Université Paris-Est-Créteil-Val de Marne

Faculté des Lettres, Langues et Sciences Humaines

Ecole Doctorale « Cultures et Sociétés »

Thèse présentée en vue de l'obtention du grade de Docteur en Philosophie par :

Danielle Zwarthoed

Le choix collectif dans la philosophie politique contemporaine

Des fondements philosophiques de la théorie du choix social à l'évaluation démocratique des capabilités d'Amartya Sen

Directeur de thèse : Ali Benmakhlouf, Professeur de Philosophie à l'Université Paris-Est-Créteil

Le 14 septembre 2013

Jury :

Didier DEBAISE, Chercheur qualifié (FNRS) en Philosophie à l'Université Libre de Bruxelles

Franck FISCHBACH, Professeur de Philosophie à l'Université de Strasbourg

Claude GAUTIER, Professeur de Philosophie à L'Ecole Normale Supérieure de Lyon

Frédéric GROS, Professeur de Philosophie à l'Université Paris-Est-Créteil

Remerciements

Je tiens à remercier avant tout M. Ali Benmakhlouf d'avoir accepté de diriger cette thèse sur un sujet que j'avais choisi et de m'avoir accordé sa confiance et son soutien tout au long de mon doctorat.

Je remercie également Jean Robelin, mon directeur de Master, qui a continué à me conseiller et à m'aider ensuite, ainsi que Franck Fischbach, qui m'a donné l'opportunité de présenter mes travaux lors du Séminaire de Philosophie Sociale à l'Université de Nice.
Sans oublier celle qui a eu la patience de relire et de corriger cette thèse, Anne-Laure Guichard, que je remercie vivement ici.

Ce travail a bénéficié des séances de tutorat avec Richard Arena, Ragip Ege et Axel Gosseries, organisées lors de la XVe Université d'Eté en Histoire de la Pensée Economique et Philosophie Economique, à Lille. Les échanges que j'ai eus avec Antoinette Baujard, Thomas Ferretti, Muriel Gilardone, Herrade Igersheim, Dorian Jullien et Soumaya Mestiri ont aussi été très enrichissants. Je tiens à les remercier tous ici.

Je suis enfin reconnaissante à ma famille et à mes amis de leur patience, de leur compréhension, de leur soutien et de leurs encouragements. Sans eux, ce projet n'aurait pu être mené à bien. Je remercie Balla, qui a été exemplaire par sa loyauté et son affection. Colette, Julie, Michelle et Pierre ont su dissiper mes doutes et me rassurer lorsque cela a été nécessaire.

Il ignorait tout de cet effroyable négoce sur les valeurs déclassées et sur l'achat de créances, il vivait plus haut, dans un songe souverain de justice. L'idée de charité le blessait, le jetait hors de lui : la charité, c'était l'aumône, l'inégalité consacrée par la bonté ; et il n'admettait que la justice ; les droits de chacun reconquis, posés en immuables principes de la nouvelle organisation sociale. Aussi, à la suite de Karl Marx, avec lequel il était en continuelle correspondance, épuisait-il ses jours à étudier cette organisation, modifiant, améliorant sans cesse sur le papier la société de demain, couvrant de chiffres d'immenses pages, basant sur la science l'échafaudage compliqué de l'universel bonheur.

Emile Zola, L'Argent

« You are to be in all things regulated and governed » said the gentleman, « by fact. We hope to have, before long, a board of fact, composed of commissioners of fact, who will force the people to be a people of fact, and of nothing but fact. You must discard the word Fancy altogether. You have nothing to do with it. […] You must use, said the gentleman, "for all these purposes, combinations and modifications (in primary colours) of mathematical figures which are susceptible of proof and demonstration. This is the new discovery. This is fact. This is taste"

Charles Dickens, Hard Times.

Résumé

Les fondements philosophiques de la théorie du choix social sont l'objet de cette thèse. La théorie du choix social est un champ de l'économie normative qui traite de l'agrégation des préférences individuelles. L'objectif de ce travail est d'analyser les postulations philosophiques de la théorie du choix social afin de comprendre dans quelle mesure celle-ci pourrait contribuer à une théorie de la justice sociale basée sur les capabilités d'Amartya Sen. Ce travail est par conséquent élaboré à partir de l'idée d'une « approche comparative » de la justice sociale, défendue par Amartya Sen, et que ce dernier oppose à l' « approche transcendantale » de John Rawls. Nous tentons d'y préciser l'interprétation de la théorie du choix social requise pour spécifier l'approche par les capabilités, et en particulier l'évaluation et la construction d'un indice de capabilités et de fonctionnements. Ce travail de thèse défend l'hypothèse selon laquelle la tension apparente entre l'agrégation des préférences et l'approche par les capabilités est due à une interprétation étroite du cadre conceptuel de la théorie du choix social. En effet, l'approche par les capabilités est généralement comprise comme incompatible avec le préférentialisme de la théorie du choix social. Cela est dû à ce que l'approche par les capabilités s'est construite en partie comme une réponse au problème des préférences adaptatives. Cette thèse consiste donc à élargir le champ des interprétations de la théorie du choix social et de son cadre conceptuel, principalement de sa base d'information.
La première partie de ce travail de recherche traite la question suivante : les préférences sont-elles déterminées par une source individuelle pouvant être pensée indépendamment de sa position sociale et économique ? Afin de montrer que ce n'est pas nécessairement le cas sur le plan logique, nous y analysons trois types de base informationnelle : les utilités cardinales, les préférences ordinales et les capabilités.
L'objectif de la seconde partie est de déterminer ce que les préférences décrivent. Nous y analysons d'une part la nature de la préférence elle-même dans ce contexte. S'agit-il d'une décision déterminant une action, d'un désir, d'un état mental ou encore d'un jugement de valeur ? D'autre part, les différents critères éthiques de préférence sont étudiés, à savoir le plaisir hédoniste, la satisfaction des désirs et un critère de bien-être objectif. Cette exploration nous amène à défendre la conclusion suivante : la conception des préférences la plus en phase avec la structure formelle de la théorie du choix social est une conception comparative requérant d'excellentes conditions cognitives pour que les préférences puissent être considérées à proprement parler comme les véritables préférences de l'agent.
La troisième partie revient à l'approche des capabilités. L'argumentation s'appuie sur les conclusions des deux parties précédentes afin de jeter les bases d'une théorie de la justice démocratique et non-idéale basée sur les capabilités d'Amartya Sen. Dans cette partie, nous montrons qu'une évaluation des capabilités indépendante des préférences tend à nier l'importance de la liberté et de la qualité d'agent dans l'approche par les capabilités. Nous défendons donc l'hypothèse selon laquelle les capabilités et les fonctionnements comme objets de préférence permettent de filtrer celles-ci afin de parer au problème des préférences adaptatives.

MOTS-CLEFS : choix social – capabilités – philosophie économique – préférence – Amartya Sen – théories de la justice

Social choice in contemporary political philosophy: from preference aggregation to a democratic theory of justice

This dissertation examines the philosophical foundations of social choice theory. Social choice theory is the area of normative economics which is concerned with the aggregation of individual preferences. The aim of this work is to investigate the philosophical assumptions of social choice theory in order to understand to what extent it can contribute to a theory of justice based on capabilities. Therefore, the dissertation is build up on Amartya Sen's idea of a "comparative approach" of justice, as opposed to the Rawlsian "transcendental approach". It is an attempt to precise which understanding of social choice theory is required to specify the capability approach, especially the evaluation and the indexing of capabilities. In this dissertation, we argue that the apparent tension between preference aggregation and capability approach is due to a narrow interpretation of social choice theory's conceptual framework. Capability approach is generally conceived as non-compatible with the preferentalism of social choice theory: after all, capabilities are seen as a response to the recurring problem of adaptive preferences. This dissertation thus consists in widening the scope of interpretations of social choice theory framework. This research deals mainly with the informational basis of social choice theory.

This dissertation is in three parts. The first part tackles the following problem: are preferences determined by an individual source that can be thought independently of its social and economic position? To answer these questions, three kinds of informational basis in social choice theory and normative economics are investigated: cardinal utilities, ordinal preferences and capabilities.

The second part aims at defining what preferences do describe in this context. Firstly, the nature of preference itself is examined: can it be assimilated to choice? Or is it a mere evaluation? A desire? A mental state? This analysis points out the comparative structure of preferences. Secondly, the various ethical criteria of preference are investigated: hedonistic pleasure, desire satisfaction and objective well-being. We argue that preferences are better conceived as comparative evaluation and require actually excellent cognitive conditions to be truly the agent's *own real* preferences.

The third part goes back to capability approach. The argument relies on the previous results to build up a first account of a democratic non-ideal theory of justice based on capabilities. In this part, we show that a preference-independent capability evaluation turns out to dismiss the importance of freedom and agency in capability approach. Then we argue that capabilities and functionings as an object for preferences do provide a first filter against adaptive preferences.

KEYWORDS: social choice – capabilities – Economics and Philosophy – preference – Amartya Sen – theory of justice

Thèse préparée au sein de l'Equipe d'Accueil Lettres Idées Savoirs (LIS, EA n° 4395) - Université Paris-Est-Créteil-Val de Marne - 61, avenue Charles de Gaulle – 94010 Créteil (France)

Table des matières

Remerciements 3
Résumé 7
Table des matières 10
Introduction 17
1. Le pari de la rationalité collective 17
2. Un peu d'histoire 20
L'étude des procédures de vote : de Pline le Jeune à Condorcet 20
L'agrégation utilitariste 25
La rupture avec la philosophie : naissance de l'économie du bien-être et de la théorie du choix social 27
Le théorème d'impossibilité du choix social (1951) 29
3. Philosophie et théorie du choix social 33
4. Un cadre conceptuel et de multiples possibilités 36
5. Plan : objet et méthodes 39
Ière partie : L'individu dans la théorie du choix social, des origines aux plus récents développements 45
Chapitre 1 : De l'attribution des utilités et des ophélimités aux individus 47
1.1. Utilités individuelles, mécanique d'agrégation et dispositifs institutionnels 47
1.1.1. Le critère d'utilité : plaisir et peine des corps individuels 48
1.1.2. Le calcul utilitariste et ses règles 50
1.1.3. Le cercle utilitariste : les dispositifs institutionnels et leurs effets sur la formulation des attentes en termes d'utilités 51
1.2. Ophélimités et préférences ordinales : l'individu dépsychologisé 54
Chapitre 2: L'individu dans la théorie du choix social arrovienne 59
2.1. L'individu, support des préférences 60
2.1.1. Arrow, défenseur de la primauté del'évaluation individuelle 60
a) L'électeur souverain 60
b) Une approche mécaniste 63
c) Les jugements de valeur seraient individuels 64
2.1.2. La justification épistémique de la primauté de l'évaluation individuelle 66
a) L'individu comme entité observable 66
b) Conditions réelles VS conditions idéales des évaluations individuelles 67
L' « idéalisme » de Rousseau 69
L' « idéalisme » de Kant 71

2.1.3. La justification normative de la primauté de l'évaluation individuelle : souveraineté de l'agent et pluralisme 72

a) Critique arrovienne de l'hypothèse de maximisation de l'intérêt personnel 73

b) Critique des préférences rawlsiennes pour les biens premiers 74

2.2. Une primauté de l'évaluation individuelle à nuancer 75

2.2.1. Le sens et la portée de la rationalité collective 75

a) L'action collective n'est pas qu'un composé d'actions individuelles : théorie du choix social VS théorie des choix publics 76

b) Jugements de bien-être collectif et choix social : il y a un sens à parler d'action collective 78

2.2.2. Evaluations individuelles et attitudes sociales 80

2.2.3. Le caractère artificiel de la factorisation de la société en composantes individuelles 83

2.3. Comparaisons interindividuelles et intégrité individuelle 86

2.3.1. Un individu sans substance 86

2.3.2. Les comparaisons interpersonnelles : enrichir la description de l'individu 89

2.3.3. Nous ne comparons pas des individus, mais des qualités 90

a) Comparaisons interpersonnelles et sympathie étendue 90

b) Comparaisons interpersonnelles, incommensurabilité des individus et autonomie 92

