



**HAL**  
open science

## L'utilitarisme: les cercles concentriques d'une définition

Claude Gamel

► **To cite this version:**

Claude Gamel. L'utilitarisme: les cercles concentriques d'une définition. *Économies et sociétés. Série PE, Histoire de la pensée économique*, 1999, 28 (4), pp.101-138. halshs-02520678

**HAL Id: halshs-02520678**

**<https://shs.hal.science/halshs-02520678v1>**

Submitted on 9 Aug 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **L'utilitarisme : les cercles concentriques d'une définition\***

Claude Gamel

GREQAM – Aix-Marseille III

*L'économie normative est directement concernée par la doctrine utilitariste, en tant qu'éthique à prétention scientifique. Cette recherche d'une véritable « arithmétique morale » ne dissipe pas pour autant une confusion persistante au sujet de la définition de l'utilitarisme. On soutiendra ici que cette confusion, source supplémentaire de controverses, résulte de l'existence de cadres d'analyse plus ou moins larges, emboîtés les uns dans les autres : trois définitions de plus en plus restreintes semblent ainsi coexister, sans que l'acceptation de la première n'implique l'adhésion aux suivantes.*

*Normative economics is directly concerned by utilitarianism, as ethics with scientific claim. The will to define a real "moral arithmetic" does not allow to clear up a persistent confusion about the definition of utilitarianism. Here it will be defended the fact that this confusion is a supplementary source of debates ; it results from the existence of analytical frameworks which are more or less broad and fitted together : three more and more restricted definitions of utilitarianism seem to coexist but the agreement on the first does not imply the same agreement on the others.*

---

\* Cet article constitue une version révisée d'une communication présentée au Colloque « L'utilitarisme : analyse et histoire » (Association Charles GIDE pour l'Étude de la Pensée Économique et Université Lille I) – Roubaix – 25 et 26 janvier 1996. Je remercie par ailleurs les rapporteurs de la revue *Économies et Sociétés* pour leurs précieux commentaires sur la première version de ce texte, mais reste évidemment seul responsable des erreurs et imprécisions qui pourraient subsister.

## INTRODUCTION : UNE ÉTHIQUE À PRÉTENTION SCIENTIFIQUE

Dans l'architecture de la pensée utilitariste, la norme de justice sociale (« le plus grand bonheur du plus grand nombre ») peut être considérée comme la clé de voûte, par rapport à laquelle se conçoivent et se situent les matériaux peu nombreux mais rigoureusement fondés d'une réflexion à la fois cohérente et systématique. Telle est du moins la perspective ici soutenue. Tel se révèle aussi l'aspect doctrinaire de la pensée utilitariste : dès ses origines son ambition est en effet de déterminer un nombre aussi réduit que possible de « principes » généraux et simples, induits d'observations sur les comportements humains dominants. Une fois découverts, ces principes, en fonction desquels tout le détail des comportements sociaux sera jaugé, pourront servir à la constitution d'une science synthétique des mœurs et de la législation, fondement d'une organisation rationnelle de la société.

Le recul permis par près de deux siècles de développement de la réflexion en la matière, ainsi que l'application pour le moins controversée des recommandations utilitaristes<sup>1</sup>, auraient pu de nos jours dévaloriser cette doctrine en repoussoir de la réflexion philosophique. Or c'est presque la constatation inverse qui s'impose : les utilitaristes contemporains restent fort nombreux. Harsanyi, Arrow, Mirrlees, Hammond en sont aujourd'hui quelques-uns des défenseurs les plus convaincus<sup>2</sup> ; leur assurance est telle que certains d'entre eux considèrent Rawls comme l'un des leurs, malgré l'énergie déployée par ce dernier pour se distinguer des premiers<sup>3</sup>. Hayek subit aussi – à un moindre degré – la même mésaventure, bien que son œuvre comporte un vigoureux réquisitoire contre les ravages qu'il attribue à la pensée utilitariste<sup>4</sup>.

Quintessence méconnue de la réflexion sociale pour certains, étiquette presque infamante pour d'autres, la doctrine utilitariste, pour le moins, ne laisse pas indifférents ceux qui sont amenés à l'étudier.

<sup>1</sup> À titre d'exemples, voir notamment Hahn (1982) et Hampshire (1982).

<sup>2</sup> Harsanyi (1977), Arrow (1973), Mirrlees (1982), Hammond (1982).

<sup>3</sup> « J'examine aussi, pour opérer un contraste éclairant, les conceptions classiques utilitaristes et intuitionnistes et j'étudie certaines des différences entre ces points de vue et la théorie de la justice comme équité. Mon objectif est d'élaborer une théorie de la justice qui soit une solution de rechange à ces doctrines qui ont dominé depuis longtemps notre tradition philosophique », Rawls (1987) p. 29 ; « J'établis dans le présent article plus explicitement ce qui peut être appelé une version ordinaliste de l'utilitarisme et je soutiens que, dans ces termes, la position de Rawls n'en diffère pas beaucoup », Arrow (1973) p. 247.

<sup>4</sup> Voir notamment Hayek (1981), chapitre 7, « L'erreur constructiviste de l'utilitarisme ». Langlois, (1982) p. 286, ne voit pas pourtant « d'incompatibilité complète » entre l'utilitarisme de règle et la pensée de Hayek.

L'attrait que les uns et les autres partagent à son égard, provient probablement du fait que tous considèrent l'utilitarisme comme la première tentative de formulation à prétention scientifique des jugements éthiques, dont sont issus de près ou de loin les débats contemporains en la matière. À l'époque de Bentham, il ne s'agissait pas moins que « de faire, pour la philosophie morale, ce que Newton [avait] fait pour la philosophie naturelle »<sup>5</sup>, c'est-à-dire fonder « une science de l'âme et une science de la société »<sup>6</sup> analogues, par leur rigueur, à la physique newtonienne ; de nos jours, on se contente surtout de décortiquer les prémices et les étapes du raisonnement, soit pour en déceler les failles, soit pour en affiner les éléments.

La volonté affichée de définir avec précision une véritable « arithmétique morale » n'a pas pour autant permis de dissiper une confusion persistante, lorsqu'il s'agit de poser une définition de l'utilitarisme acceptable par tous. Cette confusion, source supplémentaire de controverses, résulte à nos yeux de l'existence dans la pensée utilitariste de cadres d'analyse plus ou moins larges, emboîtés les uns dans les autres, qui ont chacun engendré des débats spécifiques.

Selon le plan où se situe la réflexion, trois définitions de plus en plus restreintes semblent coexister, sans que l'acceptation de la première n'implique forcément l'adhésion aux suivantes : au niveau individuel, la morale personnelle de l'utilité (I) peut conduire à l'arithmétique interpersonnelle des plaisirs et des peines (II), laquelle permet, si on l'accepte, l'énoncé de la norme controversée du plus grand bonheur (III).

#### I. — LA MORALE PERSONNELLE DE L'UTILITÉ ET SES MODALITÉS D'APPLICATION

La doctrine utilitariste, dans sa formulation première (Bentham), est d'abord une tentative d'évaluation rationnelle de la moralité de chaque acte individuel ; mais la prise de conscience des imperfections et des excès de cette première version fondée sur l'acte (A) a fait plus récemment évoluer la doctrine en un utilitarisme conçu en situation de risque<sup>7</sup> (B), avant qu'un nouveau démembrement n'accroisse sa mutation vers la définition d'un utilitarisme, générique, des règles morales<sup>8</sup> (C). Ces deux reformulations contemporaines ont pour point commun de centrer la réflexion sur des aspects « procéduraux » de la justice

<sup>5</sup> Halévy (1995), vol. 1, p. 17.

<sup>6</sup> Halévy (1995), vol. 1, p. 14.

<sup>7</sup> Harsanyi (1953), (1955), (1977).

<sup>8</sup> Brandt (1959).

sociale totalement ignorés dans la version originelle, à laquelle, en ce sens, elles s'opposent radicalement.

Bien que, chronologiquement, seule la première version nous importe ici, en tant que fondement de la doctrine, il semble opportun de lui juxtaposer les ramifications récentes, afin d'en souligner, malgré leurs divergences de fond, l'origine commune et la filiation naturelle.

### I. – 1. L'utilitarisme d'acte

On peut *a priori* considérer avec Rawls que « les deux concepts principaux de l'éthique sont ceux du juste et du bien... La structure d'une théorie éthique est donc largement déterminée par la manière dont elle définit et relie entre elles ces deux notions de base. Or il semble que la façon la plus simple de les relier est celle qu'adoptent les théories téléologiques : le bien est défini indépendamment du juste et, ensuite, le juste est défini comme ce qui maximise le bien »<sup>9</sup>.

Ce cadre général propre aux doctrines téléologiques est celui dans lequel se situe indubitablement l'utilitarisme, au moins à ses origines : il pose le bien en antériorité logique sur le juste, la justice ne relevant alors que d'une rationalité instrumentale permettant de maximiser le bien<sup>10</sup>.

#### *Une doctrine téléologique*

Unies par une démarche formelle commune, les doctrines téléologiques ne se différencieront entre elles que par le caractère exclusif du bien, que chacune retiendra<sup>11</sup>. Pour ce qui est de la doctrine utilitariste, le critère du bien est entièrement contenu dans le principe de l'utilité considéré par Bentham et ses disciples comme l'analogue en matière de philosophie morale du principe newtonien de l'attraction universelle : « La nature a placé l'humanité sous le gouvernement de deux maîtres souverains, la peine et le plaisir. C'est à eux seuls de montrer ce que nous devons faire aussi bien que ce que nous ferons. La distinction du juste et de l'injuste, d'une part, et, d'autre part, l'enchaînement des causes et des effets, sont attachés à leur trône. Le principe de l'uti-

<sup>9</sup> Rawls (1987), p. 50.

<sup>10</sup> Par opposition à une doctrine téléologique, une théorie déontologique est pour Rawls « une théorie qui soit ne définit pas le bien indépendamment du juste, soit n'interprète pas le juste comme une maximisation du bien », Rawls (1987) p. 55.

<sup>11</sup> À titre d'exemples, Rawls distingue, outre l'utilitarisme, le perfectionnisme où le critère du bien correspond à la « réalisation de ce qu'il y a d'excellent dans l'homme, dans les différentes formes de culture », l'hédonisme fondé sur le plaisir et l'eudémonisme sur le bonheur, Rawls (1987) p. 51.

lité constate cette sujétion et la prend pour fondement du système dont l'objet est d'élever l'édifice de la félicité par la main de la raison et de la loi »<sup>12</sup>.

Ainsi le bien, désigné comme « l'utilité », prend-il à ses origines une forme essentiellement hédoniste qui exprime la loi à la fois nécessaire – ce qui est – et morale – ce qui doit être – de toute action humaine : sera juste au niveau individuel le comportement qui maximisera la somme des plaisirs et des peines. Ce faisant, l'individu lui-même est seul juge de son plaisir, c'est-à-dire seul juge du souverain bien.

Bien que systématiquement utilisé par Bentham, c'est à un philosophe non utilitariste, Hume, que le principe de l'utilité semble avoir été emprunté. Dès 1738, celui-ci situe en termes très évocateurs le rôle du plaisir comme finalité de l'activité humaine : « demandez à un homme « pourquoi il prend de l'exercice », il répondra, « parce qu'il désire conserver sa santé » ; si vous demandez, alors, « pourquoi il désire la santé », il répondra sans hésiter, « parce que la maladie est pénible ». Si vous poussez plus loin votre enquête, et demandez à savoir « pour quelle raison il hait la peine », il est impossible qu'il vous en donne jamais une. C'est là une fin dernière et qui n'est jamais rapportée à un autre objet »<sup>13</sup>.

Ce passage souligne les deux traits dont la juxtaposition caractérise la pensée utilitariste dès ses origines :

– toute action humaine sera moralement jugée non en fonction des intentions qui l'animent mais à travers les conséquences qu'elle engendre,

– le degré de satisfaction, que ces effets induits procurent à l'individu, en sera le critère d'évaluation unique.

#### *Un hédonisme « conséquentialiste »*

L'aspect « conséquentialiste »<sup>14</sup> de l'utilitarisme traduit avant tout un certain pragmatisme, tandis que son aspect hédoniste en souligne le caractère individualiste.

L'évaluation rationnelle du contenu moral des actions humaines suppose d'abord l'existence de repères simples susceptibles de faciliter cette évaluation. Tel n'est pas le cas de l'intention qui précède l'acte de l'individu ; cette intention peut être ouvertement connue ou délibéré-

<sup>12</sup> Bentham, « *Introduction aux principes de la Morale et de la Législation* » (1789), cité par Halévy (1995) vol. 1 p. 37.

<sup>13</sup> Hume, « *Enquête sur l'entendement humain* » (1738), cité par Halévy (1995) vol. 1 p. 22.

<sup>14</sup> Ce néologisme n'est que la traduction peu satisfaisante du terme « conséquentialism », utilisé par Sen et Williams (1982) p. 3.

ment masquée, réelle ou superficielle, détournée de son objet ou amplifiée par l'action, autant d'obstacles nuisibles à son repérage explicite.

