_ethics/](https://en.wikipedia.org/wiki/Population_ethics/), Page consultée le 22 juillet 2023.

Cowen, Tyler. 1996, «What Do We Learn from the Repugnant Conclusion?» Ethics 106, no. 4.  
<http://www.jstor.org/stable/2382033/>

Daniels, Norman. 2020, «Reflective Equilibrium», The Stanford Encyclopedia of Philosophy (Summer 2020 Edition), Edward N. Zalta (ed.), URL =  
<https://plato.stanford.edu/archives/sum2020/entries/reflective-equilibrium>.

Feldman, Fred. 1995, «Justice, Desert, and the Repugnant Conclusion» Utilitas 7, no. 2.  
<https://doi:10.1017/S095382080000203X/>

Fin Moorhouse. 2021, «Longtermism: An Introduction | Effective Altruism»,  
<https://www.effectivealtruism.org/articles/longtermism>, Page consultée le 22 juillet 2023.

Fisher, Richard. 2022, «The Perils of Short-Termism : Civilisation’s Greatest Threat» . BBC Future.  
<https://www.bbc.com/future/article/20190109-the-perils-of-short-termism-civilisations-greatest-threat/>

Foot, Mike, and David M. Raup. 1996, «Fossil Preservation and the Stratigraphic Ranges of Taxa» Paleobiology 22, no. 2. <https://doi:10.1017/S0094837300016134/>

Fraser, Chris. 2016, «The Philosophy of the Mòzi: The First Consequentialists», Columbia University Press, ISBN-10 : 9780231149273.

Greaves, Hilary. 2017, «Population Axiology». Philosophy Compass. 12(11).

Greshko, Michael. 2017, «These Early Humans Lived 300,000 Years Ago—But Had Modern Faces»,  
<https://www.nationalgeographic.com/history/article/morocco-early-human-fossils-anthropology-science/>

Grill, Kalle. 2023, «The Sum of Averages: An Egyptology-Proof Average View» *Utilitas* 35, no. 2. <https://doi:10.1017/S0953820822000450/>

Gustafsson, Johan. 2020, «Population Axiology and the Possibility of a Fourth Category of Absolute Value» *Economics & Philosophy* 36, no. 1. <https://doi:10.1017/S0266267119000087/>

Fugulin, Vali, Saint-Onge, Marc, Landry, Stéphane. 2011, « Nauffragés des villes | E02 | L'itinérance ».

Harari, Yuval Noah. 2015, «Sapiens : A Brief History of Humankind». New York, NY: Harper.

Harsanyi, John C. 1953, «Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking», *Journal of Political Economy* 61, no. 5. <http://www.jstor.org/stable/1827289>

Heyd, David. 1988, «Procreation and value can ethics deal with futurity problems?», *Philosophia* 18. <https://doi.org/10.1007/BF02380074/>

Hoang, Lê Nguyễn, Science4All. 2023, «#ChatGPT, Top Gun, CR7, c'est... cool ? #CoolWashing #MuteNews» YouTube. <https://youtu.be/E9c2y0eUBMs/>

Huemer, Michael. 2008, «In Defence of Repugnance» *Mind* 117, no. 468. <http://www.jstor.org/stable/20532700/>

Hurka, Thomas. 1982, «Average Utilitarianisms» *Analysis* 42, no. 2. <https://doi.org/10.2307/3327924/>

Hurka, Thomas. 1982, «More Average Utilitarianisms» *Analysis* 42, no. 3. <https://doi.org/10.2307/3327578/>

Hurka, Thomas. 1983, «Value and Population Size» Ethics 93, no. 3. <http://www.jstor.org/stable/2380627/>

Kagan, Shelly. 1984, «Does Consequentialism Demand Too Much? Recent Work on the Limits of Obligation» Philosophy & Public Affairs 13, no. 3. <http://www.jstor.org/stable/2265413/>

Kahneman, Daniel, Deaton, Angus. 2010, «High income improves evaluation of life but not emotional well-being» Proc Natl Acad Sci U S A. 107(38):16489-93. <https://doi:10.1073/pnas.1011492107/>

Kaneda, Toshiko, Haub, Carl. 2022, «How Many People Have Ever Lived on Earth?», <https://www.prb.org/articles/how-many-people-have-ever-lived-on-earth/>, Page consultée le 22 juillet 2023.

Kavka, Gregory. 1982, «The Paradox of Future Individuals» Philosophy & Public Affairs 11, no. 2. <http://www.jstor.org/stable/2264924/>

Khan Academy, «First Humans : Homo Sapiens & Early Human Migration (Article)», <https://www.khanacademy.org/humanities/world-history/world-history-beginnings/origin-humans-early-societies/a/where-did-humans-come-from>, Page consultée le 22 juillet 2023.

Kosonen, Petra. 2022, «Tiny Probabilities of Vast Value» PhD thesis, University of Oxford.

La Fabrique Sociale. 2023, «Superorganisme : la seule transition possible pour l'humanité ?», YouTube, <https://youtu.be/Uju01jB0pTs/>

Lambeth, Edmund. 1986, «Committed Journalism: An Ethic for the Profession», Volume 364 de Midland Books: No. 364, Indiana University Press.



Lichtenstein, Sarah, Slovic, Paul, Fischhoff, Baruch, Layman, Mark, Combs, Barbara. 1978, «Judged frequency of lethal events. Journal of Experimental Psychology: Human Learning and Memory», 4(6). <https://doi.org/10.1037/0278-7393.4.6.551/>

MacAskill, William. 2022, «What We Owe the Future». Basic Books.