Chapitre 3 : Evaluation individuelle et évaluation sociale dans l'approche par les capabilités 95

3.1. Des préférences individuelles aux évaluations positionnelles 97

3.1.1. Les attributs objectifs et subjectifs de l'individu dans l'évaluation de l'avantage 98

3.1.2. Le problème de la distorsion des évaluations individuelles : illusion objective et capabilités 101

a) Morale traditionnelle et arguments cognitifs 101

b) Le prisme marxien 103

L'illusion objective 103

La technologie sociale 106

c) Capabilités et individualisme méthodologique 108

3.1.3. Eléments de méthode : néo-holisme et infra-individualisme 110

3.1.4. L'objectivité positionnelle 114

3.2. La construction de l'individualité 118

3.2.1. L'individuation par l'identification des privations 118

3.2.2. L'individuation par la formation des valeurs 123

3.2.3. L'autonomie par la révision des préférences 126

3.3. La place de l'individu dans le choix social senien 128
3.3.1. Le conflit Pareto-liberté : libertés individuelles et préférence sociale 129
a) La critique libertarienne de la liberté comme préférence décisive 131
b) La critique libérale de la liberté comme préférence décisive 134
c) Répondre aux critiques : liberté en exercice, droits-buts et préférence décisive 135
3.3.2. La liberté individuelle comme réalisation sociale 136
a) Responsabilité individuelle, responsabilité sociale et rareté 136
b) Contrôle direct et liberté indirecte 139
3.3.3. Un accord basé sur les évaluations positionnelles ? A propos des jugements transpositionnels 144
Conclusion de la Ière Partie. 146
IIe partie : Nature et valeur des préférences 147
Chapitre 4 : Préférences, choix et évaluations 151
4.1. Préférence et décision : de la préférence révélée à la préférence fonctionnelle .. 151
4.1.1. Les préférences révélées 152
4.1.2. Amartya Sen, critique des préférences révélées 155
4.1.3. La préférence, cause du choix ? 158
a) La préférence comme cause du choix 158
b) La causalité comme rationalisation de l'action 160
c) Deux objections à la thèse de la relation causale 162
4.1.4. La préférence fonctionnelle 165
4.2. Préférences et évaluations : des jugements de valeur comparatifs 167
4.2.1. La préférence est un jugement 167
4.2.2. Un jugement comparatif 171
4.2.3. Jugements de fait comparatifs et jugements de valeur comparatifs 173
4.2.4. Jugements de valeur comparatifs et jugements de valeur non-comparatifs . 177
Conclusion du Chapitre 4 : nature de la préférence et approche des capabilités 180
Chapitre 5 : Que décrit la préférence ? Position de l'évaluateur et critères d'évaluation 183
5.1. La position de l'évaluateur et la rationalité de l'évaluation 185
5.1.1. « Voix » et choix social 186
a) Connaît-on ses propres préférences mieux que personne ? 186
On peut s'identifier aux prescriptions plutôt qu'à l'expérience de l'autre 188
Préférences et expérience 188
Les préférences des absents 189
b) Qualité d'agent et expression des préférences 190
5.1.2. L'observateur omniscient et impartial 192

a) L'observateur idéal 193
b) La position originelle 196
Théorie de la justice et théorie du choix rationnel 196
Position originelle, choix collectif et justice des institutions 199
5.1.3. Les évaluations positionnelles : de multiples possibilités 201
a) L'attribution des préférences aux positions 202
b) La position de l'évaluateur 204
c) Le test positionnel 205
d) La position de l'évaluateur : bref bilan 208
Implications pour le formalisme de la théorie du choix social 209
Implication pratiques 210
5.2. Le critère d'évaluation 211
5.2.1. L'hédonisme 211
a) Que décrit la préférence dans la conception « hédoniste » ? 213
b) Deux conceptions de l'hédonisme 215
L'hédonisme étroit 215
L'hédonisme du désir 218
c) Les critiques de la préférence basée sur les états mentaux 220
Une mesure arbitraire 220
Une interprétation réductrice 221
5.2.2. Le bien-être comme satisfaction des préférences 223
a) Satisfaction des préférences et connaissance des états sociaux 225
Le problème des croyances erronées 226
Le problème de l'occurrence différée de l'état social 227
b) Satisfaction des préférences et réalisation des intérêts 228
Quelles préférences individuelles ? 229
Quelles préférences collectives ? 231
c) Préférences et faisabilité 238
Le faisable et le possible 240
Le processus d'ajustement 242
Une solution : considérer les préférences irréalisables 244
5.2.3. Préférences rationnelles et objectivité du bien 247
a) Des rapports entre bien objectif et préférences 248
b) L'indépendance de la conception objective du bien à l'égard de la préférence 251

Du choix collectif au bien commun, la préférence comme révélation du bien objectif 252

Les limites de la description du bien objectif par la préférence 255

c) Le bien-être comme objet déterminé par une préférence rationnelle..... 255

Les conditions de rationalité de la préférence 256

Les arguments en faveur de la conception du bien-être comme objet de préférences rationnelles 266

IIIe partie : du choix social à l'évaluation des capabilités 273

Chapitre 6 : Les capabilités : conséquentialisme et liberté réelle 277

6.1. Conséquentialisme et mesure des résultats 278

6.1.1. Pour une conception conséquentialiste de l'évaluation sociale 278

6.1.2. Le conséquentialisme utilitariste : de l'agrégation des utilités au vote sur la base de préférences 281

a) L'agrégation utilitariste 282

b) Agrégation des préférences et égalité démocratique 284

6.2. De la mesure des résultats à l'évaluation des opportunités 285

6.2.1. Les opportunités de bien-être 289

6.2.2. Capabilités et fonctionnements 292

a) Capabilités ou opportunités de bien-être ? 292

b) Les fondements des capabilités : économie normative et néo-aristotélisme 295

c) Classifier les capabilités 298

Capabilités et acte de choix : fonctionnements affinés ou ensembles d'opportunités ? 302

Les finalités : bien-être ou liberté d'agent ? 306

Chapitre 7 : Vers une approche préférentialiste de l'évaluation des capabilités. 311

7.1. Le problème de l'évaluation et de la liste des capabilités 312

7.1.1. Le problème de l'évaluation et le problème de la liste : dépendances et différences 313

7.1.2. Le silence de Sen 315

7.1.3. Une liste de capabilités nécessaires à un fonctionnement digne de l'être humain : la solution néo-aristotélicienne de Martha Nussbaum 316

7.1.4. L'évaluation démocratique des capabilités aujourd'hui : les travaux de d'Ingrid Robeyns, de Sabina Alkire et d'Elizabeth Anderson 318

a) Ingrid Robeyns : expertise, brainstorming et contextualisation..... 318

b) Sabina Alkire : le raisonnement pratique de John Finnis et les dimensions du développement humain 320

c) Elizabeth Anderson : les capabilités suffisantes pour l'égalité démocratique 321

7.2. Une approche préférentialiste de l'évaluation des capabilités 324
7.2.1. Capabilités et théorie du choix social : deux possibilités 326
7.2.2. Les évaluations comparatives de capabilités 328
7.2.3. Les préférences pour les fonctionnements 330
a) Les préférences pour les fonctionnements : problèmes et définitions 330
b) La structure holiste des préférences pour les fonctionnements 331
c) La structure narrative des préférences pour les fonctionnements 334
d) Préférences globales pour les fonctionnements et degré d'adaptation 336
7.2.4. Des préférences pour les fonctionnements à l'évaluation des ensembles capabilité 339
a) Evaluation des capabilités et préférences pour les fonctionnements 339
b) Remarques sur la délibération comme comparaison d'opportunités 342
Conclusion 345
1. Un bref bilan 346
2. De nouvelles pistes de recherche 349
Bibliographie 353
Monographies 353
Ouvrages de référence 359
Articles de revues 359
Chapitres de livres 364
Articles de dictionnaire ou d'encyclopédie 366
Thèses 366
Documents de travail 366
Communications de colloques, cours, conférences 367
Ressources en ligne 367
Rapports officiels 367
Annexes 368
Annexe 1 : notation standard en logique mathématique 368
Annexe 2 : les relations de préférence et les ensembles de choix 369
Annexe 3 : le théorème d'impossibilité d'Arrow (démonstration de Sen) 371
Annexe 4 : l'impossibilité du libéral parétien 376

Introduction

1. Le pari de la rationalité collective

Le choix collectif est l'objet de cette étude. Le projet peut paraître ambitieux, tant le champ d'investigation semble large. La conception du choix collectif que nous étudions ici est celle qui correspond à l'agrégation des préférences. Le terme « agrégation » est issu du latin *grex,* qui signifie « troupeau », et par extension toute forme de regroupement, qu'il s'agisse de personnes ou d'éléments. En informatique, en logique, en statistiques ou en économie, ce terme décrit l'opération suivante : dans un premier temps, on sélectionne des éléments hétérogènes mais relevant néanmoins de la même catégorie, ce qui permet de les considérer comme commensurables ; dans un second temps, on associe à ces éléments un exemplaire représentatif de la même catégorie, ce qui permet de généraliser la collection des éléments[1]. Par exemple, on peut sélectionner différents individus de l'espèce « mouton » : les uns sont blancs, les autres sont noir, certains sont petits. À cette collection de moutons, on associe un mouton représentatif. Ce mouton est « représentatif » en vertu de règles non arbitraires. Dans la théorie du choix social, cette opération d'agrégation est appliquée aux évaluations, aux choix et aux préférences. Par exemple, on peut collecter les votes des citoyens français au moment des élections présidentielles de 2007, et associer à ces votes, en vertu d'une règle comme la règle de la majorité, une préférence représentative (en l'occurrence, celle pour le candidat François Hollande).

C'est l'agrégation des préférences qui nous intéresse ici. Au sens courant du terme, la préférence évoque à la fois l'inclinaison et la comparaison. A la croisée des notions de jugement, de désir et de choix, la préférence s'avère éminemment ambiguë, et un travail de clarification s'impose ici.

La notion de choix collectif décrit des réalités quotidiennes : tout choix effectué par plus d'un individu est un choix collectif. Pourquoi alors étudier une notion aussi commune à l'aide d'un concept aussi technique que celui d'agrégation des préférences ? Mais ce serait confondre familiarité et clarté. La notion de « choix collectif » est familière, mais elle n'est pas claire ; celle d'agrégation des préférences est peu familière, mais elle a la clarté de la précision.

[1]MONGIN, Philippe, "Une source méconnue de la théorie de l'agrégation des jugements", *Revue économique,* 2012 /4, vol. 63, pp 645-657.

Prenons un exemple : lorsque, dans un couple, vient le moment de déterminer qui fera la vaisselle, il semble que nous ayons ici un exemple de choix collectif. Mais suffit-il que l'événement ait lieu sans aucune manifestation d'opposition de la part d'un des conjoints pour parler de choix collectif ? Des forces invisibles n'ont-elles pas mû l'un ou l'autre contre son authentique volonté ? Nous savons que la division du travail au sein du foyer est encore tributaire de pressions sociales et de perceptions du rôle des genres qui ne rendent pas forcément justice à l'idée de « choix »[2]. La notion de choix collectif est donc ici de peu de secours. Qu'en est-il de l'agrégation des préférences ? Si la répartition de la vaisselle doit résulter de l' « agrégation des préférences », alors il nous faut dans un premier temps recenser les préférences de chacun des partenaires avant de déterminer une répartition qui soit représentative des souhaits du couple. L'agrégation des préférences correspond ainsi à une exigence de description précise des procédures de choix collectif. Or, la précision est aussi garante de l'absence d'arbitraire lorsque vient le moment de voter ou de répartir les bénéfices de la coopération sociale.

L'agrégation des préférences est l'objet d'étude de la théorie moderne du choix social. En ce sens, celle-ci est tributaire de à la fois de l'analyse mathématique des procédures de vote et de la maximisation du bien-être issue de l'utilitarisme philosophique.