Au-delà de cet aspect pragmatique, la substitution des conséquences aux intentions est cohérente avec l'ambition utilitariste de codification rationnelle de la morale. On ne peut à la fois fonder une éthique scientifique et continuer de raisonner à partir d'une morale « intuitionniste » préétablie ; il n'est donc pas d'intentions mauvaises *a priori*, qu'elles obéissent à des motifs égoïstes ou philanthropiques, mesquins ou nobles, envieux ou bienveillants, qu'elles respectent ou non des prescriptions bibliques ou des coutumes ancestrales. Le critère essentiel est de savoir si ces actes engendrent des phénomènes sociaux désirables. L'affirmation souvent avancée que « la fin ne justifie pas les moyens », les utilitaristes auraient donc tendance à répliquer que « seuls les objectifs atteints peuvent justifier les moyens ».

C'est dans ce contexte que s'éclaire notamment la virulence de la critique utilitariste à l'égard de Mandeville qui, dans sa « Fable des Abeilles » (1723), remarquait que ce sont des vices humains moralement condamnés qui sont paradoxalement à l'origine de la cohésion de la société : pour le philosophe utilitariste, le paradoxe n'existe, superficiellement, que par les mots choisis pour le décrire, mots qui traduisent une morale confuse voire erronée, largement supplantée par la nouvelle science – utilitariste – de l'éthique.

Pourquoi en effet persister à tenir l'égoïsme pour un vice si, d'une part, l'égoïsme est utile au bien public et si, d'autre part, on convient d'appeler vertueuses les qualités individuelles utiles à la collectivité ? Sur ce point, l'utilitarisme se trouve donc en opposition avec les grandes morales traditionnelles et notamment religieuses, opposition que son caractère hédoniste contribue à renforcer.

L'aspect hédoniste de la morale utilitariste souligne en effet le caractère individualiste de cette doctrine téléologique, du moins au stade préliminaire de l'éthique personnelle qu'elle préconise. Peu importe que les actions humaines servent ou non, de manière incidente, le progrès des connaissances, la diffusion d'une idéologie, ou la promotion d'une nation, d'un Dieu, d'un dictateur ; de telles conséquences comptent d'un poids négligeable dès lors que leur principal effet est de satisfaire les désirs des individus.

Cette priorité marquée à l'égard des aspirations individuelles trouve sa justification dans leur autonomie supposée : de la capacité de choisir de manière indépendante dépend la valeur de référence des désirs exprimés et leur force morale ; faisant de l'homme quel qu'il soit le seul juge possible de son bien, la réflexion utilitariste accepte donc implicitement le principe de « l'individualisme méthodologique » pour lequel les

désirs spontanés de chacun ne peuvent être la traduction même partielle de l'environnement social ou *a fortiori* d'une coercition totalitaire<sup>15</sup>.

Peu souvent évoquée<sup>16</sup>, cette hypothèse est pourtant indispensable à la cohérence de la doctrine, afin de lui éviter une cyclicité perverse : si l'étude du processus de création des désirs personnels révélait l'importance prédominante du conditionnement social, la définition préalable du principe déterministe de l'utilité deviendrait en effet incongrue.

Cependant ce n'est pas cet aspect foncièrement individualiste de la morale de l'utilité, mais trois types d'imperfections, dénoncées dès ses origines, qui sont plus récemment la source d'un néo-utilitarisme conçu en termes de risque.

### I. – 2. L'utilitarisme de risque

Les objections les plus radicales à la doctrine utilitariste concernent la nature des plaisirs et des peines retenus comme critère d'évaluation morale : l'une fut émise par Hume, les deux autres par John Stuart Mill.

#### *Les objections à l'utilitarisme classique*

Chronologiquement la plus ancienne, la critique de Hume est une opposition de principe à l'égard de l'ambition éthique de l'utilitarisme qui démarque nettement son auteur de la doctrine de Bentham. Cette opposition – que Blaug qualifie de « guillotine de Hume » – est entièrement contenue dans la proposition que « l'on ne peut déduire ce qui doit être de ce qui est »<sup>17</sup> : des énoncés descriptifs factuels, même éclairés par le principe de l'utilité dont Hume est pourtant le précurseur, ne peuvent impliquer que d'autres énoncés descriptifs et jamais de normes, de jugements éthiques prescrivant de faire ou de ne pas faire. La recherche des principes qui gouvernent le comportement humain n'autorise pas pour autant à enjoindre ou à proscrire. Ainsi Hume adhère-t-il au principe de l'utilité comme loi nécessaire de l'action des hommes mais n'accepte pas pour autant de l'ériger en loi morale.

En assimilant pourtant l'une à l'autre, l'utilitarisme aboutit à juger moralement satisfaisant ce que l'individu perçoit comme un plaisir alors que nulle garantie n'existe pour considérer les plaisirs personnels comme systématiquement compatibles avec le bien commun ; en

<sup>15</sup> Cette remarque ne préjuge en rien des conclusions que tirent notamment les premiers auteurs utilitaristes de l'improbable « identité naturelle » des intérêts individuels ; de ce fait il conviendra de pratiquer une « harmonisation artificielle », le plus souvent par la contrainte de la loi (cf. *infra* II-B-1).

<sup>16</sup> À l'exception notamment d'Elster (1982) p. 219.

<sup>17</sup> Hume, « *Traité sur la nature humaine* » (1732), cité par Blaug (1982) p. 109.



d'autres termes, « si l'objet naturel de mes désirs, c'est mon plaisir, si l'objet naturel de mes aversions, c'est ma douleur, comment concevoir que le sens moral, qui m'inspire de poursuivre l'utilité générale, et non mon intérêt privé, fasse partie de ma nature ? »<sup>18</sup>.

Cette critique externe de Hume trouve son prolongement dans la controverse, interne à la doctrine classique, sur la qualité des plaisirs à prendre en compte. John Stuart Mill en est à l'origine lorsqu'il s'interroge, d'abord discrètement puis de plus en plus ouvertement, sur le « réductionnisme » de la doctrine de Bentham et de son propre père James Mill. Pour ces derniers en effet, les actions humaines ne peuvent être comparées que sur la base des intensités de plaisirs que chacune procure ; il n'existe pas entre ces actions de différences qualitatives, quelle qu'en soit l'origine : intérêt immédiat ou idéal absolu, plaisir partagé ou désir solitaire. Selon la célèbre formule de Bentham, « les quantités de plaisir étant égales, jouer à la poussette est aussi bon que s'adonner à la poésie ».

Ce nivellement qualitatif des plaisirs aboutit pour John Stuart Mill à supposer que les êtres humains ne sont pas capables d'éprouver d'autres plaisirs que ceux que peut éprouver un cochon. Or, les êtres humains ayant des facultés plus élevées que les « appétits animaux », « on ne connaît pas une seule théorie épicurienne de la vie qui n'assigne aux plaisirs que nous devons à l'intelligence, à la sensibilité, à l'imagination et aux sentiments moraux une bien plus haute valeur comme plaisirs qu'à ceux que procure la pure sensation »<sup>19</sup>.

De cette critique générale sur le « réductionnisme » de la doctrine, John Stuart Mill extrait en outre la matière pour une seconde objection plus précise : Bentham et James Mill ne parviennent pas à construire une éthique scientifique rigoureusement incontestable, car cette conception unidimensionnelle des plaisirs, jointe à leur souci primordial de pragmatisme, les amène à privilégier, dans la nature humaine, le

<sup>18</sup> Halévy (1995), vol. 1 p. 22. À cette objection de fond, Bentham et ses adeptes répondaient que c'est précisément le rôle du législateur, inspiré par l'idéal utilitariste, que de concevoir un système de règles morales et juridiques qui assure artificiellement l'identification des intérêts individuels à l'intérêt général (cf. *infra* II-B-1).

<sup>19</sup> MILL (1968), p. 50-51. Le test de la qualité entre deux plaisirs repose alors sur la préférence bien arrêtée accordée à l'un d'entre eux par tous ceux, ou presque tous ceux, qui ont l'expérience de l'un et de l'autre, préférence accordée sans qu'ils y soient poussés par un sentiment d'obligation morale. La problématique de la qualité des plaisirs, si elle est clairement exposée par J.S. Mill, ne semble pas correctement résolue par le test de qualité qu'il suggère. Viner, (1949) p. 216, propose une interprétation conforme à l'analyse marginaliste de l'utilité : les désirs étant égaux à la marge, une « dose » supplémentaire de jeu à la poussette est *a priori* équivalente à une « dose » supplémentaire de poésie. En fait ces plaisirs restent qualitativement différents si les individus expérimentés ne choisissent jamais une première « dose » de jeu à la poussette tant qu'ils ne sont pas saturés de poésie.

penchant prédominant si ce n'est exclusif, c'est-à-dire à réduire la recherche individuelle des plaisirs à la simple protection des intérêts personnels.

En laissant peu de place pour des sentiments humains habituellement considérés comme plus nobles (générosité, bienveillance, « amour » de la justice...), Bentham cherche à isoler à travers l'égoïsme « le sentiment qui [paraît] le mieux se prêter à la mesure. Or le sentiment de sympathie paraît, moins que tout autre, satisfaire à cette condition : comment dire, sans absurdité, que le sentiment de sympathie varie, selon une loi quelconque, en raison du nombre de ses objets ? Au contraire, les sentiments égoïstes admettent, mieux que tous les autres, un équivalent objectif. La crainte d'une douleur se laisse, avec quelque précision, évaluer et comparer à d'autres craintes, lorsque les douleurs considérées sont des douleurs proprement égoïstes, lorsqu'il s'agit, par exemple, de la crainte d'une amende déterminée... Une espérance se laisse, avec quelque précision, évaluer et comparer à d'autres espérances, lorsque les plaisirs espérés sont d'ordre égoïste, lorsqu'il s'agit, par exemple, de recevoir un nombre déterminé de pièces de monnaie égales entre elles : c'est pourquoi l'économie politique, la « dogmatique de l'égoïsme », constitue peut-être la plus fameuse des applications du principe de l'utilité »<sup>20</sup>.

Au total, l'influence prédominante attribuée à l'égoïsme dans les actions humaines, le nivellement de la qualité des plaisirs et la « guillotine » de Hume constituent un ensemble cohérent d'arguments de poids qui relativisent la portée éthique de l'utilitarisme classique fondé sur l'acte. C'est, nous semble-t-il, l'intérêt primordial des travaux de Harsanyi d'avoir réuni, face à ces critiques de grande ampleur, les éléments d'une reformulation de la doctrine utilitariste.

### *L'originalité de l'utilitarisme de risque*

Dans cette perspective, Harsanyi<sup>21</sup> introduit deux amendements ; le premier, relativement mineur, résulte du transfert de la pensée utilitariste de son contexte initial, centré sur une psychologie hédoniste, vers le cadre de la théorie moderne de l'utilité en situation de risque, fondée sur les postulats de Von Neumann et Morgenstern<sup>22</sup>.

Bien que cette relecture soit faite en termes formalisés, l'essentiel, malgré les apparences, est préservé : l'individu reste le meilleur juge – autonome et rationnel – de ce qui est bien pour lui ; au lieu de l'expri-

<sup>20</sup> Halévy (1995), vol. 1 p. 24-25.

<sup>21</sup> Harsanyi (1953), (1955), (1977).

<sup>22</sup> Von Neumann et Morgenstern (1947).

mer en termes hédonistes de plaisirs et de peines, il émet des préférences cardinales dont il cherche à maximiser l'utilité et dont les sources se trouvent diversifiées. En particulier, « il est au moins plausible que dans beaucoup de circonstances nous sommes plus intéressés par la réalisation de certaines situations objectives que par les sentiments subjectifs de plaisirs et de peines que nous pouvons éprouver »<sup>23</sup>.

Ceci étant admis, l'essentiel reste à faire, c'est-à-dire à conférer aux préférences individuelles un contenu éthique plus solide, permettant de surmonter les obstacles énoncés par Hume et par John Stuart Mill.

L'originalité du deuxième amendement de Harsanyi est alors de concevoir *a priori* l'existence de deux ensembles de préférences pour chaque individu et d'une procédure déontologique permettant de les séparer.

La première catégorie est constituée par les préférences « personnelles » de l'agent, celles qui guident son comportement quotidien et sont exprimées par la fonction d'utilité ; sans être totalement égoïstes, elles assignent aux intérêts propres de l'individu, à ceux de sa famille ou de ses relations proches, une pondération plus élevée qu'elles n'en attribuent aux intérêts de tiers complètement inconnus.

Les préférences « morales », quant à elles, peuvent ou non imprégner le comportement quotidien de l'agent mais, en tout cas, elles guident sa pensée dans les situations éventuellement très rares, où il est contraint à une attitude impersonnelle et impartiale : la pondération attribuée à chacun – y compris la sienne propre – est identique. Le point névralgique de la construction éthique de Harsanyi est alors la définition d'une situation de risque, où l'individu, tout en étant lui-même, est contraint de penser à autrui exactement dans les mêmes termes.