Mankiw, Gregory, Taylor, Mark (trad. de l'anglais). 2015, «Principes de l'économie 4e édition», Louvain-La-Neuve/Paris, De Boeck, ISBN 978-2-8041-9306-5.

McMahan, Jefferson. 1981, «Review of Problems of Population Theory», Ethics 92, no. 1. <http://www.jstor.org/stable/2380707/>

Narveson, Jan. 1967, «Utilitarianism and New Generations», Mind 76 (1967).

Nathan Franz, Dean Spears. 2020, «Mere Addition is equivalent to avoiding the Sadistic Conclusion in all plausible variable-population social orderings», Economics Letters, Volume 196, 109547, ISSN 0165-1765, <https://doi.org/10.1016/j.econlet.2020.109547/>

Nebel, Jacob. 2022, «Totalism without Repugnance». In Jeff McMahan, Tim Campbell, James Goodrich & Ketan Ramakrishnan (eds.), Ethics and Existence: The Legacy of Derek Parfit. Oxford: Oxford University Press.

Nozick, Robert. 1974, «Anarchy, State, and Utopia». New York: Basic Books.

Ord, Toby. 2020, «The Precipice : Existential Risk and the Future of Humanity». New York: Hachette Books.

Orwell, Georges. 1949, «Nineteen Eighty-Four».

Parfit, Derek. 1986, «Reasons and Persons» (Oxford, 1986; online edn, Oxford Academic, 1 Nov. 2003). <https://doi.org/10.1093/019824908X.003.0018/>

Pressman, Michael. 2015, «A Defence of Average Utilitarianism» *Utilitas* 27, no. 4. <https://doi:10.1017/S0953820815000072/>

Rawls, John. 1971, «A Theory of Justice: Original Edition», Harvard University Press. <https://doi.org/10.2307/j.ctvjf9z6v>

Saumon, Frédéric. 2014-2015, «Manifestation de suffragettes à Paris, place de la Bastille en 1935», *Histoire, Géographie, EMC – Strabon*, [https://histoire.ac-versailles.fr/IMG/html/louise\\_weiss2.html/](https://histoire.ac-versailles.fr/IMG/html/louise_weiss2.html/), Page consulté le 22 juillet 2023.

Selle, Gaëtan, The Flares. 2022, «Augmenter le bien-être via les biotechnologies avec David Pearce», Youtube, <https://youtu.be/V64xZPT6s9k/>

Selle, Gaëtan, The Flares. 2023, «Influencer le FUTUR LOINTAIN : est-ce possible ?», YouTube, <https://youtu.be/AOTiYMOCfY4/>

Selle, Gaëtan, The Flares. 2023, «La chronologie du FUTUR LOINTAIN : repenser les horizons temporels», YouTube, <https://youtu.be/Gx6HbANWgMI/>

Selle, Gaëtan, The Flares. 2023, «L'éthique du FUTUR LOINTAIN : les implications morales de nos décisions ?», YouTube, <https://youtu.be/vplcgBR3SSs/>

Selle, Gaëtan, The Flares. 2023, «Les progrès de l'I.A., singularité et futurisme – avec Anders Sandberg», YouTube, <https://youtu.be/K7RwE3zOHPY/>

Setiya, Kieran. 2022, «The New Moral Mathematics - Boston Review», <https://www.bostonreview.net/articles/the-new-moral-mathematics/>, Page consultée le 22 juillet 2023.

Sidgwick, Henry. 1907, «The Methods of Ethics (7th ed.)». United Kingdom. pp. Book 4, chapter 1, section 2.

Singer, Peter. 2011, «The Expanding Circle: Ethics, Evolution, and Moral Progress». Princeton University Press.

Spears, Dean, Budolfson, Mark. 2018, «Why the Repugnant Conclusion is Inescapable», Princeton University Climate Futures Initiative working paper.

Stanford Encyclopedia of Philosophy. 2017, «The repugnant conclusion (Stanford Encyclopedia of Philosophy)», <https://plato.stanford.edu/entries/repugnant-conclusion/>, Consultée le 22 juillet 2023.

Subramanian, Sreenivasan. 2023, «The Repugnant, the Sadistic, and Two “Despotic” Conclusions in Population Ethics», in « Working Papers | Chair in Public Finance | Victoria University of Wellington». Victoria University of Wellington.

Tännsjö, Torbjörn. 2020, «Why Derek Parfit Had Reasons to Accept the Repugnant Conclusion» *Utilitas* 32, no. 4. <https://doi:10.1017/S0953820820000102/>

Teruji, Thomas. 2018, «Some Possibilities in Population Axiology», *Mind*, Volume 127, Issue 507. <https://doi.org/10.1093/mind/fzx047/>

Todd, Benjamin. 2017, «Longtermism : The Moral Significance of Future Generations - 80,000 Hours»,

<https://web.archive.org/web/20230314130427/https://80000hours.org/articles/future-generations/>, Page consultée le 22 juillet 2023.

United Nations. 2022, «Day of Eight Billion, 15 November 2022 | Population Division», <https://www.un.org/development/desa/pd/events/day-eight-billion>, Page consultée le 22 juillet 2023.

Vickrey, William. 1945, «Measuring Marginal Utility by Reactions to Risk», *Econometrica* 13, no. 4. <https://doi.org/10.2307/1906925>

WWF France, «Le rhinocéros, un animal en danger | WWF France», <https://www.wwf.fr/especes-prioritaires/rhinoceros/>, Page consultée le 22 juillet 2023.