L'approche consistant à traiter le choix collectif comme une agrégation de préférences présuppose un positionnement philosophique que nous qualifierons de « pari de la rationalité collective ». Il peut être décomposé en trois propositions.

---

[2] Le terme "choix" désigne d'ores et déjà plusieurs choses :

- Il peut décrire l'acte par lequel un sujet se détermine pour une option. Si des choix contrefactuels étaient effectivement possibles, on parlera de libre arbitre. Cet « acte » est d'ordre psychologique ou conscient. Si l'on fait nôtre une conception rationaliste du choix, cet acte implique un processus cognitif (la délibération). Cet aspect du choix n'est pertinent pour une étude du choix collectif que dans la mesure où la collectivité est conçue elle-même comme un « agent » ou un « sujet » doté d'une «psychologie » ou d'une « conscience » propre, et ce, pas seulement de manière métaphorique. Dans ce cas, la distinction entre choix individuel et choix collectif n'est pas pertinente, et la notion d'agrégation d'aucune utilité. Nous verrons dans notre étude que, si la distinction entre choix individuel et choix collectif n'est pas requise pour penser le problème du choix social, en revanche, l'idée que le choix collectif doive être le produit de l'agrégation de préférences hétérogènes l'est. Or, cette « agrégation » n'a de sens que dans la mesure où la « collectivité » elle-même n'est pas caractérisée par l'unité d'une volonté ou d'une capacité de choix propre.
- Le choix peut également être le « menu » du choix, l'ensemble d'options dans lequel l'agent sélectionne celle qui correspond à ce qui lui convient le mieux. En ce sens, l'existence du choix dépendrait de l'existence de ce menu autant que de la capacité de l'agent à déterminer ce choix.
- Le choix peut enfin décrire le résultat de tout ce processus, l'utilisation de ce terme permettant de le distinguer de l' « événement » naturel.

Notre étude s'intéresse au choix collectif ; mais ce terme lui-même ne fait sens que dans la mesure où l'on peut se référer à une sorte de « volonté collective ». L'enjeu de la théorie du choix social n'est donc pas tant l'action collective elle-même que la construction d'une évaluation de la société correspondant aux préférences et aux valeurs potentiellement divergentes des membres de cette collectivité. C'est pourquoi nous nous focaliserons sur la base informationnelle du choix collectif.

La première proposition stipule la prééminence de l'évaluation sur la décision. Autrement dit, il ne peut y avoir choix collectif que déterminé par une préférence collective : la décision collective dépend toujours de l'acte idéel d'évaluation. C'est parce que la décision dépend de l'évaluation que l'on parle de « rationalité » collective – même si, dans un premier temps, il n'est pas forcément question d'une rationalité *substantielle* - . Il convient toutefois ici de dissiper des malentendus : affirmer que le choix collectif est rationnel ne présuppose rien quant à son contenu. Notre hypothèse de rationalité est très faible : elle affirme la dépendance de la décision à l'égard de l'évaluation, et rien de plus ; elle n'affirme même pas que l'évaluation entraîne nécessairement la décision[3].

La deuxième proposition est la suivante : l'approche du choix collectif comme agrégation des préférences constitue une conception générale et abstraite du choix collectif. Il s'agit d'une idéalisation permettant de faire ressortir les propriétés du choix collectif. Autrement dit, cette approche met de côté, le temps de l'analyse, les questions relatives aux conditions concrètes d'exercice des choix collectifs, telles que la forme particulière des institutions ou les circonstances historiques et sociales. Mettre entre parenthèse n'est pas éliminer : nous savons que le questionnement sur le choix collectif émerge dans un contexte intellectuel, social et économique donné ; mais l'abstraction permet de dégager des traits communs à des formes diverses de choix collectifs. De façon assez surprenante, l'abstraction peut être porteuse de critique : en effet, en se concentrant sur les structures plutôt que sur les pratiques de choix collectif, il apparaît que certaines de ces pratiques n'en sont pas, et que d'autres qui échappent à l'institutionnalisation des choix collectifs, par le vote notamment, en sont de fait. L'abstraction est aussi un moteur de créativité en matière de propositions politiques et sociales, puisqu'elle permet d'envisager des solutions en termes de choix collectif là où le cadre institutionnel est défaillant ou inapproprié.

La troisième proposition affirme la valeur normative du choix collectif. Si l'agrégation des préférences est une définition exigeante du choix collectif, c'est qu'elle ne prétend pas décrire les pratiques que nous qualifions de « choix collectif » en raison de leur « air de famille », mais instituer une définition du choix collectif permettant de discriminer parmi ces pratiques celles qui sont d'authentiques choix collectifs et celles que l'on devrait caractériser comme « événements subis » ou comme « choix d'un seul » ou d'un groupe imposé à la collectivité. Si une telle définition discriminante est requise, c'est parce que la motivation d'une investigation du choix collectif conçu comme agrégation des préférences est la thèse selon

[3] Cette question sera abordée en détail dans le Chapitre 4.

laquelle il est bon que ce qui peut être l'objet de choix concernant la société dans son ensemble soit un choix collectif : il s'agit donc d'un positionnement normatif. Cela conduit à exclure du domaine des choix collectifs certaines pratiques voisines telles que les négociations : le choix collectif entendu comme agrégation des préférences aspire à être représentatif des intérêts et opinions des membres de la société plutôt que des effets des rapports de force et des marchandages.

2. Un peu d'histoire

Un regard sur le passé éclairera les motivations à l'origine de l'étude de l'agrégation des préférences. Celle-ci peut se prévaloir d'une double filiation : d'une part, l'étude des procédures de vote, qui remonte au Moyen Âge au moins ; d'autre part, la tradition utilitariste d'évaluation sociale.

L'étude des procédures de vote : de Pline le Jeune à Condorcet

Une première apparition des réflexions sur le vote se trouve dans les *Lettres de Pline le Jeune :* Pline rapporte le sort de suspects accusés d'avoir tué un consul romain. Le Sénat romain devant décider du sort des suspects, il peut soit les condamner à mort, soit les bannir, soit les acquitter. De peur que les partisans de l'acquittement ne gagnent, ceux qui étaient favorables à la peine de mort énoncèrent une préférence pour le bannissement plutôt que pour la condamnation à mort[4]. Ce cas de « vote utile » motive une première réflexion sur l'influence des procédures de vote sur le choix collectif final, réflexion qui constitue l'objet même de la théorie du choix social. En effet, le problème du choix social apparaît lorsque les évaluations des membres de la société ne sont plus homogènes. L'objet du choix social consiste alors à déterminer des procédures permettant de dériver une évaluation unique de cette pluralité de préférences et de valeurs. L'intérêt de cette étude devient manifeste lorsque, à l'instar de Pline, nous nous rendons compte que le choix même de la procédure n'est pas sans conséquence sur le résultat final.

Bien que l'on tende à associer la démocratie à l'Antiquité et à la civilisation grecque, la réflexion sur les procédures de vote et de choix collectif ne commence véritablement qu'au Moyen Âge. Comme le soulignent McLean et London, les élections telles que celle du Pape, du Saint Empereur Romain Germanique ou encore du Doge de Venise reposaient sur des

[4] Cf. GAERTNER, Wulf, *A Primer in Social Choice Theory*, Oxford : Oxford University Press, 2006, p 69.

schèmes électoraux complexes impliquant un questionnement logique[5]. Cette remarque historique est une pierre de plus apportée à l'argument en faveur de la deuxième proposition du « pari de la rationalité collective » : nous devons nous méfier et ne pas associer précipitamment les institutions dites démocratiques à la pratique des choix collectifs. Des sociétés n'ayant jamais revendiqué le statut de démocratie peuvent être profondément structurées par la prise de décision en groupe, tandis que des « démocraties » seront davantage caractérisées par certaines structures institutionnelles (la garantie de droits individuels ou la présence d'une constitution, par exemple) que par la fréquence des délibérations et des décisions collectives.

McLean et London recensent trois occurrences médiévales de réflexion sur les procédures de choix collectif. La première est une méthode décrite dans le *Libre d'Evast e Blanquerna* écrit entre 1282 et 1287 par le philosophe, poète et théologien catalan Raymond Lulle. Le contexte est celui de l'élection d'une abbesse : Natana, une des soeurs, propose une méthode électorale par laquelle « la vérité peut être trouvée », et donc la meilleure et la plus convenable des abbesses. La méthode en question consiste en comparaisons par paires, la gagnante étant celle qui a reçu le plus de votes une fois agrégées ces comparaisons. La deuxième occurrence se trouve dans un autre texte de Lulle, explicitement intitulé *De Arte Electionis* (1299). Ce texte concerne les élections au sein de l'Eglise et préconise une méthode apparentée au « gagnant de Condorcet » : le candidat retenu doit être celui qui a été classé au-dessus de chacun des autres dans les comparaisons par paires. La troisième occurrence est un texte du penseur Nicolas de Cues. Au XVe siècle, celui-ci s'intéresse aux procédures d'élection de l'Empereur germanique, visant, selon ses propres termes, à éviter les pratiques absurdes ou malhonnêtes. La méthode préconisée est celle de Borda, avec un scrutin secret. Soulignons que Nicolas de Cues préconise que les électeurs agissent en liberté, sans peur, et réfléchissent en toute conscience, « au nom de Dieu », afin de déterminer le meilleur candidat. Ces trois occurrences ne résument sans doute pas toute la pensée médiévale à ce sujet. McLean et London rappellent notamment que Lulle était arabisant et avait de fréquents contacts avec l'Espagne islamique : il se pourrait que ses idées sur les mathématiques du vote aient des origines arabes.

[5]Cf. McLEAN, Ian and LONDON, J., "The Borda and Condorcet Principles. Three Medieval Applications", *Social Choice and Welfare*, 7, 1990, pp 99-108.

Mention est faite également d'une réflexion sur les procédures de vote dans l'ouvrage de Pufendorf, *De Jure Naturae et Gentium* (1672). Mais ce sont les travaux de Borda et de Condorcet qui inaugurent véritablement l'étude mathématique des procédures de vote.

Le chevalier Jean-Charles de Borda donna en 1770 une conférence à l'Académie des Sciences, qui fut publiée en 1784, et dans laquelle il proposait une étude comparative des différents systèmes de vote. Borda avait remarqué que la méthode consistant à désigner comme vainqueur de l'élection le candidat ayant été classé au-dessus de chaque autre candidat lors de comparaisons par paire pouvait mener à un choix non représentatif de la majorité, dans les cas où il y avait plus de deux candidats (cette méthode, dite du gagnant de Condorcet, est également celle décrite par Lulle dans *Blanquerna*). Il proposa donc la procédure suivante[6] : chaque électeur classe les candidats par ordre de préférence, et assigne à chacun une valeur égale à son rang dans le classement. Comme le rappelle Arrow, par cette méthode :

> *« Borda souleva ainsi la plupart des questions qui ont occupé les analyses ultérieures : (1) baser les choix sociaux sur les classements entiers des candidats disponibles par tous les individus, et pas seulement sur le premier du classement de choix; (2) la mesurabilité des utilités individuelles ; et (3) la comparabilité interpersonnelle des préférences (Borda fit de la comparabilité interpersonnelle un jugement éthique d'égalité, pas un jugement empirique). »*[7]

Borda proposait en effet de prendre en compte dans le calcul électoral des informations sur la position occupée par chaque candidat dans le classement de préférences des électeurs. Cette approche ordinale de la base informationnelle du choix collectif était complétée par une mesure cardinale de la « valeur » de chaque candidat, que, de manière peut-être arbitraire, Borda stipulait équivalente à son rang dans le classement de l'électeur. La comparabilité relève d'un jugement « éthique » dans la mesure où c'est le principe d'égalité des votants, et non pas le souci de refléter le degré d'attente espéré pour chaque candidat par chaque électeur, qui détermine l'attribution de la valeur selon le rang occupé dans le classement.

Mais le véritable point de départ de la théorie du choix social est le fameux paradoxe de Condorcet, qui apparaît dans son ouvrage de 1785, l'*Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Condorcet souhaitait étudier les

---

[6] Cf. GAERTNER, Wulf, *A Primer in Social Choice Theory*, Oxford : Oxford University Press, 2006, p 92-98.