Autrement dit, la moralité du comportement de chacun ne dépendra pas d'un quelconque sursaut intérieur, d'un effort sur soi-même mais bien d'une pression extérieure. Cette intervention extérieure, Harsanyi la trouve dans une situation de risque, où l'individu est contraint à la moralité par l'ignorance de sa propre situation ou, du moins, agit comme s'il ignorait les caractères de celle-ci. L'impartialité dont il est obligé de faire preuve l'amène ainsi à attribuer la même probabilité à la réalisation de chaque situation possible, afin de n'en privilégier aucune.

Nous qualifierons d'utilitarisme de risque l'ensemble de la réflexion issue de la définition d'une situation d'ignorance avec équiprobabilité. Cet édifice néo-utilitariste, s'il prend en compte les objections majeures de Hume et de John Stuart Mill, permet en outre de retrouver la norme

---

<sup>23</sup> Harsanyi (1977), p. 54.

sociale du plus grand bonheur du plus grand nombre à partir d'une conception procédurale de la justice, laquelle peut être comparée, à bien des égards, avec la conception alternative du « voile d'ignorance » proposée par Rawls <sup>24</sup>.

Fondée sur l'acte ou fondée sur le risque, les deux formes d'utilitarisme qui viennent d'être évoquées ont pour point commun d'analyser la nature éthique de chaque comportement particulier. Tel n'est pas le cas du dernier démembrement de la réflexion utilitariste, centré d'abord sur la genèse et la cohérence des règles morales.

### I. – 3. L'utilitarisme de règle

L'utilitarisme de règle, comme la version précédente fondée sur le risque, a pour origine le constat d'imperfections essentielles de la doctrine classique initiale. Celles-ci résident dans le manque d'épaisseur temporelle de l'utilitarisme d'acte, dont la morale est rarement située en perspective.

#### *L'atemporalité de l'utilitarisme classique*

La doctrine classique, tout en appliquant un seul et même principe – la maximisation de l'utilité individuelle –, aboutit à une morale de l'instant, c'est-à-dire à des prescriptions tout à fait contingentes, variables dans le temps et selon les individus concernés.

Dans certains cas, la morale utilitariste classique peut même heurter de front les principes les plus élémentaires de la morale intuitive. À la suite notamment de Kelman <sup>25</sup>, on parvient ainsi à construire des situations hypothétiques dans lesquelles la somme des plaisirs et des peines aboutit à la moralité du mensonge, à la condamnation d'un innocent ou au non-respect de la parole donnée.

Dans ce dernier cas, par exemple, le contexte est le suivant : deux amis très liés partent ensemble pour une expédition polaire. L'un d'eux tombe épuisé dans la neige et meurt rapidement, sans qu'aucun secours ne soit possible. Au moment de mourir, il ne demande qu'une chose à son ami : lui promettre que, dans dix ans jour pour jour, il reviendra au même endroit et allumera une bougie à sa mémoire. Sur l'instant, l'ami promet solennellement mais, dix ans plus tard, se pose pour lui la question de tenir réellement sa promesse, car il lui serait pénible de refaire le long voyage. En outre, puisqu'il n'a parlé à personne de cette promesse, son renoncement n'ébranlerait pas la confiance générale dans le

<sup>24</sup> Sur ce point, cf. Gamel (1991) p. 70 à 82.

<sup>25</sup> Kelman (1981).

respect de la parole donnée. Enfin, l'individu juge les circonstances suffisamment particulières pour penser que son renoncement ne l'incitera pas à ne pas tenir ultérieurement d'autres promesses. Au total, les inconvénients liés au respect de son engagement l'emportant sur ses avantages, il serait moralement inconvenant d'aller sur la banquise allumer la chandelle.

Si particulier qu'il soit, cet exemple a notamment le mérite de souligner en creux le rôle décisif d'une information complète sur toutes les conséquences d'un acte individuel. Dans le cas d'espèce, cet impératif est peu exigeant, car les incidences ultérieures de l'attitude de l'individu sur le comportement d'autrui, d'une part, et sur le sien propre, d'autre part, sont supposées nulles. C'est pourquoi le bilan des plaisirs et des peines est relativement aisé à faire : il suffit de prendre en compte les avantages et les inconvénients immédiats pour le seul individu concerné par le respect de la promesse, ce qui explique d'ailleurs la prescription morale tout à fait insolite à laquelle on parvient.

Par contre si, comme c'est le cas le plus souvent, il convient de tenir compte non seulement des incidences immédiates aisément perceptibles mais aussi de toutes les répercussions ultérieures, sur soi-même et aussi sur autrui, d'un acte individuel isolé, le vertige peut s'emparer du moraliste devant l'ampleur de la collecte d'informations à laquelle il doit s'astreindre, avant de pouvoir statuer.

L'ampleur démesurée d'une telle tâche aboutissant à une nouvelle impasse de la doctrine classique, un certain nombre d'utilitaristes ont progressivement pris conscience de la nécessité de modifier sensiblement l'axe de leur réflexion, en conservant certes le principe de l'utilité, mais en changeant l'objet sur lequel l'appliquer.

#### *La révision doctrinale*

Les prémices de cette révision doctrinale sont controversées ; certains les trouvent dans la pensée de John Stuart Mill, voire même dans l'œuvre de Hume<sup>26</sup>. Cependant un consensus semble se dégager pour considérer que Harrod<sup>27</sup> fut le premier à énoncer clairement les avantages d'une nouvelle version, générique, de l'utilitarisme par rapport à la vision particulariste, traditionnelle, depuis Bentham ; mais le clivage définitif « règle/acte » ne paraît avoir été introduit qu'en 1959<sup>28</sup>.

Dans cette nouvelle version, le bilan des plaisirs et des peines n'est plus appliqué directement à chaque acte individuel, dont il est impos-

<sup>26</sup> Controverse évoquée notamment par Langlois (1982).

<sup>27</sup> Harrod (1936).

<sup>28</sup> Brandt (1959).

sible d'entrevoir toutes les incidences, mais d'abord à la règle générale de conduite qui le déterminera. Désormais toute action humaine tire sa valeur éthique, indirectement, de sa conformité à la règle morale idoine, c'est-à-dire à celle qui maximiserait dans le temps l'utilité collective, dans la mesure où elle serait suivie par tous dans les situations aux caractères génériques comparables. En outre, cette prescription morale spécifique doit s'insérer correctement dans l'ensemble du « code » moral en vigueur.

Ce faisant, on conçoit que l'utilitarisme de règle pose d'abord une question institutionnelle : celle de la sélection et de la genèse d'un système de règles cohérentes entre elles, dont l'effet conjugué doit permettre de garantir dans le temps le plus grand bonheur du plus grand nombre, c'est-à-dire de retrouver la norme de justice sociale de la doctrine classique.

En d'autres termes, comme la version de la théorie centrée sur le risque, l'utilitarisme de règle adopte une démarche déontologique : il tente d'abord d'en définir la norme de justice et d'en déduire ensuite le bien, par simple conformité du comportement individuel. Toutefois, dans la mesure où la moralité des actions humaines réside ici dans le respect de règles préalablement définies, cette version de l'utilitarisme est plutôt à comparer avec la problématique de l'émergence des « règles de juste conduite », propre à la pensée de Hayek <sup>29</sup>.

Avec cette dernière remarque s'achève la présentation du « premier cercle » de notre définition de l'utilitarisme, qui avait surtout pour buts de souligner la continuité de cette doctrine au-delà de ses reformulations successives et, incidemment, de révéler la nature du principal fondement philosophique de l'analyse économique.

La continuité de la pensée utilitariste réside dans la référence constante au principe de l'utilité, c'est-à-dire à la conviction que, au moins au premier abord, chaque individu est le mieux à même de définir ce qui est bien pour lui, que ce soit en cherchant le plaisir et en évitant la peine, que ce soit, plus généralement, en formulant des choix ou que ce soit, encore, en adhérant à des règles morales dont la qualité essentielle est de garantir dans le temps la validité de ce principe. Appliqué à l'acte, au risque ou à la règle, un tel postulat aboutit systématiquement à privilégier le comportement dont les effets en termes d'utilité immédiate, d'utilité espérée ou d'utilité intertemporelle, seront maximaux.

On conçoit alors en quoi la réflexion utilitariste a pu féconder et habiter toujours la pensée économique et on mesure mieux la dette de la seconde à l'égard de la première :

<sup>29</sup> Sur ce point, cf. Gamel (1992) p. 87 à 89.

– En tant que loi psychologique descriptive du comportement humain, le principe de l'utilité constitue le fondement sur lequel repose toute l'analyse positive des choix du consommateur et des autres agents économiques.

– En tant que prescription normative imposée à l'individu, le principe de l'utilité constitue la référence commune des économistes du bien-être, soucieux des aspects éthiques de l'allocation des ressources rares.

En ce sens très large, c'est-à-dire en définissant la doctrine par rapport à ce principe, l'analyse économique, que ce soit dans ses applications de politique économique ou dans ses études théoriques de normes, a pour référence l'utilitarisme, par rapport auquel peuvent ainsi se rapprocher Walras et John Stuart Mill, Pareto et Edgeworth, Pigou et Hicks ainsi que, de nos jours, Arrow, Harsanyi, Hayek ou Rawls<sup>30</sup>.

Cependant ce rapprochement ne peut être que momentané et superficiel, car l'ambition éthique de l'utilitarisme a toujours dépassé, depuis ses origines, ce dénominateur commun de la morale personnelle pour parvenir à la définition d'une norme sociale de justice. Ce sont dans les modalités de passage de l'une à l'autre qu'apparaît un clivage irrémédiable, donnant naissance à une seconde définition beaucoup plus restreinte de l'utilitarisme : si tous les économistes du bien-être acceptent la morale personnelle issue de l'utilitarisme d'acte, en revanche beaucoup en refuseront la norme sociale de justice ou, plus précisément, l'arithmétique interindividuelle des plaisirs nécessaire, dans la doctrine classique, à sa définition.

## II. – L'ARITHMÉTIQUE INTERINDIVIDUELLE ET SES SUBSTITUTS DE SECOND RANG

Pour progresser vers le cœur de la doctrine utilitariste, nous pouvons ignorer, en toute rigueur chronologique, les refontes contemporaines en termes de risque et de règle ; l'utilitarisme, réduit à sa version originelle fondée sur l'acte, est donc une théorie où le passage du bien individuel à la norme de justice se fait par « simple » maximisation du bien.

Dans cette étape caractéristique d'une réflexion téléologique, la doctrine utilitariste se heurte à deux obstacles successifs :

– La maximisation du bien suppose d'abord que chacun puisse estimer sur une échelle propre la quantité de bien liée à chaque état social.

<sup>30</sup> Rawls lui-même, reconnaît l'existence de ce « consensus » minimal : « je suppose avec l'utilitarisme que le bien est défini comme la satisfaction du désir rationnel » 1987 p. 56.

En l'occurrence, il s'agit donc de définir pour chacun un indice cardinal susceptible de quantifier les utilités individuelles (A).

– La résolution nécessaire de cette première difficulté n'est cependant pas suffisante ; pour procéder à l'agrégation des niveaux de satisfaction individuelle, il faut encore pouvoir les comparer. La nécessité d'un « arbitre impartial et bien informé », capable d'effectuer des comparaisons interpersonnelles d'utilité, se fait alors sentir (B).

## II. – 1. La nécessité d'un indice cardinal d'utilité

L'application du principe de l'utilité au niveau individuel a permis de définir la morale personnelle de l'utilité : est juste pour un être capable de sensations le comportement qui maximise l'excès d'une somme de plaisirs sur une somme de peines. Par simple répétition du principe au niveau collectif, il est possible de définir le bonheur du tout social qu'est la société et d'en déduire l'organisation juste : « chacun comptant pour un et personne pour plus d'un », le bonheur sera la somme des plaisirs et des peines ressentis par l'ensemble des individus. Sera donc socialement juste l'organisation de la société qui maximisera cette somme, c'est-à-dire celle qui assure « le plus grand bonheur du plus grand nombre ».

Ce passage de l'individuel au collectif caractérise l'utilitarisme qui va ainsi bien au-delà de la morale hédoniste initiale.

### *Au-delà de l'hédonisme*

Deux remarques préliminaires sont ici nécessaires pour mettre en valeur la portée spécifique de l'utilitarisme :

– Comme l'observation en a déjà été faite, il est maintenant devenu habituel en théorie économique de définir l'utilité comme une mesure du bien-être individuel et non plus nécessairement comme une simple mesure du plaisir au sens de Bentham. L'approfondissement de cette notion de bien-être et notamment l'analyse de ses sources – purement individuelles ou interpersonnelles – interviendront ultérieurement, en particulier avec la distinction fameuse de Pareto entre utilité et ophélimité. Au stade actuel de notre réflexion, l'expression utilitarisme est donc employée dans un sens large, où le bien-être est identifié à l'utilité définie comme le sentiment psychologique de satisfaction personnelle, quelles qu'en soient les origines ;

– Même ainsi modifiée, la définition de l'utilité et son interprétation au niveau collectif suggèrent une seconde observation ; ainsi que le remarque Halévy, « pour que l'idée d'un bonheur maximum ait un sens,



il faut qu'un calcul des plaisirs et des peines soit possible : c'est ce que nous appellerons le postulat rationaliste de la doctrine utilitariste. Postulat indépendant du principe de l'utilité, et nécessaire seulement si l'on veut qu'une science sociale, fondée sur ce principe, soit possible. Car les hédonistes anciens, selon qui le plaisir est la fin naturelle des actions humaines, ne se préoccupaient nullement de savoir, après cela, si une comparaison quantitative des plaisirs et des peines était possible »<sup>31</sup>.

Dans ce passage, Halévy souligne la principale différence entre l'hédonisme antique et l'utilitarisme du XIX<sup>e</sup> siècle : par rapport aux hédonistes, Bentham et ses disciples manifestent l'ambition supplémentaire de dégager un classement des états sociaux selon le critère de bonheur collectif issu de l'agrégation des utilités individuelles. Autrement dit, la morale personnelle, issue de l'hédonisme, n'a pas de prolongement au plan collectif pour définir une quelconque norme de justice sociale.

C'est précisément dans l'élaboration d'un tel prolongement que vont s'amonceler des difficultés que Bentham n'a d'ailleurs pas vraiment résolues, même s'il les a, semble-t-il, entrevues<sup>32</sup>.

L'analyse des conditions nécessaires à la définition d'un optimum social selon la norme utilitariste révèle d'abord que chaque individu doit être capable de classer les états sociaux selon ses préférences personnelles. Un classement ordinal pourrait *a priori* être envisagé. En ce cas, chacun est capable d'attribuer à chaque état social un nombre d'autant plus élevé que l'utilité qu'il associe à cet état est forte. Toutefois, chaque nombre ainsi défini traduit non pas le niveau exact d'utilité procuré par un état social donné, mais son rang dans le classement d'ensemble de tous les états sociaux, tel que l'estime l'individu.

Même si une telle hiérarchie était établie par chaque sociétaire, elle ne suffirait pas pour autant à rendre opératoire l'application du principe de l'utilité au plan collectif, car celui-ci suppose l'agrégation des utilités individuelles ; en d'autres termes, l'indice numérique traduisant l'utilité collective de chaque état social sera construit à partir de la som-

<sup>31</sup> Halévy (1900), p. 312. Sur ce point, cf. également Halévy (1995) vol. 3 p. 222 et 226.

<sup>32</sup> Selon Halévy, (1900) p. 314, « dans un fragment manuscrit qui n'a passé dans aucune de ces œuvres imprimées », Bentham écrit : « Il est inutile de parler d'additionner des quantités qui, après l'addition, resteront distinctes comme elles étaient avant. Le bonheur d'un homme ne sera jamais le bonheur d'un autre homme ; le gain d'un homme n'est pas le gain d'un autre. Vous pourriez aussi bien prétendre ajouter vingt pommes à vingt poires, et obtenir ainsi quarante unités de même nature ; vous auriez toujours, en réalité, vingt objets de l'une et de l'autre espèces. Cette additivité du bonheur des sujets différents, quoiqu'à le considérer rigoureusement, il puisse paraître fictif, est un postulat faute duquel tout raisonnement politique est rendu impossible ; il n'est pas d'ailleurs plus fictif que celui de l'égalité du probable et du réel sur lequel est établie toute la branche des mathématiques qu'on appelle la théorie des probabilités ». Sur ce point, cf. également Halévy (1995) vol. 3 p. 227.

mation des indices individuels attribués par chacun aux états sociaux en cause. Pour que cette sommation ait un sens, il faut donc que les indices individuels d'utilité traduisent non le rang de classement mais le niveau absolu d'utilité attribué par chacun à chaque état social. Ainsi la construction d'un indice cardinal d'utilité propre à chacun est-elle la première exigence imposée par la norme de justice sociale, exigence dont la difficulté va susciter diverses variantes du cardinalisme.

#### *Les variantes du cardinalisme*

La mesure de l'utilité individuelle soulève *a priori* des problèmes analogues à ceux rencontrés, par exemple, dans la mesure physique de la température : en particulier, il convient de choisir une origine et une unité étalon pour l'échelle d'utilité de l'individu. Comme ce double choix, affaire de conventions, est susceptible à tout moment de modifications, on admet généralement que l'indice cardinal d'utilité individuelle ne peut être connu, au mieux, qu'à une transformation affine positive près<sup>33</sup>.

A l'origine de la réflexion utilitariste, Bentham n'a pas consacré de longs développements à préciser ainsi les modalités de la cardinalisation de l'utilité. Ainsi que l'écrit par exemple Arrow, « Bentham considérait comme établi que l'utilité était une grandeur mesurable. Il définissait de différentes manières les facteurs qui déterminent l'utilité, comme la proximité dans le temps et la certitude, mais nulle part on ne trouve un procédé clairement défini de mesure de l'utilité, tel que l'exigera ultérieurement la philosophie scientifique moderne »<sup>34</sup>.

Pour Bentham en effet, l'intensité d'un plaisir n'était qu'une dimension parmi d'autres : la durée, la certitude ou l'incertitude, la proximité ou l'éloignement dans le temps. « L'arithmétique morale » consiste alors à définir les règles de calcul spécifiques à chaque dimension ; en ce qui concerne l'intensité, l'idée maîtresse réside dans l'observation que celle-ci a un minimum : le plus faible degré de plaisir qui se laisse distinguer d'un état d'insensibilité. A partir de ce minimum pris comme unité et de cet état insensible pris comme origine, l'intensité du plaisir est une grandeur susceptible de croître sans limite.

<sup>33</sup> Il existe pour la température au moins deux échelles de mesure, en degrés Celsius et en degrés Fahrenheit et l'on passe de l'une à l'autre par la transformation affine :  $^{\circ}F = 32 + 1,8 \text{ } ^{\circ}C$  ; de même, pour un individu quelconque, deux indices cardinaux d'utilité  $V$  et  $U$  relatifs à un état social  $j$  se déduisent l'un de l'autre par la relation :  $V(j) = aU(j) + b$  avec :  $a, b$  constantes et  $a > 0$ . Un coefficient  $b$  non nul implique un déplacement du « degré 0 » de l'échelle d'utilité. Un coefficient  $a$ , différent de 1, traduit une variation de son unité étalon. Le signe positif de  $a$  évite la transformation de la mesure positive dans une échelle (utilité) en une mesure négative dans l'autre (désutilité).

<sup>34</sup> Arrow, (1984) p. 118.

Faute de précisions supplémentaires, notamment sur les modalités d'application de cette arithmétique, les recherches ultérieures dans ce domaine adoptèrent des démarches très différentes que l'on peut néanmoins regrouper selon deux axes principaux.

Le premier axe de recherches repose sur l'hypothèse plus réaliste que chaque personne ne peut pas réellement comparer avec précision un nombre infini de niveaux d'utilité eu égard à une *capacité limitée de discrimination*.

La différence entre deux niveaux d'utilité consécutifs demeure la variation d'utilité minimale que l'individu peut discerner ; celui-ci est donc indifférent entre les choix situés sur un même niveau d'utilité et on peut donc obtenir une mesure cardinale de la différence d'utilité entre deux états sociaux en faisant le compte des niveaux de discrimination qui les séparent. L'échelle cardinale ainsi obtenue est naturellement unique à une transformation affine près, compte tenu du choix, qui reste arbitraire, d'une unité et d'une origine.

Conceptuellement proche de la pensée de Bentham tout en s'en démarquant par une démarche plus pragmatique, ce premier axe de recherches semble avoir ses prémices dans les travaux de Borda<sup>35</sup> concernant les systèmes électoraux et a bénéficié plus près de nous d'un regain d'intérêt, avec notamment les travaux d'Edgeworth, Armstrong, Goodman et Markovitz ou Rothenberg<sup>36</sup>. Ces travaux n'ont pas pour préoccupation commune de renouveler les fondements de l'approche mais d'en extraire les applications, par exemple en termes de fiscalité optimale. Cependant, ce type de recherches reste obéré par les difficultés pratiques rencontrées dans le repérage des différents niveaux de discrimination ; par ailleurs, lorsqu'elle est appliquée à l'économie du bien-être, la méthode de cardinalisation utilisée soulève une difficulté supplémentaire : il est en effet impossible d'observer l'estimation par l'individu de l'utilité d'états sociaux alternatifs qui restent totalement virtuels et ne peuvent faire l'objet que d'une expérimentation par la pensée.

Le second axe de recherches, plus sophistiqué mais aux origines au moins aussi anciennes, repose sur l'introduction d'hypothèses additionnelles sur la nature des choix qu'effectue l'individu. Le principe général réside dans l'affirmation que les différents biens ou, plus largement, les différentes décisions peuvent être répartis en plusieurs catégories étanches, de telle sorte que l'utilité de chaque bien d'une classe donnée est indépendante des quantités de biens disponibles dans les autres classes.

<sup>35</sup> Borda, « *Mémoire sur les élections au scrutin* » (1781).

<sup>36</sup> Edgeworth (1881), Armstrong (1951), Goodman et Markovitz (1952) et Rothenberg (1961).

De cette manière, il est possible d'attribuer un niveau cardinal d'utilité à des ensembles de biens (à l'intérieur de chaque classe) et de définir l'utilité du panier entier comme la somme des niveaux d'utilité de chaque classe. La séparabilité des biens en ensembles indépendants permet d'introduire la propriété d'additivité de leur utilité et d'estimer simultanément le niveau de satisfaction de chaque sous-ensemble et de l'ensemble total ; d'où l'appellation de ce procédé connu généralement sous le nom de *mesure conjointe de l'utilité*.

La « troïka » de l'école marginaliste, Jevons, Menger, Walras, souvent reconnus comme les véritables pères de l'usage en théorie économique de la psychologie utilitariste<sup>37</sup>, proposait déjà une version rudimentaire de fonctions d'utilité additives, dont les propriétés formelles ont été depuis l'objet de nombreux approfondissements. Certes l'absence de toute complémentarité entre les différents arguments de la fonction d'utilité est souvent considérée comme une hypothèse non justifiable par rapport à la réalité, mais cette critique perd une partie de son acuité lorsqu'elle vise une méthode particulière qui, à l'intérieur de la problématique de la mesure conjointe, a pris une place essentielle. La méthode est particulière dans la mesure où elle s'attache à la mesure de l'utilité en situation de risque ; la place est essentielle car les prolongements en furent multiples, tant dans la théorie de la décision qu'en économie du bien-être.

La première tentative de définition d'une fonction d'utilité à partir des réactions individuelles face à des situations seulement connues en probabilité remonte au mémoire de Bernouilli sur le célèbre Paradoxe de Saint-Pétersbourg. Comme on le sait, une façon de résoudre ce paradoxe consiste à reconnaître que le joueur fonde son attitude non pas sur l'espérance mathématique des gains mais, selon l'expression même de Bernouilli, sur « l'espérance morale » de gagner ; en ce cas les probabilités ne pondèrent plus directement les niveaux de revenu mais les degrés d'utilité attachés à chaque niveau de revenu. Toutefois, c'est évidemment à Ramsey<sup>38</sup> et surtout à Von Neumann et Morgenstern<sup>39</sup> que l'on doit d'avoir précisé au XX<sup>e</sup> siècle les postulats suffisants à la mesure cardinale de l'utilité en situation aléatoire et d'avoir généralisé les conclusions de Bernouilli concernant l'aversion pour le risque. Dans ce cadre, le principe de la mesure conjointe, précédemment évoqué, suppose l'indépendance stochastique de la mesure de l'utilité par rapport aux différentes « loteries canoniques », auxquelles l'individu a pu successivement participer et dont sa mémoire a gardé le souvenir.

<sup>37</sup> Cf. notamment Arrow (1984) p. 118.

<sup>38</sup> Ramsey (1931).

<sup>39</sup> Von Neumann et Morgenstern (1947).

Appliquée au problème du choix social, cette hypothèse d'indépendance<sup>40</sup> s'est heurtée à une objection importante<sup>41</sup> : l'individu doit juger équivalent d'avoir maintenant dans la vie la même chance qu'un autre d'atteindre un certain statut dans la société et d'avoir eu, par hypothèse, à l'origine la même chance d'être la personne vouée avec certitude à cette destinée. Par exemple, « dans les canots de sauvetage où la règle est « les femmes d'abord », est-ce une consolation pour un homme de savoir qu'il avait la même chance d'être une femme que n'importe quelle femme, avant l'affectation de chacun à sa place dans la société ? Que dire aussi des Noirs obligés de s'asseoir à l'arrière des autocars ? »<sup>42</sup>. Plus récemment est apparue une controverse plus générale sur le sens à donner à l'agrégation de fonctions d'utilité individuelle issues des postulats de Von Neumann et Morgenstern : s'agit-il, comme le soutient Harsanyi, d'une axiomatisation de la norme utilitariste fondée sur la théorie moderne de l'utilité cardinale ou, comme l'affirme Sen, d'une simple « représentation » des préférences sociales, qui restent ordinales, à partir de loteries respectant les postulats neumaniens ?<sup>43</sup>.