[7] ARROW, Kenneth, « Formal Theories of Social Welfare », in WIENER, Philip P. (ed), *Dictionary of the History of Idea*s, Charles Scribner's Sons, 1973, vol. 4, pp 276-284. Rééd. Dans ARROW, Kenneth, *Collected Papers*, Cambridge, MA: The Belknap Press of Harvard University Press, 1983.

procédures de décision collective afin de déterminer le système de vote ayant le plus de probabilité de générer une décision selon la raison.

> *« Il est donc de l'intérêt de ceux qui disposent de la force publique, de n'employer cette force que pour soutenir des décisions conformes à la vérité, et de donner aux Représentans, qu'ils ont chargés de prononcer pour eux, des règles qui répondent à la bonté de leurs décisions. »*[8]

Comme Raymond Lulle ou Nicolas de Cues, la préoccupation de Condorcet est d'abord d'ordre épistémique : une bonne procédure de vote conduit à une décision conforme à la « raison » et représentative des intérêts ou des opinions des électeurs. Elle échappe à la manipulation comme à la domination. Le refus de l'arbitraire relève également d'un souci normatif : la procédure de choix collectif est en effet appelée à légitimer son résultat, et le souci de justice et d'équité réapparaît régulièrement dans ces textes fondateurs de la théorie du choix social.

Le paradoxe de Condorcet est le résultat dont la portée pour la future théorie du choix social fut la plus importante[9]. Condorcet l'exemplifie par le cas d'une élection à trois candidats, désignés par les lettres A, B et C. Trois électeurs sont appelés à classer ces candidats. Le problème est alors que le cas de figure suivant est possible (Condorcet remarque que, parmi les huit combinaisons possibles, deux d'entre elles donnent un résultat contradictoire de ce type) :

- Le premier électeur préfère A à B et B à C ;
- Le second électeur préfère B à C et C à A ;
- Le troisième électeur préfère C à A et A à B.

Nous avons ici affaire à un cycle, puisque l'agrégation de ces classements donne le résultat suivant : le groupe préfère A à B, B à C et C à A. Ce cycle nous empêche de déterminer « le premier du classement » et donc le choix collectif qui résulterait de ce vote. Lors de l'élection présidentielle du 22 avril 2007, les sondages montrèrent que le candidat Nicolas Sarkozy était préféré à Ségolène Royal, Ségolène Royal à François Bayrou et François Bayrou à Nicolas Sarkozy. Sans le scrutin à deux tours et la restriction de l'information au premier choix, il

---

[8] CONDORCET, Marie-Jean-Antoine de Caritat, marquis de, *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Paris, 1785, Discours Préliminaire.Rééd.Paris : Fayard, 1986.

[9] Cf. GAERTNER, Wulf, *A Primer in Social Choice Theory*, Oxford : Oxford University Press, 2006, p 100-101

aurait été difficile de désigner le nouveau Président. Le paradoxe de Condorcet n'est pas un problème trivial[10].

Une contribution de C. L. Dodgson, mieux connu sous le nom de Lewis Carroll, rend également compte des majorités cycliques[11]. Cette contribution n'était pas destinée à être publiée ; c'était un pamphlet sur les élections à Oxford. Le logicien y propose une règle de vote consistant à choisir le candidat qui aurait eu la majorité sur tous les autres si les classements de préférences des électeurs étaient intervertis[12]. Dodgson suggère également la possibilité de ne pas faire d'élection. Il estime de ce fait que ce résultat d'impossibilité équivaut à une préférence pour le *statu quo.*

[10] Cf. FANSTEN, Michel, « L'échec électoral est une science exacte », *Mathématiques et sciences humaines*, 185, Printemps 2009

[11] DODGSON, C. L., *A Method on taking votes on more than two issues,* in P. Black, 1958.

[12] Cf. GAERTNER, Wulf, *A Primer in Social Choice Theory*, Oxford : Oxford University Press, 2006, p 100-101.

L'agrégation utilitariste

La seconde source du choix collectif conçu comme agrégation des préférences est le calcul utilitariste.

Si Jeremy Bentham (1748-1832) a le premier systématisé l'utilitarisme, l'idée selon laquelle l'évaluation des actions et des comportements devrait être fondée sur le bonheur ou l'utilité se retrouve chez des précurseurs tels que Shaftesbury, Hutcheson ou Hume[13].

Le siècle de Bentham est marqué par les découvertes de Newton[14] : une nouvelle conception de la science émerge, caractérisée par la connaissance de lois naturelles tirée de l'observation et de l'expérience, par une approche qui mathématise la nature, et par un principe d'association : le principe de l'attraction universelle. Le rêve des philosophes admirateurs de Newton est de transposer cette « science » dans le domaine de la psychologie, de la morale, de la politique et du droit[15].

Au principe d'attraction universelle qui régit les phénomènes naturels correspond pour Hume le principe d'utilité qui régit l'action[16]. D'après ce principe, une action *bonne* est une action conforme à l'intérêt de l'individu. Cet effet est observable, le principe d'utilité permet donc la constitution d'une science de la morale objective. Le principe d'utilité a valeur d'axiome. Plus une action produit de plaisir, de satisfaction ou de bonheur (tous ces termes étant associés à une expérience positive), plus son utilité est grande.

Lorsque nous passons à l'échelle collective, le principe d'utilité devient synonyme de principe du plus grand bonheur du plus grand nombre (énoncé en ces termes par Helvétius[17]). On passe ici d'un sens descriptif du principe, énoncé par Hume comme une loi de la nature humaine, à un sens normatif : la tâche de la politique est de réaliser le plus grand bonheur du plus grand nombre.

---

[13] Cf. HUME, David, *A Treatise of Human Nature*, 1739. Rééd par L.A. Selby-Bigge, Oxford : Oxford University Press, 2000 ; HUTCHESON, Francis, *An Inquiry into the Original of Our Ideas of Beauty and Virtue,* London , 1725 ; SHAFTESBURY, Anthony Ashley Cooper, 3rd Earl of Shaftesbury, *Characteristics of Men, Manners, Opinions, Times*, edited by Lawrence E. Klein, Cambridge: Cambridge University Press, 1999.
Pour l'histoire de l'utilitarisme, nous nous sommes référés à : AUDARD, Catherine, *Anthologie historique et critique de l'utilitarisme,* Paris, PUF, 1999, et DRIVER, Julia, "The History of Utilitarianism", *The Stanford Encyclopedia of Philosophy* ,Edward N. Zalta (ed.), été 2009 ;<http://plato.stanford.edu/archives/sum2009/entries/utilitarianism-history/>.

[14] 1688

[15] Cette ambition apparaît clairement au début du *Traité de la Nature humaine* de Hume.

[16] HUME, *op. cit.*, III, I, II.

[17] HELVETIUS, *De l'Esprit,* Paris [Liège], Durand [Bassompierre], 1759.

Cette tâche requiert le « calcul utilitariste » : les utilités individuelles sont recensées. Par « utilité », Bentham et ses successeurs entendent la magnitude mesurable de satisfaction résultant d'une action ou d'une configuration sociale donnée. Cette recension est suivie d'une sommation ; pour chaque action ou chaque politique possible, les utilités individuelles sont agrégées afin de calculer la satisfaction totale qui en découle. La finalité de l'opération étant le plus grand bonheur du plus grand nombre, l'agrégation sera effectuée en vue de la maximisation du bien-être[18]. La tradition utilitariste reprendra, affinera ou réinterprétera l'approche agrégative du bien-être, les contributions de John Stuart Mill et de Henry Sigdwick étant les plus riches et les plus étudiées[19].

[18] La maximisation n'est pas nécessairement impliquée par l'opération d'agrégation. On peut – logiquement - collecter les utilités individuelles en vue de définir la configuration dans laquelle la somme totale de satisfaction est la plus *basse* (cela supposerait l'hypothèse selon laquelle la norme d'action doit être la douleur, l'insatisfaction ou le malheur) ; on peut également définir un seuil de satisfaction, sans chercher à aller au-delà. Cf. BROOME, John, *Weighing Goods*, Oxford : Blackwell, 1991, p 7.

[19] L'agrégation utilitarisme telle que la conçoit Bentham fera l'objet de la section 1.1. de ce travail.

La rupture avec la philosophie : naissance de l'économie du bien-être et de la théorie du choix social

A la fin du XIXe siècle, l'économiste irlandais Edgeworth applique le critère de sommation des utilités à l'évaluation des politiques fiscales[20]. Il s'inscrit ainsi dans la « Révolution marginaliste », au cours de laquelle la théorie de la valeur comme quantité de travail fourni est remplacée par une théorie de la valeur comme utilité, au sens que les utilitaristes donnent à ce terme. Les implications de cette opération d'agrégation sont égalitaristes : puisqu'on observe que les pauvres tirent plus de satisfaction de la possession d'argent que les riches, il convient de redistribuer les richesses de façon égalitaire[21].

Pour des raisons épistémologiques, Pareto puis Lionel Robbins et Abram Bergson rejetèrent la thèse de la possibilité d'assigner une mesure cardinale à l'utilité individuelle[22], ainsi que celle de la possibilité d'effectuer des comparaisons interpersonnelles d'utilités. L'économie du bien-être naissante culmine avec l'article d'Abram Bergson portant sur les « fonctions de bien-être social » [23] [24] : la fonction de bien-être social fait dépendre les jugements de valeur collectifs de la configuration de l'utilité des individus donnés[25]. La signification du terme « utilité » est devenue ici tout à fait abstraite, et ne se rattache plus nécessairement à la description d'une expérience de satisfaction : en fait, cette question est considérée par les auteurs comme ne relevant pas de l'économie.

---

[20] Cf. EDGEWORTH, F. Y., *Mathematical Psychics,* London , et « The Pure Theory of Taxation », in *Papers Relating to Political Economy,* London : 1925.

[21] Cela correspond à la « loi de l'utilité marginale décroissante », généralisation d'une observation selon laquelle la satisfaction tirée de la possession d'un bien décroît au fur et à mesure que l'on possède davantage de quantité de ce même bien. Si l'on m'offre un pull jaune, cela me fera plaisir ; si l'on m'en offre deux, j'aurais davantage de plaisir, mais pas deux fois plus. Si l'on m'offre cinquante pulls jaunes, je tenterai probablement d'en revendre une partie sur un site Web de vente aux enchères.

[22] Les sections 1.1.3. et 1.2. de ce travail en exposent les raisons

[23] Celles-ci font l'objet des sections 2.2.2. et 5.1.2. de ce travail

[24] Cf. BERGSON, Abram, "A Reformulation of certain aspects of Welfare Economics", *Quarterly Journal of Economics*, 52, pp 310-334, 1938 et SAMUELSON, P.A., *Foundations of Economic Analysis,* Cambridge, MA: Harvard University Press, 1947.

[25] En termes techniques, Gaertner définit ainsi le concept de fonction de bien-être social : il s'agit d' « une application à valeur réelle *W,* dont la valeur dépend de tous les éléments qui affectent le bien-être (*welfare*) d'une communauté durant une période donnée, c'est-à-dire la quantité des différentes marchandises consommées, la quantité des différentes sortes de travail effectuées, la quantité de facteurs autres que le travail dans chacune des unités de production, etc. » (GAERTNER, Wulf, *A Primer in Social Choice Theory*, Oxford : Oxford University Press, 2006, p 5).

Les travaux de Black orientèrent l'attention de la science économique vers les élections[26]. Mais l'idée d'associer le paradoxe de Condorcet à la fonction de bien-être social fut celle d'Arrow, et donna naissance au fameux théorème d'impossibilité à l'origine de la théorie du choix social[27]. Dans la mesure où ce théorème occupe une place importante dans ce travail, nous en donnons ici une présentation informelle (une démonstration formelle se trouve dans l'Annexe 3).

[26]BLACK, Duncan, "On The Rationale of Group Decision Making", *The Journal of Political Economy,* 56, 1948, pp 23-34.