Au total, que l'individu parvienne à mesurer cardinalement l'utilité selon cette méthode ou selon l'hypothèse alternative de discrimination limitée, le problème posé par la norme utilitariste de justice n'est qu'en partie résolu. Pour choisir l'état social qui assure le plus grand bonheur du plus grand nombre, il faut encore harmoniser ces échelles individuelles d'utilité, ce qui nécessite peu ou prou l'intervention d'un « arbitre impartial et bien informé ».

## II. – 2. La nécessité d'un « arbitre impartial et bien informé »

L'agrégation des utilités individuelles évaluées cardinalement par chaque membre de la société peut être appréhendée de deux points de vue complémentaires : il s'agit d'abord de définir sur le plan logique les éléments d'information nécessaires, puis d'évaluer la signification éthique de cette collecte d'informations.

En ce qui concerne le premier point, il faut noter que l'application de la norme utilitariste du plus grand bonheur du plus grand nombre n'implique pas une collecte d'informations exhaustive. Le contexte idéal serait celui dans lequel l'utilité admettrait une échelle naturelle de mesure qui s'imposerait d'évidence à tous, chacun évaluant ses satisfactions

<sup>40</sup> Appelée aussi « principe de la chose sûre ».

<sup>41</sup> Formulée initialement par Diamond (1967), cette objection a été par la suite amplement commentée, cf. notamment Sen (1970) p. 143-145.

<sup>42</sup> Phelps, (1973) p. 24.

<sup>43</sup> Cf. respectivement Harsanyi (1955) et Sen (1986) ; pour un exposé de cette controverse, voir Weymark (1991) et Roemer (1996) p. 138-147.

dans une même unité naturelle et à partir d'une même origine. Une telle situation simplifierait certes l'arithmétique utilitariste mais n'est pas nécessaire : on peut admettre que chacun ait sa propre échelle de mesure (unité étalon et origine propres).

Pour procéder à l'agrégation des utilités, il conviendrait alors, au préalable, de convertir dans une même échelle de référence, dont le choix reste arbitraire, l'évaluation cardinale de ces utilités. En d'autres termes, cette conversion suppose connue la clé de passage de chaque échelle individuelle à l'échelle de référence, exactement de la même manière que l'on peut comparer la température en degrés Celsius avec celle mesurée en degrés Fahrenheit, à condition de connaître l'équation de conversion de l'une à l'autre <sup>44</sup>.

En fait, la recherche de l'optimum social utilitariste est encore moins exigeante : pour déterminer cet état optimal, il suffit en effet de connaître le solde net des gains et des pertes d'utilité qu'éprouveraient les individus dans la transformation sociale de la société d'un état  $x$  en un état  $y$  ; de proche en proche, la multiplication de telles comparaisons binaires permet de mettre à jour l'optimum recherché ; or, avec cette méthode, il convient seulement de comparer, pour pouvoir les additionner, non pas les niveaux d'utilité mais les variations positives de ceux qui profitent de la transformation et les variations négatives de ceux qui la subissent. En d'autres termes, la norme de justice utilitariste, dont l'objectif est uniquement de hiérarchiser les états sociaux, n'exige que la comparabilité des unités étalon des différentes échelles d'utilité, quelle qu'en soit par ailleurs l'origine.

Cet élément d'information minimal, qu'il convient d'obtenir de chaque membre de la société, constitue néanmoins pour la doctrine utilitariste un redoutable obstacle qui ne pourra être surmonté, de manière plus ou moins élégante, que par l'intervention directe d'un *deus ex machina* extérieur à l'analyse ; pour que la réflexion progresse, il faut en effet admettre qu'il existe quelque part dans la société, « en un lieu défini mais peut-être inaccessible à la raison et à la sagesse limitées des hommes » <sup>45</sup>, un médiateur capable de collecter cette information sans la déformer. Il s'agit donc d'admettre la fiction d'un « arbitre impartial et bien informé » capable, par une sorte de don d'ubiquité, de s'identifier en même temps à chacun des membres qui composent la société.

Cette notion fort ancienne de spectateur impartial et rationnel semble remonter à Hume et à Smith <sup>46</sup> ; pour eux la notion de justice, appliquée

<sup>44</sup> En l'occurrence :  $^{\circ}F = 32 + 1,8 \text{ }^{\circ}C$ .

<sup>45</sup> Dupuy (1992), p. 110.

<sup>46</sup> Sur ce point, cf. notamment Rawls (1987) p. 215 à 217 ; Hume et Smith mettent surtout en avant le caractère « bien intentionné » et non « bien informé » de cet arbitre impartial.

par exemple à une institution ou à un système social, s'identifie à l'approbation de cette institution par un observateur rationnel et impartial, dont l'avantage majeur serait de posséder la connaissance parfaite des éléments nécessaires à l'application de la norme de justice. Ce faisant, il s'agit de transposer au niveau sociétal la rationalité des choix définie au plan individuel ; tout se passe comme si cet observateur macroscopique « avait le cœur si gros qu'il se réjouissait de tout le bonheur du monde et s'attristait de son malheur ; son impartialité postulée lui faisant accorder le même poids à chacun – Bentham disait : « chacun compte pour un et personne pour plus de un » -, cet être hypothétique souffre et jouit de la même manière que la société considérée comme un tout »<sup>47</sup>.

Si les pionniers de la doctrine utilitariste admettaient sans complexe l'intervention de cette fiction anthropomorphe et la justifiaient, nombre d'auteurs plus récents tenteront d'échapper à son intervention au point de résoudre le problème de la collecte d'informations, en le supprimant, par une hypothèse encore plus forte.

#### *L'argumentation des pionniers*

Pour Bentham et ses disciples, la fiction d'un arbitre impartial et bien informé non seulement est nécessaire mais doit en outre devenir réalité à travers l'intervention du « législateur », dans la mesure où, comme on l'a déjà souligné<sup>48</sup>, la compatibilité de la morale personnelle de l'utilité avec la norme du bien commun n'est pas systématiquement établie.

La fusion naturelle des intérêts à l'intérieur du bien commun est en effet peu probable, dès lors que ne prédomine pas en chacun un sentiment de sympathie qui l'intéresserait au bonheur d'autrui. Si cette éventualité est écartée, le principe de l'utilité peut encore être interprété de deux manières distinctes<sup>49</sup>.

On peut d'abord raisonner de la manière suivante : puisqu'il est reconnu que les mobiles égoïstes prédominent dans la nature humaine et que, d'ailleurs, l'espèce humaine vit, subsiste et se développe, il faut admettre que les égoïsmes s'harmonisent d'eux-mêmes et produisent mécaniquement le bien de l'espèce.

C'est sur ce principe de « *l'identité naturelle des intérêts* » que repose depuis Smith l'économie politique : par le mécanisme de l'échange

<sup>47</sup> Dupuy (1992), p. 114-115.

<sup>48</sup> Cf. I-B-1.

<sup>49</sup> Nous reprenons ici l'analyse ternaire de Halévy, (1995) vol. 1 p. 22 à 27 notamment, qui distingue d'une part le principe de la « fusion des intérêts » et, d'autre part, les principes de « l'identité naturelle » et de « l'identification artificielle des intérêts ».

et de la division du travail, les individus sans le vouloir, sans même le savoir, en poursuivant chacun leur intérêt propre, travaillent à réaliser d'une manière immédiate l'intérêt général. *A priori* Bentham semble retenir l'essentiel des conclusions de Smith sur le « laisser-faire » qui doit caractériser les activités économiques. Son point de vue sera cependant beaucoup plus nuancé et empirique : l'intervention du gouvernement en matière économique est le plus souvent source de désutilité car, dans ce domaine, les individus connaissent beaucoup mieux que quiconque leur propre intérêt et agissent donc avec plus d'ardeur.

Cependant, ceci n'est qu'une règle et non pas un principe : en tant que telle, elle peut souffrir de nombreuses exceptions, dès lors que le plus grand bonheur du plus grand nombre l'exige et Bentham ne limite pas ces exceptions à des cas fortuits ou à des situations transitoires ; bien au contraire, il développe une analyse élaborée des circonstances dans lesquelles le gouvernement devrait intervenir (« agenda ») et de celles dans lesquelles il ne le devrait pas (« non agenda »).

Viner remarque ainsi que « l'analyse peut, aux goûts de certains <sup>50</sup>, pencher trop fortement du côté des « non agenda » mais elle est libre de tout dogme à l'exception du dogme utilitariste » <sup>51</sup> ; et Viner de préciser : « je ne peux trouver dans ses écrits la moindre trace d'une formulation explicite par Bentham de la doctrine de l'harmonie naturelle des intérêts et une telle doctrine serait en parfaite contradiction non seulement avec sa philosophie du droit mais plus généralement avec l'ensemble de sa vision du monde. La foi dans l'harmonie naturelle provient toujours soit de la confiance dans l'intervention permanente d'un grand horloger salutaire, soit de la confiance dans les réalisations d'un processus évolutif ; or les « benthamistes » rejetaient la première et n'avaient pas encore entendu parler de la seconde » <sup>52</sup>.

Bentham semble en effet avoir raisonné autrement : on peut continuer d'admettre que les individus sont principalement, ou même exclusivement, égoïstes et nier cependant l'harmonie soit immédiate, soit progressive des intérêts. On déclare alors que, dans l'intérêt des individus, il faut identifier l'intérêt de l'individu avec l'intérêt général et qu'il appartient au législateur d'opérer cette identification : c'est ce processus – qualifié par Halévy de principe de « *l'identification artificielle des intérêts* » – qu'adopte pour l'essentiel Bentham.

<sup>50</sup> Viner semble viser ici la thèse de Halévy, selon laquelle Bentham adhérerait au principe de l'harmonie naturelle des intérêts en matière économique mais conclurait à la nécessité d'une identification artificielle des intérêts uniquement en matière juridique. Pour Viner, ce contraste, à l'intérieur même de la pensée de Bentham, serait excessif, voire même « tendancieux », Viner (1949) p. 205.

<sup>51</sup> Viner (1949), p. 207.

<sup>52</sup> Viner (1949), p. 205.



L'arbitre impartial et bien intentionné, en devenant réalité, doit sortir de son rôle passif d'observateur pour légiférer : « c'est uniquement par de bonnes lois qu'on peut former des hommes vertueux : tout l'art du législateur consiste à forcer les hommes, par le sentiment de l'amour d'eux-mêmes, d'être toujours justes les uns envers les autres »<sup>53</sup>. Cette phrase de Helvétius, dont Bentham s'est largement inspiré<sup>54</sup>, résume assez bien le défi constructiviste que ce dernier s'impose : bâtir de toute pièce un système juridique civil mais surtout pénal permettant d'assurer, par un dosage subtil de peines et de récompenses, cette convergence d'intérêts qui n'est pas spontanément acquise.

Ainsi le passage de la morale personnelle de l'utilité à la norme de justice sociale est-il assuré explicitement chez les utilitaristes primitifs par l'intervention régulatrice de l'État législateur ; l'éthique se confond désormais avec l'art du législateur et son champ de compétences n'a pas *a priori* d'autres limites que celles de l'application de la norme du plus grand bonheur du plus grand nombre. Cette norme définissant aussi bien l'éthique de l'élection à la majorité que celle du service public, Dupuy peut alors conclure que, « si le libéralisme des Welfare States d'aujourd'hui ressemble davantage à un interventionnisme omniprésent et tatillon qu'au laisser-faire, c'est autant à la vulgate utilitariste qu'aux politiques keynésiennes qu'il le doit »<sup>55</sup>.

C'est peut-être aussi cette prise de conscience que la fiction d'un arbitre impartial et bien informé pouvait justifier des réalités controversées, qui a conduit certains économistes utilitaristes à remettre en question la nécessité d'une fiction potentiellement si encombrante.

#### *L'embaras des successeurs*

La nécessité éthique d'un observateur impartial résulte de la contrainte logique, dans l'application de la norme utilitariste de justice, de collecter, auprès des membres de la société, un minimum d'informations ; si une telle contrainte pouvait être supprimée, la fiction de l'arbitre impartial n'aurait plus de raison d'être.

Pour parvenir à ce résultat, Vickrey, entre autres, n'a, semble-t-il, pas hésité à substituer à la fiction précédente une fiction au moins aussi précieuse : la collecte centralisée d'informations ne serait pas nécessaire, si l'idée pouvait être admise que les membres de la société possèdent le même profil psychologique, ce qui permet de leur attribuer des fonctions d'utilité identiques. Quant à Lerner, il refuse une telle simplifica-

<sup>53</sup> Helvétius, « *De l'esprit* » (1758), cité par Halévy (1995) vol. 1 p. 30.