[27]ARROW, Kenneth, *Social Choice and Individual Values*, New Haven, Yale university Press, 1951 (2ème édition révisée, 1961).Traduction française par l'Association de Traduction Economique de l'Université de Montpellier, Paris, Calmann-Levy, 1974

Le théorème d'impossibilité du choix social (1951)

Le point de départ de la réflexion d'Arrow est qu'un choix qui concerne une collectivité ne peut être dit « collectif » que s'il dépend des préférences de tous les membres de cette collectivité. Le choix collectif doit donc être fonction de ces préférences individuelles[28]. Il a pour objet un ensemble d'états sociaux alternatifs, décrits de manière aussi pertinente et complète que possible[29] Aussi bien le choix collectif que les préférences individuelles sont représentés par Arrow sous la forme de relations d'ordre[30]. Ces relations d'ordre sont des relations de préférence faible, de préférence stricte ou d'indifférence. Une relation de préférence stricte stipule qu'un état social est strictement préféré à un autre, c'est-à-dire qu'il ne lui est pas indifférent. Une relation d'indifférence signifie qu'un état social est indifférent à un autre. Une relation de préférence faible signifie que l'état social faiblement préféré est soit préféré, soit indifférent à l'autre.

Arrow spécifie dans un premier temps les propriétés des relations d'ordre qui représentent tant les préférences individuelles que la préférence collective :

- REFLEXIVITE : Une option ne peut être strictement préférée à elle-même.

- TRANSITIVITE : Si une option *x* est préférée à une option *y*, et *y* à l'option *z*, alors *x* doit être préférée à *z*.

- COMPLETUDE : Pour chaque alternative, il doit être possible d'établir une relation de préférence ou d'indifférence entre les deux membres de l'alternative (l'ordre de préférence ne peut donc être incomplet).

Arrow essaie donc de déterminer une « règle de choix social », c'est-à-dire une règle permettant de dériver une préférence collective des préférences individuelles (un exemple bien connu de cette règle est celui de la règle de la majorité). Une fois cette règle déterminée, il suffirait de connaître le contenu des préférences individuelles (le « profil » de ces préférences) pour déterminer la préférence collective. La question à laquelle Arrow essaie de répondre est donc la suivante : comment, selon quelle règle, la préférence collective dépend-elle des préférences individuelles, quelles qu'elles soient ?

Ce ne peut être n'importe quelle règle. Arrow formule alors les conditions que devrait respecter cette règle pour que l'on puisse, selon lui, parler de choix collectif lorsque cette

[28] La signification des termes « individuel » et « collectif » est étudiée dans une perspective historique, dans la Ière partie de ce travail.

[29] La question de la description des états sociaux et de son impact sur les préférences est traitée en partie dans le Chapitre 7 de ce travail.

[30] Cf. Annexe 2. L'interprétation de ces « préférences » est l'objet de la IIe partie de ce travail.

règle est respectée. Ces conditions sont dites « raisonnables » dans la mesure où elles se justifient par la définition même de ce qu'est un choix collectif (par exemple, elles excluent la dictature). Elles revêtent le statut d'axiome dans la formalisation de la règle de choix social. Dans la version de Sen, que l'on rencontre le souvent, ces conditions sont au nombre de quatre :

Arrow essaie donc de déterminer une « règle de choix social », c'est-à-dire une règle permettant de dériver une préférence collective des préférences individuelles (un exemple bien connu de cette règle est celui de la règle de la majorité). Une fois cette règle déterminée, il suffirait de connaître le contenu des préférences individuelles (le « profil » de ces préférences) pour déterminer la préférence collective. La question à laquelle Arrow essaie de répondre est donc la suivante : comment, selon quelle règle, la préférence collective dépend-elle des préférences individuelles, quelles qu'elles soient.

Arrow essaie donc de déterminer une « règle de choix social », c'est-à-dire une règle permettant de dériver une préférence collective des préférences individuelles (un exemple bien connu de cette règle est celui de la règle de la majorité). Une fois cette règle déterminée, il suffirait de connaître le contenu des préférences individuelles (le « profil » de ces préférences) pour déterminer la préférence collective. La question à laquelle Arrow essaie de répondre est donc la suivante : comment, selon quelle règle, la préférence collective dépend-elle des préférences individuelles, quelles qu'elles soient.

- CONDITION U (Domaine Non-Restreint) : la règle de choix social doit être valable pour n'importe quel ensemble logiquement possible de préférences individuelles. On ne peut exclure en amont des préférences individuelles de la procédure de choix collectif.
- CONDITION P (Pareto-Faible) : si chaque membre de la société préfère *x* à *y*, alors la société doit préférer *x* à *y*. C'est une forme d'unanimité.
- CONDITION I (Indépendance des Alternatives Non-Pertinentes) : la préférence collective se prononçant sur un ensemble donné d'options (aussi appelé « environnement de choix) *ne doit dépendre que* des préférences individuelles sur ce même ensemble. Cela signifie qu'elle ne peut dépendre ni d'informations autres que les préférences comprises comme des classements par ordre de préférence justement, ni de préférences se prononçant sur des options ne figurant pas dans l'ensemble proposé au choix. Cette condition exclut donc aussi bien des informations comme

l'intensité de la satisfaction ressentie que des préférences pour des candidats morts, par exemple.

- CONDITION D (Non-Dictature) : Il ne devrait pas y avoir de « dictateur », c'est-à-dire d'individu tel que, si il préfère une option à une autre, alors, quelles que soient les préférences des autres membres, la collectivité devra aussi préférer cette option.

Le théorème d'Arrow démontre l'impossibilité du choix social pour toutes les règles de vote satisfaisant ses conditions raisonnables, et pas seulement la règle de la majorité. Aucune règle, qu'il s'agisse de la règle de la majorité ou de celle de l'unanimité, par exemple, ne peut se conformer aux conditions qui garantissent que le choix collectif est bel et bien représentatif des évaluations des membres de la société[31].

La théorie du choix social devient à la suite d'Arrow un champ d'investigation à part, avec de nombreux résultats. Parmi ceux qui ont eu le plus de retentissement, figurent, outre le théorème d'Arrow et les travaux de Black susmentionnés, le théorème d'impossibilité du parétien libéral de Sen, sur lequel nous reviendrons et le théorème de Gibbard-Satterthwaite[32]. Une première recension de l'état de la discipline fut effectuée par Amartya Sen, dans l'ouvrage devenu classique *Collective Choice and Social Welfare*. Ce texte déploie de nombreux aspects du choix collectif conçu comme agrégation des préférences. Mais il se distingue surtout par le souci de ne jamais perdre de vue les aspects philosophiques du choix collectif, et a d'ailleurs été entièrement révisé par John Rawls :

> *« La théorie du choix collectif appartient à de nombreuses disciplines. L'économie en est une mais pas la seule. Bien que ce livre appartienne à une série de « textes en mathématique économique », aucune tentative de confiner cette étude aux seuls problèmes économiques n'a été faite. En fait, l'approche de ce livre repose sur la croyance que le problème ne peut être discuté de manière satisfaisante à l'intérieur des frontières de l'économie. Bien que le choix collectif soit un aspect crucial de l'économie (notamment de l'économie du bien-être, de la théorie de la planification et de l'économie publique), le sujet est étroitement lié aux sciences politiques, en particulier à la théorie de l'Etat et à la*

[31] Le récit de la découverte d'Arrow par lui-même est rapporté dans : ARROW, Kenneth, *Collected Papers*, Cambridge, MA: The Belknap Press of Harvard University Press, 1983, pp 1-4.

[32] Pour une présentation des résultats les plus importants de la théorie du choix social, on se référera à : SALLES, Maurice, articles « agrégation des préférences » et « bien-être et choix social », in BEITONE, Alain, CAZORLA, Antoine, DOLLO, Christine, DRAI, Anne-Mary, *Dictionnaire des Sciences Economiques* , Paris : Armand Colin, 2006 ; SALLES, Maurice, « La théorie du choix social : de l'importance des mathématiques », Center for Research in Economics and Management (CREM), University of Rennes 1, University of Caen and CNRS, Economics Working Paper Archive (University of Rennes 1 & University of Caen) no 200617 ; SEN, Amartya, *Collective Choice and Social Welfare,* San Francisco, Holden day, 1970; FLEURBAEY, Marc, *Théories économiques de la justice*, Paris, Economica, 1996 ; GAERTNER, Wulf, *A Primer in Social Choice Theory*, Oxford : Oxford University Press, 2006.

théorie des procédures de décision. Il a également des aspects philosophiques importants, liés à l'éthique et notamment aux théories de la justice. »[33]

Certains des aspects philosophiques dont parle Sen font l'objet de ce travail de recherche.

Notre analyse s'intéresse plus spécifiquement à la base d'information de la théorie du choix social et aux théorèmes d'Arrow et de Sen (les démonstrations formelles sont reproduites en annexe). Nous avons écarté le paradoxe de Condorcet, ainsi que les développements sur la manipulation des règles de décision, initiés par le théorème de Gibbard-Sattherthwaite. Ce théorème prend en compte la possibilité pour les individus d'avoir des comportements stratégiques et de ne pas énoncer leurs véritables préférences afin d'influer le résultat du vote en leur sens. Gibbard et Sattherthwaite ont démontré que, lorsque les individus ont la connaissance des préférences des autres votants, alors aucune fonction de choix social permettant de déterminer un choix social à partir d'un ensemble de préférences individuelles ne peut satisfaire simultanément une condition de non-manipulation et une condition d'absence de dictature. En d'autres termes, il faut soit que la règle de vote puisse être manipulable, soit qu'il y ait un « dictateur », au sens technique du terme. Ce théorème, qui associe à la théorie du choix social la théorie des jeux, est d'un intérêt certain pour la philosophie politique et les sciences politiques, ainsi que pour l'économie de l'information, le droit et la théorie des contrats. Néanmoins, notre problématique relevant davantage de l'évaluation sociale et de la question du bien-être collectif, nous ne l'évoquerons pas davantage[34].

[33]SEN, Amartya, *Collective Choice and Social Welfare*, San Francisco, Holden day, 1970, Préface.Traduction libre.

[34] Le théorème est exposé dans les articles suivants: GIBBARD, Allan, "Manipulation of voting schemes: a general result", *Econometrica* 41 (4), 1973, pp 587–601 ; SATTERTHWAITE, Mark, "Strategy-proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions", *Journal of Economic Theory* 10, 1975, pp 187–217. Sur l'exposé initial du problème, cf. FARQUHARSON, R., *Theory of voting*, New Haven : Yale University Press, 1969, pp 20-23. Pour une approche comparative des systèmes électoraux dans cette perspective, cf. COX, GW, *Making Votes Count : Strategic Interaction in the World's Electoral System*, New York: Cambridge University Press, 1997. Sur les enjeux éthiques, on peut se referer à BRENNAN, Jason, *The Ethics of Voting*, Princeton University Press, 2011.

3. Philosophie et théorie du choix social

Dans une notice présentant le choix social et l'économie normative, Marc Fleurbaey et Philippe Mongin mettent en évidence trois aspects philosophiques de cette discipline, aspects jusqu'alors négligés :

> *« Ni les distinctions internes à l'économie normative contemporaine, ni les arguments qui la rassemblent polémiquement contre les théories précédentes du choix social ou du bien-être collectif, ne doivent faire oublier que toutes [ces conceptions] participent de certaines postulations philosophiques communes, d'autant plus massives qu'elles sont rarement explicitées. En premier lieu, toutes s'accordent à voir dans l'individu sinon la source unique des jugements de valeurs portés au nom de la collectivité, du moins leur source primordiale et le seul objet dont ils traitent en réalité. Dans le calcul parétien du bien-être, en théorie du choix social, aussi bien que dans les théories de la négociation ou de la position originaire, les individus sont à l'origine des jugements de valeurs substantiels, qu'il s'agit alors seulement d'« agréger », en s'aidant d'autres jugements de valeur, subordonnés parce que purement formels (par exemple, le principe de Pareto, l'axiome arrovien d'indépendance). S'il est vrai que le schéma se brouille dans les travaux les plus récents d'économie normative, il demeure que les jugements de valeur substantiels ne concernent finalement que l'existence individuelle, dans telle ou telle de ses modalités -- consommation des ressources, « réalisations », etc. En second lieu, [ces constructions] entretiennent une relation étroite, mais mal définie et insuffisamment problématisée, avec la philosophie morale traditionnelle. Les normes explicitées par la théorie arrovienne du choix social concernent aussi bien les fondements de la vie politique, que l'appréciation morale de la répartition des biens. Lorsque le projet d'une théorie économique de l'« équité » s'autonomise, l'interprétation éthique devient sans doute prédominante; mais on peut avoir alors l'impression que la moralité, aux yeux de certains auteurs qui ne sont pas des moindres, se réduirait aux seuls aspects de justice distributive. Ainsi, [ces conceptions] tantôt excèdent, tantôt limitent, le champ d'investigation ordinaire de l'éthique. Enfin, parce qu'ils emploient les méthodes de l'économiste théoricien, et héritent de concepts techniques qu'ils n'analysent pas toujours complètement, les auteurs contemporains en répercutent certaines ambiguïtés caractéristiques. En particulier, même si la critique du « welfarisme » a permis d'en mieux délimiter le rôle, la notion de préférence apparaît à la fois omniprésente et mystérieuse: elle intervient tantôt dans la prise de décision et le choix, tantôt dans ces actes purement idéels que sont l'évaluation et le classement; elle décrit tantôt le bien de l'individu de façon quasi objective, tantôt son bien-être ou son bonheur, tantôt, seulement, ses goûts. Les économistes du 20ème siècle ont sans doute affermi le*

fondement de leurs théories en subordonnant l'utilité à la préférence; mais ce remaniement n'a pas dissipé entièrement l'obscurité de leurs concepts premiers. »[35]

Cette citation un peu longue constitue un véritable programme de recherche. Le premier axe de questionnements est celui de l'individualisme présumé de la théorie du choix social et de l'économie normative : l'agrégation des préférences est-elle nécessairement l'agrégation de préférences *individuelles* ? Et, si oui, en quel sens l'individu est-il l'auteur ou la cause de ces préférences ? Le deuxième axe, déjà mis en évidence par Sen, concerne l'objet de ces théories : se réduit-il aux normes de répartition des ressources ? Le rappel historique met en évidence la filiation utilitariste, et un rapprochement de cette branche de l'économie avec les débats contemporains en philosophie morale ne peut qu'être fécond. Le troisième axe porte sur la préférence. Que décrit-elle ? Quel est son rapport avec les critères d'évaluation traditionnellement utilisés en éthique ? L'investigation menée sur les préférences ne peut faire l'économie d'une réflexion sur l'objectivité des jugements moraux dans le contexte du choix collectif.

Ce « programme de recherche » a inspiré ce travail de thèse, même si le deuxième axe a été mis de côté, dans la mesure où l'étude de tous les champs d'application possibles de l'agrégation des préférences serait un objectif bien trop ambitieux. En revanche, nous avons tenté de rétablir la connexion entre les débats contemporains en philosophie morale et les différents avatars de l'utilitarisme, d'une part, et le cadre conceptuel commun à l'économie normative et à la théorie du choix social, d'autre part.

Si nous avons choisi d'étudier le choix collectif comme agrégation des préférences du point de vue de l'analyse des présuppositions philosophiques du cadre conceptuel commun à la théorie du choix social et à l'économie normative, il ne s'agit sans doute pas de la seule manière d'entreprendre une réflexion de philosophie à partir de celles-ci. Nous avons exclu trois pistes, que nous ne mentionnerons qu'en introduction, où lorsqu'un de leurs résultats sert notre propre argumentation.

Une première piste fait actuellement l'objet de travaux de recherche très actifs : il s'agit de la théorie de l'agrégation des jugements[36]. Cette théorie propose un traitement logique du problème de l'agrégation des jugements, qui constitue une généralisation de l'agrégation des

[35] FLEURBAEY, Marc, et MONGIN, Philippe, "Théorie du choix social et économie normative", in *Dictionnaire de philosophie morale*, sous la dir. de M. Canto-Sperber, Paris, P. U.F., 1996, p. 243-251.

[36] Pour une vue d'ensemble, cf. MONGIN, Philippe et DIETRICH, Franz, « Un bilan interprétatif de la théorie de l'agrégation logique », *Revue d'économie politique*, 120 (6), pp 929-972.

préférences. Le traitement logique permet à la fois de généraliser le cadre conceptuel et de clarifier les problèmes relatifs à la production de jugements collectifs.

Une deuxième manière de lier philosophie et choix collectif a été proposée par Jon Elster lors de ses cours du Collège de France en 2008 et 2009. Tandis que les théoriciens du choix social partent d'un concept abstrait d'agrégation des préférences pour ensuite vérifier dans quelle mesure les choix collectifs concrets correspondent bel et bien à la définition de ce concept, Elster adopte le cheminement inverse : il commence par décrire des formes de décisions collectives, en s'appuyant sur une définition de travail relativement intuitive. La description se focalise sur les mécanismes conduisant aux décisions en question, et vise à dégager des catégories telles que l'argumentation ou la négociation. Le parti pris étant descriptif et inductif, Elster peut inclure la négociation dans le domaine des décisions collectives, tandis que les approches déductives et normatives basées sur l'agrégation des préférences sont trop exigeantes pour cela (la négociation est laissée aux soins de la théorie des jeux).

La dernière piste consiste en une approche socio-historique de la théorie du choix social et de l'économie normative. Il s'agirait ici d'interpréter l'évolution de cette discipline afin de diagnostiquer certaines tendances, que l'on pourrait retrouver dans d'autres champs du savoir, et qui permettraient de dessiner une évolution sociale plus générale. A propos de la théorie du choix social et plus généralement de la structure économique dite néoclassique, Philippe Mirowski, économiste et philosophe, a entrepris de mettre en parallèle l'évolution de la pensée économique après la Seconde Guerre Mondiale avec l'Histoire des recherches menées par l'armée américaine pendant la guerre froide. La thèse de Mirowski est que la science économique a été profondément modifiée par l'incursion des technologies destinées dans un premier temps à la guerre (notamment l'informatique). Son analyse s'inscrit dans la suite des travaux menés par Donna Haraway sur la « science cyborg »[37] : l'économie orthodoxe est, selon Mirowski, une « science cyborg ». Dans cette perspective, Mirowski interprète le théorème d'Arrow comme une illustration de la conception de la rationalité caractéristique de la politique de la Guerre Froide, une rationalité de type computationnel (le théorème est rebaptisé par Mirowski « l'impossibilité de l'ordinateur démocratique »). Notre objet n'est pas de discuter la forme historique et sociale de rationalité caractérisant le théorème d'Arrow, d'autant que cette rationalité s'avère, de l'aveu même de Mirowski, particulièrement

[37]HARAWAY, Donna, *Manifeste cyborg et autres essais. Sciences - Fictions - Féminismes*, Anthologie établie par Laurence Allard, Delphine Gardey et Nathalie Magnan. Éditions Exils, 2007.

indéterminée. En revanche, le travail d'analyse des présuppositions philosophiques du théorème d'Arrow que nous entendons mener pourrait par la suite nourrir une telle démarche.

4. Un cadre conceptuel et de multiples possibilités

Ce travail discute le cadre conceptuel de la théorie du choix social et de l'économie normative. Ce « cadre conceptuel » correspond à l'agrégation de préférences attribuées à des individus. Il est généralement caractérisé comme « welfariste ». Le terme « welfarisme » vient de l'anglais *welfare,* que l'on traduit par « bien-être ». « Welfarisme » est un néologisme créé par Sen. Selon l'économiste indien, le welfarisme consiste dans la thèse selon laquelle la valeur d'un état de choses ne doit dépendre que des utilités ou des préférences des individus situés dans cet état de chose. La difficulté est que cette thèse peut recevoir deux interprétations indépendantes l'une de l'autre, même si elles sont fréquemment – et souvent implicitement – associées.

La première interprétation associe le welfarisme à ses origines utilitaristes : elle correspond à ce que Marc Fleurbaey appelle le « welfarisme réel »[38] et Antoinette Baujard le « welfarisme philosophique ou moral »[39]. Le welfarisme philosophique associe aux concepts d'utilité ou de préférence un contenu descriptif existant indépendamment de la représentation formelle ; il s'agit généralement d'une expérience de satisfaction. Le chapitre 5 de ce travail s'attachera à développer les significations multiples de cette notion d' « expérience de satisfaction ».

La seconde interprétation s'en tient au contraire strictement au cadre formel. Le welfarisme formel désigne simplement une représentation selon laquelle la fonction de bien-être social dépend exclusivement des utilités individuelles. C'est cette relation logique de dépendance qui intéresse le welfarisme formel, plus que le contenu descriptif qu'on peut lui associer. Notons que, techniquement, le welfarisme est décrit par un lemme réunissant trois conditions arroviennes, la condition de domaine non restreint, la condition d'indépendance des alternatives non pertinentes et la condition de Pareto-faible[40].

---

[38] FLEURBAEY, Marc, « On the informational basis of social choice », *Social Choice and Welfare*, 21, 2003, p 375.

[39] BAUJARD, Antoinette, "A return to Bentham's felicific calculus: from moral welfarism to technical non-welfarism", *European Journal of the History of Economic Thought,* vol 16, no 3, 2009, pp 431-453.

[40] Lemme welfariste :

Soit *f* une fonction de préférences sociales définie sur un ensemble *E* de préférences individuelles. Cette fonction détermine une préférence sociale pour un profil donné de préférences individuelles.

Soit U la condition de domaine non-restreint, stipulant que tous les profils de préférences individuelles sont admis dans le calcul ;

Cette séparation du welfarisme philosophique ou réel et du welfarisme formel rend possible l'association de ce dernier à d'autres interprétations. C'est cette piste que nous avons explorée dans ce travail de recherche, en gardant à l'esprit le questionnement exposé au paragraphe précédent : l'individu est-il nécessairement la source de l'évaluation dans le choix collectif comme agrégation des préférences ? Quel « bien » décrit la préférence ?

L'interprétation du cadre conceptuel de la théorie du choix social inspirée par une vue générale et souvent grossière de l'utilitarisme philosophique s'est avérée n'être en fait qu'une possibilité interprétative particulière, et pas nécessairement celle se trouvant le plus en phase avec les ambitions originelles des « pères fondateurs » que sont Arrow et Sen. Ainsi, il s'est avéré que le cadre conceptuel de la théorie du choix social n'est pas nécessairement individualiste, et que le préférentialisme requis pour une conception du choix collectif basée sur l'agrégation des préférences n'implique pas que lesdites « préférences » en question soient des jugements arbitraires, des caprices, ni même un désir de maximiser son propre bien-être.

Cet élargissement des possibilités interprétatives de l'agrégation des préférences permettrait de réconcilier deux pans des travaux d'Amartya Sen, à savoir la théorie du choix social et l'approche par les capabilités. Ces deux champs d'investigation sont encore peu étudiés conjointement. Les travaux d'Amartya Sen en théorie du choix social restent assez peu connus des philosophes, alors que les capabilités ont trouvé depuis plusieurs années déjà leur place dans les débats de philosophie politique et sociale. En élargissant le champ des interprétations possibles de la théorie du choix social, nous souhaitons montrer que ce cadre conceptuel est « philosophiquement compatible » avec les capabilités, et qu'il est dès lors possible de jeter les bases d'une théorie démocratique de la justice ayant pour métrique les capabilités et les fonctionnements. Ici, nous passons d'une perspective interprétative à une perspective constructive. Cela implique un renouvellement du rapport entre philosophie et économie, qui sera de deux ordres : d'une part, le philosophe fournit la réflexion conceptuelle

---

Soit I la condition d'indépendance des alternatives non pertinentes, selon laquelle la préférence sociale sur une alternative ne doit dépendre que des préférences individuelles sur cette même alternative (et non pas d'information sur les préférences pour d'autres options, ni d'informations sur le degré de préférence en général) ;

Soit P la condition de Pareto : pour chaque paire d'options *x, y,* appartenant à l'environnement de choix, si tous les individus préfèrent strictement *x* à *y,* alors la société préfère aussi *x* à *y*.