<sup>54</sup> Sur ce point, cf. notamment Halévy (1995) vol. 1 p. 27 à 31.

<sup>55</sup> Dupuy (1992), p. 114.

tion et, plus, généralement, le principe même des comparaisons interpersonnelles d'utilité, tout en essayant de sauvegarder l'application de la norme utilitariste de justice.

*L'identité des préférences individuelles* a évidemment l'avantage important d'éliminer le problème des comparaisons interpersonnelles d'utilité, en définissant le seul cas pour lequel ces comparaisons ne soulèvent aucune difficulté. En utilisant la théorie neumanienne de l'utilité, Vickrey<sup>56</sup> considère ainsi le cas d'un candidat à l'émigration cherchant à évaluer l'attrait relatif des différentes « colonies » susceptibles de l'accueillir. Les individus se différenciant les uns des autres par leurs capacités productives, il existe une certaine inégalité dans la répartition des revenus à l'intérieur de chaque communauté, mais la distribution des talents entre « colons » est la même dans chacun des pays d'accueil. En aucun cas, l'immigrant potentiel ne sait quelle serait sa position dans la hiérarchie sociale mais il doit néanmoins déterminer le pays offrant les meilleures opportunités, sachant que chacun a mis en place une politique propre de transferts visant à corriger ces inégalités de revenus.

Autrement dit, l'immigrant de Vickrey va remplacer l'observateur impartial habituel avec la différence fondamentale que, par hypothèse, cet immigrant aura une structure des préférences identique à celle de chaque individu dans chaque pays d'accueil (ce qui suppose aussi une attitude identique vis-à-vis du risque).

Dans ces conditions, Vickrey soutient que « l'immigrant peut prendre sa décision en maximisant l'utilité espérée : les niveaux d'utilité sont ceux des différents membres d'une communauté donnée, les probabilités qui leur sont attachées correspondant à l'estimation par l'immigrant de la chance de pouvoir atteindre une situation comparable au cas où il choisirait la colonie en question. Si nous identifions le niveau de bien-être social d'une communauté avec l'attrait qu'elle exerce sur l'immigrant, nous observons que la fonction de bien-être social prend la forme d'une somme pondérée d'utilités individuelles. Si l'immigrant ignore complètement le rôle qu'il jouera dans son nouveau cadre de vie, il pondère à l'identique la situation de chaque individu et nous aboutissons ainsi à la somme benthamiste des utilités individuelles, celles-ci étant de type bernouillien »<sup>57</sup>.

L'obstacle des comparaisons interpersonnelles est donc contourné, voire supprimé, mais on peut pour le moins s'interroger sur la portée éthique du résultat : il est en effet généralement admis, comme le rappelle Pattanaik, que « toute tentative authentique d'évaluation du bien-

<sup>56</sup> Vickrey (1960), p. 459 à 461.

<sup>57</sup> Vickrey (1960), p. 460-461.

être social doit prendre en compte les différences dans la structure des préférences des individus »<sup>58</sup>.

C'est précisément ici que le modèle de Vickrey s'écarte radicalement de celui de Harsanyi, bien que tous deux adoptent l'analyse neumanienne de l'utilité en situation de risque : Harsanyi rejette l'hypothèse réductrice de préférences identiques, ce qui rend son analyse beaucoup plus riche et lui permet d'affronter les obstacles majeurs d'une réflexion éthique approfondie<sup>59</sup>.

Vickrey, quant à lui, a néanmoins la probité intellectuelle de reconnaître son embarras : « naturellement, écrit-il, tout ceci ne dit rien sur le sujet vexant des comparaisons interpersonnelles. L'exemple du colon immigrant évacue ce problème par l'hypothèse de goûts identiques et d'une mobilité sociale totale à long terme. Mais il vaut mieux, à mon avis, laisser la question des comparaisons interpersonnelles ... aux théologiens et aux métaphysiciens, jusqu'à ce que quelqu'un découvre un moyen nouveau de les définir de manière efficace »<sup>60</sup>.

Avant même la tentative de Vickrey, « à l'époque de Marshall, il était admis que le « calcul de félicité » reposait sur l'hypothèse selon laquelle tous les individus ont des fonctions d'utilité du revenu identiques »<sup>61</sup> ; mais l'introduction de cette hypothèse fut loin d'être anodine. À l'époque en effet, la réflexion concernait notamment la question de la répartition d'un revenu total fixe à l'aide d'impôts forfaitaires et en l'absence d'effets de revenu ou de substitution susceptibles de modifier l'incitation au travail, c'est-à-dire le niveau du revenu total à partager. Dans ce contexte épuré, si l'on retient d'une part l'hypothèse courante d'une utilité marginale décroissante du revenu et, d'autre part, l'hypothèse restrictive que chaque individu a la même fonction d'utilité, la maximisation de la somme des utilités individuelles, sous la contrainte d'un revenu total fixe, requiert non seulement l'égalisation des utilités marginales mais encore l'égalisation des utilités totales et donc celle des revenus.

Ainsi, « à partir de cette coïncidence très spéciale dans un cadre d'hypothèses extrêmement restrictives »<sup>62</sup>, l'analyse utilitariste débouche-t-elle sur la répartition strictement égalitaire du revenu. Mais, alors que ce résultat n'est obtenu que dans un cadre analytique schématique, celui-ci a conféré à l'utilitarisme une réputation égalitariste au point que « des adeptes de la réflexion utilitariste tels que Marshall et

<sup>58</sup> Pattanaik (1968), p. 300.

<sup>59</sup> Cf. I-B-2.

<sup>60</sup> Vickrey (1961), p. 296-297.

<sup>61</sup> Blaug (1981), p. 690.

<sup>62</sup> Sen (1973), p. 16.

Pigou furent attaqués par Robbins et d'autres pour leur usage de l'utilitarisme à des fins égalitaristes »<sup>63</sup>.

Bien que cette réputation se révèle en fait tout à fait usurpée<sup>64</sup>, son existence souligne l'importance de l'effet pervers engendré par l'hypothèse héroïque de l'identité des profils psychologiques individuels, dont le seul intérêt est d'éviter l'intervention intempestive d'un arbitre impartial et bien informé. Toutefois, face à cette alternative pour le moins contraignante, il convient de souligner l'originalité solitaire d'un auteur comme Lerner<sup>65</sup>.

Lerner refuse de choisir entre ces deux inconvénients que sont l'hypothèse de l'identité des fonctions d'utilité individuelle et la fiction d'un observateur impartial. Il suppose en revanche que nous ne pouvons jamais réellement découvrir quelle exacte satisfaction éprouve un individu donné en comparaison de celle que ressent un autre. Face à cette incertitude fondamentale et en dépit de son *refus des comparaisons interpersonnelles*, l'objectif de Lerner sera de proposer un critère de choix social de second rang permettant néanmoins à l'autorité compétente d'appliquer la norme utilitariste de justice.

L'analyse de Lerner concerne à nouveau la question du partage pur d'un revenu total fixe entre un nombre déterminé de personnes. En outre l'utilité de chaque individu est une fonction concave de son revenu personnel (utilité marginale décroissante) et indépendante des revenus d'autrui ; d'autre part, l'utilité sociale est bien la somme équipondérée des utilités individuelles. Lerner démontre alors la proposition suivante : « s'il est impossible, pour toute division du revenu, de découvrir lequel de deux individus a une utilité marginale du revenu plus élevée que celle de l'autre, la valeur probable des satisfactions totales est maximisée en divisant le revenu de façon égale »<sup>66</sup>.

L'interprétation de ce résultat, c'est-à-dire la restauration de l'égalitarisme en équiprobabilité, a été l'objet de nombreux commentaires. En particulier, la comparaison de la problématique de Lerner avec les approches déontologiques du type Harsanyi ou Rawls, semble *a priori* féconde : le recours à une situation de risque ou d'incertitude constitue pour chacun d'entre eux l'étape préalable de leur analyse. Cependant cette analogie reste superficielle, car les approches de Rawls et de Harsanyi d'une part et de Lerner d'autre part n'ont pas la même ambition : les premiers, loin d'éviter les comparaisons interpersonnelles, les intègrent à leur recherche d'une norme absolue de

<sup>63</sup> Sen (1973), p. 16.

<sup>64</sup> Cf. *infra* III-B.

<sup>65</sup> Lerner (1944).

<sup>66</sup> Lerner, (1944) p. 29 ; sur ce résultat, cf. également Sen (1968) p. 195.

justice<sup>67</sup>, alors que le second se contente de définir un critère de second rang relatif à un contexte épuré, en considérant comme vaine et insoluble la question des comparaisons interpersonnelles.

En définitive donc, cette question des comparaisons interpersonnelles d'utilité revient, à chaque étape de notre analyse de l'arithmétique interindividuelle des plaisirs, comme un leitmotiv obsédant. C'est elle qui rendait d'abord nécessaire la cardinalisation des préférences, c'est encore elle qui justifiait le recours à la fiction d'un observateur impartial. En l'absence de celui-ci, la réflexion utilitariste se trouve en effet minée par des hypothèses excessivement pauvres (identité des préférences individuelles) et/ou par des résultats tout à fait contingents (égalitarisme strict dans le partage « pur »).

Enfin et surtout, cette question des comparaisons interpersonnelles est à l'origine de la fracture décisive dans l'histoire de l'économie du bien-être entre les « anciens » et les « modernes » : si ce problème est jugé correctement résolu, la voie est alors ouverte pour la doctrine utilitariste, lui permettant de fonder en raison la norme sociale du plus grand bonheur du plus grand nombre. Si, en revanche, sa résolution n'est pas jugée satisfaisante, le passage de la morale personnelle de l'utilité à la norme utilitariste de justice reste définitivement obstrué.

On pourrait à ce stade affiner notre réflexion, en regroupant dans notre deuxième cercle de définition, inclus dans le premier, ceux qui acceptent seulement le principe des comparaisons interpersonnelles mais pas le troisième niveau de la réflexion utilitariste, c'est-à-dire la norme de justice sociale. En fait, la position largement dominante consiste à ne pas distinguer les deuxième et troisième niveaux, à l'image de la nouvelle économie du bien-être qui, avec le critère de Pareto, refusera tout autant la méthode – l'arithmétique interpersonnelle – que la conclusion – la norme du plus grand bonheur et ses dérapages potentiels – .

---

<sup>67</sup> Chez Harsanyi comme chez Rawls, la recherche d'une norme absolue de justice passe en effet par d'authentiques comparaisons interpersonnelles, effectuées par l'individu placé en position d'émettre un jugement moral : situation de risques équiprobables pour le premier (cf. I-B-2), situation d'incertitude non probabilisable caractéristique du « voile d'ignorance » pour le second. Ces comparaisons interpersonnelles sont en outre facilitées par le recours d'Harsanyi à un « postulat de similitude » entre les préférences personnelles de tous les individus « soumises aux mêmes lois psychologiques fondamentales » ; Rawls admet, quant à lui, une entorse au principe d'ignorance complète dans la position originelle, en introduisant des « biens premiers » dont « un homme rationnel préfère avoir plus que moins ».

## III. – LA NORME DU PLUS GRAND BONHEUR ET SES DÉRAPAGES POTENTIELS

Une fois définis les fondements de la doctrine – la morale personnelle – et son matériau essentiel – l'arithmétique des plaisirs –, on parvient au sommet de l'édifice : la norme de justice. Cette dernière étape est essentielle car elle permet d'évaluer l'ensemble de la réflexion utilitariste en fonction de sa finalité.

Or, à cet égard, la prudence est de mise, même si – hypothèse la plus favorable – le « calcul de félicité » a pu être mené avec toute la rigueur requise. Une retraite éventuelle de l'économiste vers une forme larvée d'intuitionnisme n'est pas en effet à exclure, compte tenu de la difficulté de supporter en pleine lumière l'éclat du jugement de valeur utilitariste : la norme du plus grand bonheur du plus grand nombre peut impliquer, d'une part, le sacrifice des droits de l'homme à la Société prise comme une entité supérieure et, d'autre part, l'exacerbation des différences interpersonnelles de condition en l'absence d'une norme plus explicite de répartition.

## A. Le sacrifice des droits de l'homme

Il est essentiel de repérer par quels détours du raisonnement une doctrine initialement fondée sur un individualisme débridé peut en arriver à nier l'individu, même dans ses prérogatives les plus élémentaires, par une norme de justice d'essence totalitaire. Du rapprochement détonant de ces deux caractéristiques, il résultera l'impossibilité de les concilier sans, dans le même temps, les hiérarchiser.

Pour ce faire, la matière première de notre réflexion sera issue des travaux de Sidgwick, dont la finalité constante est de défendre la norme du plus grand bonheur du plus grand nombre contre les morales intuitives ou égoïstes. L'ambition de Sidgwick<sup>68</sup> est en effet de définir des principes moraux à l'évidence incontestable et d'application universelle. Leur choix doit reposer sur leur compatibilité externe avec le sens moral commun et leur cohérence interne, quels qu'en soient le lieu et l'époque d'application.