Si *f* satisfait les conditions U, I et P, alors il existe une préférence sociale telle que, quelle que soit l'alternative évaluée, quelles que soient les préférences individuelles, la préférence sociale est fonction de ces préférences individuelles.

Cf. ASPREMONT, Claude, "Economie du bien-être et utilitarisme », in L.A. Gérard-Varet et J.C. Passeron, *Le modèle et l'enquête. Les usages du principe de rationalité dans les sciences sociales,* Paris, Editions de la Maison des Sciences de l'Homme, 1995 ; FLEURBAEY, Marc, *Théories économiques de la justice*, Paris, Economica, 1996, p 63 ; FLEURBAEY, Marc, « On the informational basis of social choice », *Social Choice and Welfare*, 21, 2003, p 347-384.

et éthique, qui prendra corps ensuite dans les axiomes de la théorie du choix social[41], et en ce sens la philosophie « nourrit » l'économie ; d'autre part, nos recherches nous ont permis de percevoir ce que John Broome exprime ici clairement :

> *« Je pense qu'il est déjà largement reconnu que les méthodes formelles dérivées de l'économie peuvent contribuer à l'éthique. »*[42]

Il n'est certes pas question ici d'incorporer une vision réductrice de l'*homo oeconomicus* dans une philosophie morale appauvrie. Il est question de se servir du pouvoir clarificateur des concepts formels utilisés en économie, tels que les relations d'ordre ou l'agrégation, sans jamais perdre de vue les possibilités et les impossibilités d'association de ce cadre avec une interprétation donnée. Par exemple, la démarche de Broome dans l'ouvrage cité consiste à se servir de la relation d'ordre, utilisée en économie pour représenter les préférences, afin de clarifier le concept de « bien ». C'est en ce sens que certaines des méthodes économiques « nourrissent » la philosophie.

[41] Cf. FLEURBAEY, Marc, "Social Choice and Just Institutions", *Economics and Philosophy,* 23, 2007, p. 15-43.
[42] BROOME, John, *Weighing Goods*, Oxford : Blackwell, 1991, p ix. Traduction libre.

5. Plan : objet et méthodes

La thèse que nous défendons dans ce travail de recherche est la suivante : les principales postulations philosophiques de la théorie du choix social sont compatibles avec celle de l'approche par les capabilités. Cette thèse s'inscrit donc dans l'esprit de la démarche d'Amartya Sen, qui voit l'ensemble de ses contributions, tant en philosophie qu'en économie, comme constituant un tout[43]. Dans son ouvrage *l'Idée de Justice,* Sen a suggéré d'utiliser le cadre conceptuel de la théorie du choix social comme un cadre de base pour le raisonnement en philosophie sociale et politique et en éthique[44]. C'est cette suggestion que nous avons souhaité approfondir et mettre à l'épreuve de la question du choix démocratique des capabilités. Nous avons donc enquêté sur les postulations philosophiques de la théorie du choix social et de l'économie normative dans cette perspective.

Ces principales postulations concernent la place du point de vue de l'individu dans le choix social (Ière partie) et la nature de la préférence (IIe partie). L'articulation avec l'approche par les capabilités à travers la résolution du problème de la sélection des capabilités pertinentes s'effectue dans la IIe partie.

La Ière partie de cette thèse traite la question suivante : dans la théorie du choix social et l'économie normative, l'individu est-il nécessairement à l'origine des préférences, jugements, utilités ou choix, qui seront par la suite agrégés afin de produire un choix collectif ? La méthode utilisée pour y répondre correspond principalement ici à un travail d'histoire des idées, assorti d'une reconstruction conceptuelle. Nous analyserons en effet trois moments importants de l'histoire de l'agrégation des préférences, l'utilitarisme philosophique, la nouvelle économie du bien-être et son idéal positiviste, et enfin les développements dits « post-welfaristes » visant à élargir la base d'information du choix social en y intégrant une description enrichie des attentes des membres de la société. L'étude de l'évolution des idées à

[43] La première monographie entière consacrée aux capabilités, *Commodities and capabilities*, associe explicitement les capabilités aux outils conceptuels de la théorie du choix social, même si Sen ne pousse pas plus avant l'enquête sur la compatibilité de leurs présupposés conceptuels respectifs. Dans *Development as Freedom,* Sen suggère de voir le problème de l'évaluation des capabilités comme un problème de choix social et de discussion démocratique. Dans l'*Idée de Justice,* la théorie du choix social est présentée comme un cadre de référence pour l' « approche comparative » de la justice qu'il développe.

Cela dit, les contributions de Sen en philosophie et en théorie du choix social sont distinctes, par le langage qu'elles utilisent, par les revues dans lesquelles elles sont publiées, par les chercheurs qui travaillent dessus. Et l'articulation de leurs présupposés conceptuels ne va pas de soi aux yeux des philosophes comme des économistes.

[44] SEN, Amartya K., *The Idea of Justice*, London: Allen Lane, 2009. Traduction française Paul Chemla et Eloi Laurent, *L'Idée de Justice,* Paris : Flammarion, 2009.

propos des comparaisons interpersonnelles met en évidence une première découverte : l'individu n'est pas nécessairement la source substantielle des jugements de valeur et des préférences agrégées pour constituer le choix collectif. Ce qui fait le caractère « collectif » du choix, ce n'est pas le regroupement d'individualités séparées, mais le caractère hétérogène des attentes.

Dans cette première partie, le chapitre 1 est consacré à un premier tournant des conceptions du choix collectif comme agrégation des préférences, à savoir le passage de l'utilité cardinale et comparable à la préférence ordinale non-comparable. Ce tournant est réinterprété à la lumière de la problématique de l'individualisme. Nous verrons que, si la métrique de l'utilité cardinale identifie l'individu à un support de plaisir et de peine, elle n'en présuppose toutefois pas l'existence substantielle. Les données relatives au plaisir et à la peine s'avèrent en fait ne prendre sens que dans et par le calcul institutionnalisé. Cela rend possible le déni de psychologie inauguré par Pareto et les défenseurs de la Nouvelle Economie du Bien-Être, dans une perspective déflationniste qui ne laisse subsister de l'individu qu'un ordre de préférences attribuable à toutes sortes d'entités.

Le chapitre 2 étudie la place de l'individu dans le théorème d'Arrow, en se fondant sur l'étude précise du propos même d'Arrow à ce sujet. Or, il s'avèrera que, pour Arrow, le choix de faire de l'individu la source des préférences relève de considérations éthiques plus que de l'adoption d'une ontologie ou de l'acceptation d'une éventuelle découverte empirique. En effet, le cadre conceptuel qu'il met en place ne requiert pas, par lui-même, que l'individu soit l'auteur ou la cause de ses préférences. Même la réintroduction des comparaisons interpersonnelles ne l'exige pas, puisque que celles-ci ont pour objet des qualités et non des individus substantiels.

Le chapitre 3 analyse la contribution d'Amartya Sen à la lumière de cette même problématique. Le souci de Sen est de réintroduire l'éthique dans l'économie, et cela passe par l'enrichissement de la base informationnelle de l'évaluation sociale et du choix collectif. Cet enrichissement passe par la multiplication des perspectives sur l'individu, multiplication qui culmine dans le concept d'objectivité positionnelle : cette hypothèse relative aux conditions de validité des jugements tant de connaissance que moraux entraîne la possibilité d'attribuer les jugements et les préférences non pas à des individus substantiels, mais à des « positions ». De l'individu auteur du jugement, nous passons à la position comme facteur explicatif des éventuels biais cognitifs dans ces mêmes jugements. Si Sen ne renonce pas à un

« individualisme du sens commun » thématisé par Philip Pettit, il fournit cependant suffisamment d'outils pour une conception de l'individu comme construit plutôt que donné.

Notre IIe partie s'attache à déterminer la nature de la préférence et son objet afin de rendre celle-ci compatible avec les présupposés et les ambitions de l'approche par les capabilités. La méthode adoptée ici est l'analyse conceptuelle[45]. Cette analyse s'attache à valider une définition de la préférence qui soit cohérente avec sa structure formelle, en faisant abstraction des interprétations reçues que l'analyse fait apparaître trop restrictives. Elle vise à définir un ensemble de conditions dans lesquelles cette définition est correcte. Lorsque cela est nécessaire, nous aurons recours aux intuitions communes à propos de ce qu'est une préférence dans le contexte d'un choix collectif, tout en conservant un regard critique sur celles-ci. Il s'avèrera que la définition retenue de la préférence identifie celle-ci à un jugement de valeur comparatif déterminé par une connaissance précise des options proposées au choix.

Dans cette IIe partie, le chapitre 4 traite de la nature de la préférence. L'objectif consiste à situer son statut parmi un vaste réseau de notions incluant les jugements de valeur, les intentions, les décisions ou encore les actions. L'analyse conceptuelle éliminera successivement la préférence définie comme choix et la préférence définie comme cause efficace du choix. Il s'avèrera que, si la préférence peut remplir la fonction de disposition à choisir, son statut premier est celui d'un jugement de valeur, que nous nous attacherons ensuite à caractériser comme essentiellement comparatif, et donc fondé sur la description factuelle des options.

Le chapitre suivant étudie les normes et les critères qui déterminent le jugement de préférence. Il nous faudra à la fois statuer sur les conditions de validité de ces critères et sur leur contenu. Le problème est celui de l'objectivité de ce jugement de valeur spécifique qu'est la préférence. Dans ce chapitre, nous analysons dans un premier temps le problème de la position de l'évaluateur, afin de déterminer si l'objectivité de la description d'une préférence dépend de l'impartialité de celui qui attribue ladite préférence à un agent particulier. Le concept d'objectivité positionnelle de Sen amorcera un retour ici, dans le but de dépasser l'opposition entre jugements de bien-être social émis par un observateur impartial et indépendant, d'une part, et choix sociaux produits par les préférences particulières et subjectives d'individus pris dans le processus d'évaluation et de décision, d'autre part. Dans un second temps, nous étudierons les différentes conceptions de la norme qui détermine la

[45] C'est un point dont nous ne sommes toutefois pas complètement certaine.

préférence, conceptions généralement décrites sous le terme vague et englobant de « bien-être », qui traduit tour à tour les mots anglais *welfare* et *well-being.* Ces conceptions sont les suivantes : le bien-être comme plaisir, le bien-être comme satisfaction des préférences, et enfin le bien-être comme bien objectif[46] (cette dernière conception impliquant deux interprétation possibles de la préférence, l'une comme révélation d'un bien déjà existant, l'autre comme jugement remplissant certaines conditions de rationalité spécifiques). Cette analyse a permis une nouvelle découverte surprenante : pour être tout à fait cohérente, la conception du bien-être comme satisfaction des préférences requiert de la part de l'agent une très bonne connaissance des options qu'elle évalue. En fait, par définition, il apparaît que la préférence échappe en partie aux critiques qui lui sont fréquemment adressées, et peut servir de base pour une théorie démocratique de la justice.