*Trois concepts de référence*

Sidgwick se plaît à amalgamer logique et éthique en définissant des ensembles de type mathématique ou philosophique, dont les éléments sont parfaitement identiques les uns aux autres ; considérons avec lui

<sup>68</sup> Sidgwick (1907).

successivement trois références éthiques : le « bien », le « bien d'une vie dans son ensemble » et enfin le « bien universel ».

En ce qui concerne le « bien », un premier principe s'applique aux éléments qui le composent : « si une ligne de conduite qui est juste (ou injuste) à mon égard n'est pas juste (ou injuste) à l'égard de quelqu'un d'autre, la raison doit en incomber à une différence entre les deux situations qui n'est pas simplement le fait que moi et lui sommes des personnes différentes »<sup>69</sup>. Le premier principe éthique est celui d'impartialité à l'égard des personnes dans l'application des règles morales, qui va au-delà de la maxime traditionnelle « agis avec autrui comme tu voudrais qu'il agisse avec toi » : il introduit l'éventualité de différences de nature entre les situations qui viennent tempérer et rendre plus précise son application.

Mis en perspective sur l'ensemble de la vie d'un individu, ce premier principe en engendre un deuxième, lequel impose le traitement impartial de toutes les étapes de notre vie consciente : « après ne doit être, en tant que tel, ni valorisé, ni dévalorisé par rapport à maintenant »<sup>70</sup>. Cela ne signifie pas que des éléments d'appréciation supplémentaires – incertitude, évolution des situations, ... – ne permettent pas de trancher dans un sens ou dans un autre, mais la simple antériorité ou postériorité d'un « bien » n'est pas une justification suffisante pour considérer avec plus d'égards une période donnée par rapport à une autre. Dans la pratique, ce principe de « prudence » implique que le « plaisir ou le bonheur présents doivent avec raison être sacrifiés dans la perspective ultérieure d'un plaisir ou d'un bonheur plus grands »<sup>71</sup>.

La dernière étape est cruciale pour la mise en évidence de la rationalité éthique de la norme utilitariste ; « jusqu'ici, écrit Sidgwick, nous n'avons envisagé que le « bien dans son ensemble » d'un seul individu : mais, de même que cette notion était construite par comparaison et intégration des différents « biens » qui se succèdent dans la série de nos états conscients, de même nous pouvons bâtir la notion de « bien universel » par comparaison et intégration des biens de tous les êtres humains ou même de tous les êtres capables de sensations. Et ici à nouveau, comme dans le cas précédent, j'obtiens le principe évident en soi que le bien d'un individu quelconque n'a pas plus d'importance du point de vue (si je puis dire) de l'univers, que le bien de n'importe quel autre individu »<sup>72</sup>. Chacun étant astreint au bien en général et non à une de ses formes particulières, cette observation prend la forme d'un prin-

<sup>69</sup> Sidgwick (1907), p. 222-223.

<sup>70</sup> Sidgwick (1907), p. 224.

<sup>71</sup> Sidgwick (1907), p. 224.

<sup>72</sup> Sidgwick (1907), p. 225.

cipe de « bienveillance rationnelle » : sauf raisons particulières qui amèneraient à conclure différemment, « chacun est moralement tenu de considérer autant que le sien propre le bien de n'importe quel autre individu »<sup>73</sup>.

Cette notion de bienveillance rationnelle semble *a priori* correspondre à l'extrapolation, dans une dimension non plus intertemporelle mais interpersonnelle, du premier principe : si, individuellement, nous sommes capables de nous imposer sur l'instant un sacrifice en prévision d'un avantage futur bien supérieur, nous devons être aussi capables de sacrifier notre situation, si cette attitude est indispensable à la réalisation d'un plus grand bonheur pour un plus grand nombre, « du point de vue de l'univers » selon l'expression même de Sidgwick.

Or cette formule révèle un changement radical de perspective : il ne s'agit plus de l'optique de l'individu mais de celle de l'univers, plus précisément de celle de la société prise dans son ensemble. De même qu'un individu compare ses plaisirs présents et futurs avec ses peines présentes et futures, de même la société, par le bras séculier de l'Etat législateur, peut mettre en balance les utilités et les désutilités des membres qui la composent et en déduire le mode d'organisation sociale qui assure le plus grand bonheur du plus grand nombre. « Le principe de choix pour un groupe, souligne Rawls<sup>74</sup>, est interprété comme étant une extension du principe de choix valable pour un individu. »

#### *Le nœud de la contradiction*

Ce faisant, se noue alors la formidable contradiction de la doctrine utilitariste : l'individu est invité, dans un premier temps, à « rechercher le plaisir et à éviter la peine » sans se préoccuper outre mesure de la compatibilité de son comportement avec celui d'autrui ; cet individualisme insouciant a pour revers, dans un second temps, la subordination de la personne humaine à la société, par la contrainte qu'elle peut lui imposer afin d'assurer artificiellement l'identification des intérêts au bien commun.

L'alternative est alors la suivante :

– Ou bien chaque individu applique volontairement le principe de « bienveillance rationnelle », ce qui suppose une haute conscience civique et morale, mais un asservissement conscient et total des individus multiples à l'égard de l'Entité collective.

<sup>73</sup> Sidgwick (1907), p. 225.

<sup>74</sup> Rawls (1987), p. 50.



– Ou bien, éventualité la plus réaliste, ce principe n'est pas spontanément suivi, auquel cas l'intervention du législateur assure de manière coercitive son respect.

Excessivement individualiste dans ses prémices, la doctrine utilitariste peut ainsi justifier les modes d'organisation sociale les plus totalitaires ; ce qualificatif n'est pas excessif car, dans les termes de l'alternative précédente, l'individu ne voit aucune prérogative élémentaire lui être garantie. Face à la « raison d'État », l'impartialité de la norme utilitariste ne signifie pas en effet le respect des droits fondamentaux de la personne humaine (intégrité physique, liberté d'expression,...), même si un utilitariste pourrait objecter qu'un état social qui ne respecterait pas de tels droits – en tolérant la torture, par exemple – serait probablement inapte à assurer le plus grand bonheur du plus grand nombre.

Ce principe d'impartialité implique simplement que, « chacun comptant pour un et personne pour plus d'un », les hommes sont parfaitement indifférenciés et donc substituables au regard de la collectivité ; de ce fait, le sacrifice de n'importe lequel d'entre eux peut être exigé si le respect de la norme utilitariste de justice en dépend. « Nous devons accepter les plus grands avantages des autres comme une raison suffisante pour des attentes plus faibles dans tout le cours de notre vie. Il s'agit sûrement là d'une exigence extrême »<sup>75</sup>.

Dans cette comptabilité froide et anonyme où les minorités s'inclinent devant le plus grand nombre, où une unanimité factice peut même se faire contre un seul, l'individu n'est pas une fin en soi mais un simple moyen, c'est-à-dire une cellule ou une particule élémentaire du Léviathan social. C'est peut-être Edgeworth qui a le mieux exprimé, de manière involontaire mais prémonitoire, la modernité technocratique de la norme utilitariste : « la « mécanique sociale » peut un jour prendre sa place à côté de la « mécanique céleste »<sup>76</sup> sous l'empire d'un principe supérieur ambivalent, référence suprême de la science morale comme de la science physique : comme les mouvements de chaque particule, contraints ou libres, dépendent en permanence, dans l'univers matériel, d'un potentiel maximum d'énergie accumulée, de même, les évolutions de chaque être, qu'elles soient isolées de manière égoïste ou liées par sympathie, peuvent en permanence réaliser le potentiel maximum de plaisir, passion divine de l'univers »<sup>77</sup>.

Outre cette réduction de l'individu à l'état de particule dans le mouvement « brownien » de la société, la doctrine utilitariste a pour autre

<sup>75</sup> Rawls (1987), p. 208.

<sup>76</sup> Les expressions « mécanique sociale » et « mécanique céleste » apparaissent en français dans le texte original en anglais.

<sup>77</sup> Edgeworth (1881), p. 12.

inconvenient d'impliquer, le cas échéant, l'exacerbation des différences interindividuelles de condition.

### B. L'exacerbation des différences de condition

Quand se trouve défini l'état social assurant le plus grand bonheur du plus grand nombre, reste encore à préciser le critère selon lequel la répartition entre les membres de la collectivité va être effectuée. Or il est clair que la norme utilitariste s'intéresse d'abord à la maximisation de la somme des utilités individuelles, sans vraiment se préoccuper de la répartition interpersonnelle de cette somme. Par rapport à cette question centrale pour la dimension sociale de la justice, seules quelques précisions périphériques sont abordées, tandis que l'exacerbation éventuelle des différences de condition est une simple conséquence des différences de profil psychologique généralement admises entre individus.

#### *Quelques précisions périphériques*

Le mode d'agrégation des utilités individuelles et la nature de la population concernée constituent les deux caractéristiques indispensables au « calcul de félicité », qui conduisent les utilitaristes à n'évoquer que de manière très marginale la question de la répartition du plus grand bonheur.

En ce qui concerne le premier point, le principe d'impartialité de Sidgwick joue un rôle essentiel : dans l'agrégation des utilités individuelles, une pondération strictement identique doit être accordée aux préférences de chacun. On peut d'ailleurs remarquer que cette agrégation équi pondérée peut prendre, en toute rigueur, une forme différente – multiplicative, par exemple – de la somme des utilités individuelles qu'avait envisagée Bentham.

Comme le souligne Arrow, « l'argument, selon lequel tous les individus devraient être considérés sur un pied d'égalité dans le jugement social, aboutit uniquement à la conclusion que la fonction de bien-être social devrait être une fonction symétrique des utilités individuelles et non qu'elle devrait en être la somme »<sup>78</sup>.

Quant à la population concernée par l'application de la norme utilitariste, elle peut être composée des seuls êtres humains ou bien élargie, comme l'envisage Sidgwick<sup>79</sup>, à tous les êtres « capables de sensations » de plaisirs et de peines. Cette dernière éventualité peut *a priori* sembler incongrue, dans la mesure où elle aboutit notamment à intégrer

<sup>78</sup> Arrow (1984), p. 121.

<sup>79</sup> Cf. la définition par Sidgwick du « bien universel » (*supra* III-A-1).

dans l'arithmétique utilitariste le sort des animaux ; en fait, elle ouvre la voie à un débat philosophique important : « est-on prêt, remarque Dupuy, à justifier en raison le mal qu'on fait aux animaux pour le plus grand bien des humains ? L'éthique du gavage des oies ou du veau aux hormones ne serait pas prise très au sérieux chez nous, mais ne nous y trompons pas : c'est bel et bien d'un sujet grave qu'il s'agit, puisque c'est la rationalité du passage du sacrifice humain au sacrifice animal qui est en question. L'immolation du bouc est-elle plus légitime que celle de l'étranger ? »<sup>80</sup>. Le principe d'impartialité à l'égard d'individus reconnus égaux et substituables suppose donc que soit au préalable défini l'ensemble sur lequel il s'applique et la nature des individus « autres » qui en sont exclus<sup>81</sup>.

#### *Une conséquence des différences de profil psychologique*

L'ensemble des observations précédentes reste insuffisante pour définir la répartition interpersonnelle du plus grand bonheur. Dans le cas général où les profils psychologiques sont considérés comme différents, le caractère essentiel de cette répartition peut néanmoins être mis en lumière à partir du problème simplifié du partage pur d'un revenu total fixe.

Supposons une collectivité réduite à deux individus dont l'un (*i*) obtient un niveau d'utilité deux fois plus élevé que l'autre (*j*), pour n'importe quel niveau de revenu. On pourrait penser *a priori* qu'une norme de justice, quelle qu'elle soit, aurait au moins pour effet de ne pas renforcer la disparité « naturelle » entre les niveaux de satisfaction de *i* et de *j*. Tel n'est pas le cas avec la norme utilitariste : comme la maximisation de l'utilité totale de la collectivité sous contrainte d'un revenu fixe à partager implique l'égalisation des utilités marginales, le revenu le plus élevé doit donc être attribué à celui qui, par sa plus grande capacité de jouissance, permet à la satisfaction collective d'atteindre un niveau plus élevé.

En d'autres termes, la répartition du plus grand bonheur du plus grand nombre peut aboutir à exacerber les différences originelles de condition entre individus : si, par exemple, l'individu *i*, en bonne santé, est susceptible à ce titre de retirer plus de satisfactions d'un même revenu que *j*, handicapé, la norme utilitariste, insensible à la genèse des

<sup>80</sup> Dupuy (1992), p. 123.

<sup>81</sup> Une autre difficulté apparaît lorsqu'il s'agit d'appliquer la norme du plus grand bonheur du plus grand nombre au problème de la taille optimale de la population. Sur cette controverse d'ordre démographique à l'intérieur de la doctrine utilitariste, cf. notamment Sidgwick (1907) p. 230-231 et Rawls (1987) p. 191 à 197.

situations personnelles, conduira à opérer un transfert de revenus du handicapé vers le valide. « On peut craindre, conclut non sans euphémisme Wolfelsperger, qu'un tel résultat ne corresponde à aucune idée courante de la justice »<sup>82</sup>.

Cet aspect rédhibitoire de la norme utilitariste pousse également Sen à proposer une disposition minimale dont le respect empêcherait la réflexion éthique de dérapier vers de telles extrémités. Cette disposition qu'il appelle « axiome faible d'équité » prend la forme suivante : « Soit un individu  $i$  disposant pour chaque niveau de revenu personnel d'un niveau de bien-être inférieur à celui d'un individu  $j$ . Dans la répartition d'un revenu global entre  $n$  personnes comprenant  $i$  et  $j$ , la solution optimale doit attribuer à  $i$  un niveau de revenu plus élevé qu'à  $j$  »<sup>83</sup>.

Cet axiome peut être qualifié de faible, dans la mesure où il se contente d'indiquer une tendance – l'homogénéisation des conditions –, sans préciser jusqu'à quel point cette tendance doit être poursuivie ; mais son introduction reste relativement arbitraire : un tel axiome aurait mérité un fondement rationnel approfondi, en l'absence duquel il n'a d'autre justification que sa conformité avec un principe de morale intuitive de simple « bon sens »<sup>84</sup>.

Quoi qu'il en soit, cette remarque sur la contribution de Sen n'atténue pas les observations critiques à l'égard de l'utilitarisme, dont l'axiome faible d'équité ne constitue qu'un correctif. Sauf à accepter l'hypothèse par défaut de l'identité des profils psychologiques individuels<sup>85</sup>, cette doctrine justifie si nécessaire l'exacerbation des différences de bien-être, ce qui ne peut manquer de saper un peu plus sa portée éthique d'ensemble.

#### CONCLUSION / L'EXTRAORDINAIRE AMBIVALENCE DE L'UTILITARISME

Au total, le trait caractéristique de l'utilitarisme, au moins dans sa version centrée sur l'acte, est bien, nous semble-t-il, son extraordinaire ambivalence :

– D'un côté, une foi inébranlable, typique de la modernité occidentale, dans la raison humaine, capable de féconder une science authentique des mœurs et de la législation ;

– De l'autre, une norme sociale digne des « théories anciennes où le tout social (et politique) est premier (par contraste avec) les théories

<sup>82</sup> Wolfelsperger (1980), p. 98.

<sup>83</sup> Sen (1973), p. 18.

<sup>84</sup> Pour de plus amples développements sur l'axiome faible d'équité, cf. notamment Sen (1973) – p. 18-23 et 43-45 – et Sen (1974).

<sup>85</sup> Cf. *supra* II-B-2.

modernes où ce sont les droits de l'homme individuel qui sont premiers et déterminent la nature des bonnes institutions politiques »<sup>86</sup>.

C'est ce curieux alliage entre rationalité instrumentale et « holisme » social qui fascine et qui effraie dans l'analyse de la doctrine utilitariste. Ainsi que le remarque Scanlon, « l'utilitarisme occupe une place centrale dans la philosophie morale de notre temps. Ce n'est pas le point de vue que la plupart des individus partagent ; certainement un très petit nombre se réclamerait de l'utilitarisme d'acte. Mais, pour une beaucoup plus grande partie d'entre eux, c'est la position vers laquelle ils se trouvent eux-mêmes acculés quand ils tentent de donner une référence théorique à leurs convictions éthiques. À l'intérieur de la philosophie morale, l'utilitarisme représente une position contre laquelle il faut lutter si on veut la contourner »<sup>87</sup>.

Pour notre part, c'est par la reconnaissance explicite de trois cercles concentriques dans la réflexion utilitariste, susceptibles de mieux en repérer les controverses et les variantes principales, que nous avons tenté de mettre en pratique la sage recommandation de Scanlon. Au-delà de la salutaire remise en ordre de nos propres convictions éthiques, notre démarche n'avait en effet pas d'autre but que de permettre éventuellement au lecteur de mieux évaluer soit son degré d'adhésion au moins implicite, soit sa capacité propre de résistance à l'attraction de cette référence essentielle pour l'économie normative et la philosophie morale contemporaines.

#### RÉFÉRENCES BIBLIOGRAPHIQUES

- ARMSTRONG W. E., « Utility and the Theory of Welfare », *Oxford Economic Papers*, n° 3. 1951.
- ARROW K., « Some Ordinalist-Utilitarian Notes on Rawls's Theory of justice », *Journal of Philosophy* (mai) 1973.
- ARROW K., « Formal Theories of Social Welfare », in « *Social Choice and Justice – Collected Papers of K. J. Arrow* », Basil Blackwell. 1984.
- ARROW K., SCITOVSKY T., Ed., « *Readings in Welfare Economics* », Londres, George Allen and Unwin. 1969.
- ARROW K. J., INTRILIGATOR M. D., Ed., « *Handbook of Mathematical Economics : Volume III* », Amsterdam, North-Holland. 1986.
- BLAUG M., « *La pensée économique – origine et développement* », Paris, Economica. 1981.
- BLAUG M., « *La méthodologie économique* », Paris, Economica. 1982.
- BRANDT R. B., « *Ethical Theory* », Londres, Englewood Cliffs N. J., Prentice Hall Inc. 1959.
- BRAITHWAITE R., Ed., « *The Foundations of Probability and Other Logical Essays* », Londres, Kegan Paul. 1931.

<sup>86</sup> Dumont (1978), p. 19.

<sup>87</sup> Scanlon (1982), p. 103.

- DIAMOND P. A., « Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility : Comment », *Journal of Political Economy*, n° 75 (octobre). 1967.
- DUMONT L., « La conception moderne de l'individu », *Esprit* (février). 1978.
- DUPUY J.-P., « *Le sacrifice et l'envie – Le libéralisme aux prises avec la justice sociale* », Paris, Calmann-Lévy. 1992.
- EDGEWORTH F. T., « *Mathematical Psychics* », Londres, Kegan Paul. 1881.
- ELSTER J., « Sour Grapes – Utilitarianism and the Genesis of Wants », in Sen A. K. and Williams B., Ed., *Utilitarianism and Beyond*, Paris, Maison des Sciences de l'Homme. Cambridge University Press. 1982.
- ELSTER J., ROEMER J. E., Ed., « *Interpersonal Comparisons of Well-Being* », Cambridge University Press. 1991.
- GAMEL C., « L'apport de la philosophie sociale à l'économie du bien-être : une question de « procédure » ? », *Économies et Sociétés*, série PE n° 14 (janvier). 1991.
- GAMEL C., « *Économie de la justice sociale – Repères éthiques du capitalisme* », Paris, Cujas. 1992.
- GOODMAN L. AND MARKOWITZ H., « Social Welfare Functions Based on Individual Rankings », *American Journal of Sociology*, n° 58. 1952.
- GUITTON H., MARGOLIS J., Ed., « *Économie publique* », CNRS 1968.
- HAHN F., « On Some Difficulties of the Utilitarian Economist », in Sen A. K. and Williams B., Ed., *Utilitarianism and Beyond*, Paris, Maison des Sciences de l'Homme. Cambridge University Press. 1982.
- HALÉVY E., « La révolution et la doctrine de l'utilité », Thèse lettres, Paris, Félix Alcan Editeur. 1900.
- HALÉVY E., « *La formation du radicalisme philosophique* », volume 1, « *La jeunesse de Bentham 1776-1789* », volume 2, « *L'évolution de la doctrine utilitaire (1789-1815)* », volume 3, « *Le radicalisme philosophique* », Paris, PUF ; 1<sup>re</sup> édition (1901) pour les deux premiers volumes et (1904) pour le troisième. Paris, Félix Alcan Editeur. 1995.
- HAMMOND P., « Utilitarianism, Uncertainty and Information », in Sen A. K. and Williams B., Ed., *Utilitarianism and Beyond*, Paris, Maison des Sciences de l'Homme. Cambridge University Press. 1982.
- HAMPSHIRE S., « Morality and Convention », in Sen A. K. and Williams B., Ed., *Utilitarianism and Beyond*, Paris, Maison des Sciences de l'Homme. Cambridge University Press. 1982.
- HARROD R. F., « Utilitarianism Revised », *Mind*, n° 45. 1936.
- HARSANYI J. C., « Cardinal Utility in Welfare Economics and in the Theory of Risk Taking », *Journal of Political Economy* (octobre) 1953.
- HARSANYI J. C., « Cardinal Welfare, Individualistic Ethics and Interpersonal Comparisons of Utility », *Journal of Political Economy* (août) 1955, in Phelps. Ed., *Economic Justice*, Penguin Education, 1973.
- HARSANYI J. C., « Morality and the Theory of Rational Behaviour », 1977, in Sen A. K. and Williams B., Ed., *Utilitarianism and Beyond*, Paris, Maison des Sciences de l'Homme. Cambridge University Press. 1982.
- HAYEK F. A., « *Droit, législation et liberté* », tome 2, « *Le mirage de la justice sociale* », Paris, PUF, 1981 ; traduction de « *Law, Legislation and Liberty* » (1976).
- KELMAN S., « Cost-Benefit Analysis : An Ethical Critique », *Regulation* (janvier). 1981.

- LANGLOIS R., « Cost-Benefit Analysis, Environmentalism and Rights », *Cato Journal* (printemps). 1982.
- LERNER A. P., « *The Economics of Control* », MacMillan. 1944.
- MILL J. S., « *L'utilitarisme* », Paris, Garnier-Flammarion, 1968 ; traduction de « *Utilitarianism* » (1861).
- MIRRELES J. A., « The Economic Uses of Utilitarianism », in Sen A. K. and Williams B., Ed., *Utilitarianism and Beyond*, Paris, Maison des Sciences de l'Homme. Cambridge University Press. 1982.
- PATTANAIK P. K., « Risk, Impersonality and the Social Welfare Function », *Journal of Political Economy* (décembre) 1960, in Phelps, Ed., *Economic Justice*, Penguin Education, 1973.
- PHELPS E., « *Economic Justice* », Ed., Penguin Education, 1973.
- RAMSEY F. P., « Truth and Probability », in Braithwaite R., Ed., *The Foundations of Probability and Other Logical Essays*, Londres, Kegan Paul, 1931.
- RAWLS J., « *Théorie de la justice* », Paris, Éditions du Seuil, 1987 ; traduction de « *A Theory of Justice* », 1971.
- ROEMER J. E., « *Theories of Distributive Justice* », Harvard University Press, 1996.
- ROTHENBERG J., « *The Measurement of Social Welfare* », Englewood Cliffs N. J., Prentice Hall Inc, 1961.
- SCANLON T. M., « Contractualism and Utilitarianism », in Sen A. K. and Williams B., Ed., *Utilitarianism and Beyond*, Paris, Maison des Sciences de l'Homme. Cambridge University Press. 1982.
- SEN A. K., « Les préférences des planificateurs : optimalité, répartition et utilité sociale », in Guitton H. et Margolis J. Ed., *Économie publique*, CNRS, 1968.
- SEN A. K., « *Collective Choice and Social Welfare* », Holden Day, 1970.
- SEN A. K., « *On Economic Inequality* », Oxford, Clarendon Press, 1973.
- SEN A. K., « Rawls versus Bentham : An Axiomatic Examination of the Pure Distribution Problem », *Theory and Decision* (février-avril) 1974.
- SEN A. K., WILLIAMS B. Ed., « *Utilitarianism and Beyond* », Paris, Maison des Sciences de l'Homme et Cambridge University Press, 1982.
- SEN A. K., « Social Choice Theory », in Arrow K. J. et Intriligator M. D., Ed., *Handbook of Mathematical Economics : Volume III*, Amsterdam, North-Holland, 1986.
- SIDGWICK H., « *The Method of Ethics* », MacMillan 1907, 7<sup>e</sup> édition partiellement reproduite in Phelps E. Ed., *Economic Justice*, Penguin Education, 1973.
- VICKREY W. S., « Utility, Strategy and Social Decision Rules », *Quarterly Journal of Economics*, n° 74, 1960, in Arrow K. and Scitovsky T., Ed., *Readings in Welfare Economics*, Londres, George allen and Unwin, 1969.
- VICKREY W. S., « Risk, Utility, and Social Policy », *Social Research* (été) 1961, in Phelps E. Ed., *Economic Justice*, Penguin Education, 1973.
- VINER J., « Bentham and J. S. Mill : The Utilitarian Background », *American Economic Review* (mars) 1949, in Phelps E., Ed., *Economic Justice*, Penguin Education, 1973.
- VON NEUMANN J. ET MORGENSTERN O., « *Theory of Games and Economic Behaviour* », Princeton 1947.
- WEYMARK J. A., « A Reconsideration of the Harsanyi-Sen Debate on Utilitarianism », in Elster J. et Roemer J. E., Ed., *Interpersonal Comparisons of Well-Being*, Cambridge University Press, 1991.
- WOLFELSPERGER A., « *Économie des inégalités de revenus* », Paris, PUF, 1980.