Les fondations d'une telle théorie constituent l'objet de la IIIe partie. Si la méthode de l'analyse conceptuelle y est conservée, en revanche, la finalité cesse d'être interprétative pour devenir constructive. L'objet de cette dernière partie est de jeter les bases d'une évaluation des capabilités fondée sur les préférences. Ces bases sont nécessaires pour spécifier l'approche par les capabilités, valider des listes de capabilités et définir les procédures d'arbitrage entre capabilités. Si cette dernière partie ne constitue pas un travail interprétatif de l'oeuvre d'Amartya Sen à proprement parler, elle s'inscrit néanmoins dans la continuité de son ouvrage *The Idea of Justice,* qui met en place les outils conceptuels que nous utilisons, ainsi que l'idée d'une théorie de la justice basée sur une procédure de comparaison d'états sociaux décrits empiriquement plutôt que sur un raisonnement hypothético-déductif tel que celui mené par Rawls dans la *Théorie de la Justice.*

Le chapitre 6, premier chapitre de cette dernière partie, rappelle et reconstitue lorsque c'est nécessaire l'argumentation en faveur de la métrique des capabilités. Il s'agit d'une argumentation par comparaison, de laquelle il ressort successivement qu'une métrique conséquentialiste est préférable à une métrique déontologique, le conséquentialisme basé sur la description des opportunités est préférable à celui basé sur les seuls résultats, et enfin les opportunités de fonctionnement sont plus intéressantes que les opportunités de bien-être « tout court ». Les débats suscités par la question « quelle égalité » (*Equality of What ?*) et les nombreux arguments qui en ont émergé seront mobilisés à cet effet. Mais notre argumentation, ainsi que le choix d'inscrire cette réflexion sur les capabilités dans le cadre de

[46] Nous verrons alors comment on peut caractériser ce "bien objectif".

l'agrégation des préférences, nous inciteront à choisir une compréhension spécifique du concept de « capabilité », celle qui a prévalu dans les premiers travaux de Sen et qui correspond aux ensembles d'opportunités de fonctionnements. Nous profiterons de cette spécification pour mettre en évidence la diversité des compréhensions de cette notion de capabilité.

Le chapitre 7 propose le concept de « préférence pour des fonctionnements » comme source de l'évaluation des capabilités. Ce chapitre repose pour une grande part sur les analyses menées dans le chapitre 5. Après avoir exposé les termes du problème général de l'évaluation des capabilités et du choix d'un indice des capabilités et des fonctionnements pertinents, nous défendrons la pertinence et la nécessité du concept de « préférence pour des fonctionnements » comme solution à ce problème. Il s'avère que les capabilités et les fonctionnements comme objets des préférences jouent le rôle de filtre à l'égard de ces dernières. En d'autres termes, la valeur de la préférence dépend étroitement du paradigme descriptif adopté pour caractériser « ce qui est préféré ». On peut dire qu'une compréhension adéquate de l'agrégation des préférences devrait se consacrer à ce qui est préféré plus qu'à ceux qui préfèrent, aux objets du choix social plus qu'aux sujets de ce dernier, les raisons de choisir dépendant étroitement de la nature de ces objets et de leur mode d'appréhension.

Au fil de cette étude, il apparaîtra qu'une approche formelle du welfarisme caractéristique de la théorie du choix social permet d'accueil la problématique de l'évaluation des capabilités. Cette évaluation démocratique des capabilités requiert toutefois un renforcement conséquent de la qualité cognitive des préférences. Tant l'évaluation des capabilités que la théorie du choix social ont tout à gagner d'une articulation plus fine des préférences à la situation concrète de l'agent et à la connaissance des moyens à sa disposition dans le monde.

Ière partie : L'individu dans la théorie du choix social, des origines aux plus récents développements

Dans la théorie du choix social, le choix social dépend de préférences généralement attribuées aux individus seuls. Ainsi, Jon Elster attribue à la théorie du choix social une conception de la politique orientée vers l'intérêt individuel privé:

> *"[La théorie du choix social] appartient à une classe plus large de théories partageant certaines caractéristiques. Ces théories ont notamment en commun une conception instrumentale du processus politique, au lieu d'en faire une fin en soi, et la thèse selon laquelle l'acte politique de décision est une action privée plutôt que publique : voyez le vote secret et individuel. Avec cela vient généralement l'idée que le but de la politique est le compromis optimal entre des intérêts privés donnés et irréductiblement opposés. »*[47]

Prenant acte du fait que la base du choix social selon Arrow est l'évaluation individuelle des états sociaux, Elster en déduit que ces évaluations représentent nécessairement les intérêts de ces individus, et que ces intérêts sont fatalement différents de l'intérêt public.

En revanche, l'approche par les capabilités a mis l'accent dès ses origines sur la dépendance et la vulnérabilité[48]. Par ailleurs, elle a souligné l'importance des engagements désintéressés pour la qualité de vie.

L'objet de cette première partie est l'analyse de la conception de l'individu présupposée par la théorie du choix social telle que l'ont développée Arrow et Sen. Plus précisément, la question à laquelle nous tenterons de répondre est la suivante : dans la théorie du choix social, l'individu est nécessairement *à l'origine* des préférences, jugements, utilités ou choix, qui seront par la suite agrégés afin de produire un choix collectif ? Et, s'il l'est, est-ce de manière exclusive ? Le cadre conceptuel de la théorie du choix social requiert-il cet individualisme méthodologique ?

[47] ELSTER, Jon, "The market and the forum: three varieties of political theory", in ELSTER, Jon et HYLLAND, Aanund, eds, *Foundations of Social Choice Theory*, Cambridge University Press, 1986. Traduction libre..

[48] Cf. SEN, Amartya, "Equality of What?", in S. McMURRIN (éd.), *Tanner Lectures on Human Values*, Cambridge: Cambridge University Press, 1980. Sen inaugure la reflexion sur les capabilités en donnant l'exemple d'un individu handicapé.

La théorie du choix social est ici comprise en un sens large, puisque nous intégrons dans notre réflexion aussi bien l'utilitarisme de Bentham que l'approche par les capabilités. L'utilitarisme de Bentham est à l'origine de l'idée d'agrégation des préférences. L'approche par les capabilités fait partie des développements les plus récents de l'économie normative, tout au moins dans la version qu'a proposée Sen.

Cette première partie fait apparaître que le cadre conceptuel de la théorie du choix social au sens large ne présuppose pas un individualisme substantiel : il ne présuppose pas l'existence d'un individu autosuffisant et souverain, capable de déterminer par lui-même ses préférences. La primauté de l'individu n'est pas une nécessité logique pour la théorie du choix social : le formalisme peut s'en passer. Ce n'est pas non plus une nécessité épistémique, puisque les volontés individuelles ne sont pas observables. Ce peut être une nécessité éthique, lorsque les théoriciens du choix social revendiquent la valeur de la perspective évaluative de l'individu[49].

Le premier chapitre de cette première partie est consacré à l'analyse de la conception de l'individu dans l'utilitarisme classique de Bentham. Le passage de cet utilitarisme à la Nouvelle Economie du Bien-Etre et ses conséquences pour la conception de l'individu sont également abordés. Le deuxième chapitre traite de la place de l'individu dans le cadre conceptuel que met en place Arrow pour son théorème d'impossibilité. Le troisième chapitre renouvelle l'exercice en tentant d'articuler les différentes contributions d'Amartya Sen afin de déterminer la conception de l'individu qu'elles doivent présupposer.

[49] En ce sens, ils seraient libéraux, dans la mesure où le libéralisme correspond aux doctrines de philosophie politique qui affirment la valeur du jugement individuel pour déterminer l'action collective.

Chapitre 1 : De l'attribution des utilités et des ophélimités aux individus

L'utilitarisme de Bentham est à l'origine du cadre conceptuel de la théorie du choix social. Il repose sur le principe d'utilité. Ce principe aurait pour une première fois été formulé par Hutcheson, dans *An Inquiry into the Original of Our Ideas of Beauty and Virtue* en 1725. L'utilité décrit et quantifie le plaisir et la peine individuels. Le principe d'utilité affirme à la fois que les individus recherchent de fait l'accroissement de leur utilité, et que la société devrait maximiser l'utilité agrégée de tous ses membres.

Au début du XXe siècle, l'économie du bien-être s'est coupée de la philosophie utilitariste. Les économistes Vilfredo Pareto, puis Lionel Robbins, promeuvent alors le remplacement de l'utilité par les ordres de préférence.

De l'utilitarisme à la Nouvelle Economie du Bien-Etre, quelle est la place de l'individu, et quelle est la signification de l'attribution des utilités comme des préférences aux individus ? Notre enquête sur les textes de Bentham comme de Pareto a mis en évidence que leurs approches ne considèrent pas la psychologie individuelle comme une donnée indépendante du contexte social et institutionnel d'énonciation des utilités et des préférences des individus. Il est donc inexact d'affirmer à leur propos qu'une psychologie individuelle détermine les préférences ou les utilités indépendamment de l'environnement social et institutionnel.

Nous verrons donc que, si les individus de Bentham semblent à première vue être des supports donnés d'expériences de plaisir et de peine, la vision du philosophe anglais est en fait plus complexe. En effet, une lecture plus fine de son œuvre fait apparaître la nécessité de dispositifs institutionnels et sociaux pour formuler les attentes individuelles en termes d'utilités (1.1.). La psychologie et la physiologie de l'individu ne sont pas des données indépendantes de cette formulation. Cette psychologie et cette physiologie disparaissent ensuite, dans la mesure où le passage de l'utilitarisme classique à la formulation des attentes en termes de préférences ordinales revient à dépsychologiser complètement l'individu (1.2.).

1.1. Utilités individuelles, mécanique d'agrégation et dispositifs institutionnels

Si les expériences individuelles de plaisir et de douleur semblent être la base première de l'évaluation sociale utilitariste élaborée par J. Bentham (1.1.1.), la mécanique même

d'agrégation, c'est-à-dire de collecte de ces expériences, met en évidence l'existence de jugements de valeur supra-individuels, incarnés par la figure du « législateur », réel ou fictif (1.1.2.). En fait, comme le montre Christian Laval, on peut parler d'un cercle utilitariste : l'individu ne s'exprime dans le choix collectif qu'au travers d'une construction à laquelle participe le processus même de choix collectif (1.1.3.).

1.1.1. Le critère d'utilité : plaisir et peine des corps individuels

L'utilitarisme classique fait de l'individu la source de l'évaluation sociale en lui attribuant une « utilité », c'est-à-dire la mesure de son degré de satisfaction dans une configuration sociale donnée. Recueillir l'information relative aux utilités est indispensable pour produire une évaluation sociale substantielle.

Un tel choix théorique procède à l'origine d'une certaine conception de ce que peut être une « science de la nature humaine » ; Jeremy Bentham définit d'abord l'homme comme un être capable d'éprouver du plaisir et de la peine. La nature de l'homme est celle d'un être doué des cinq sens ; toute expérience humaine ayant de la valeur peut être réduite à une expérience de plaisir ou de peine. Cette expérience est, d'après Bentham, mesurable : il y a des degrés de plaisir et de peine. L'utilité est la mesure de cette expérience ; elle constitue l'information nécessaire pour évaluer une société et pour prescrire l'action publique, c'est-à-dire l'action effectuée en vue d'accroître le bien-être collectif.

La mesure de l'utilité d'un individu est un jugement de valeur : il s'agit d'évaluer le degré de satisfaction procuré par une politique économique, l'application d'une nouvelle loi ou encore un spectacle.

Chaque individu constitue ainsi une unité indépendante des autres individus ; il n'y a pas d'évaluation de la société indépendamment des aspirations de ses membres, l'évaluation sociale passant par l'évaluation individuelle : le processus d'agrégation permet de passer du niveau individuel au niveau collectif. Par la suite, l'art de la politique consiste à harmoniser ces intérêts individuels, au besoin au moyen de « fictions » ; l'intérêt général est alors soit le résultat d'un calcul, d'une arithmétique sociale, soit une notion illégitime qui ne peut avoir de valeur que si elle accroît l'utilité agrégée de la société.

> *« L'intérêt de la communauté est une des expressions les plus approximatives que l'on puisse trouver dans la phraséologie de la morale: il n'est pas étonnant que son sens soit si souvent perdu. Quand elle a un sens, c'est celui-ci: la communauté est un corps fictif, composé des personnes individuelles qui*

sont considérées comme constituant ce corps, comme s'ils étaient ses membres. Qu'est-ce alors que l'intérêt de la communauté, si ce n'est la somme des intérêts des membres qui la composent? »[50]

Cette conception atomiste de la collectivité est fondée sur l'observation, et ce en deux sens :

- L'observation élémentaire de la séparation des corps physiques, chaque individu correspondant à un corps capable d'éprouver le plaisir et la peine qui font la valeur d'une configuration sociale donnée. De ce point de vue, l'atomisme social de Bentham est fidèle au projet newtonien de départ.
- L'observation du fait que les individus sont d'abord motivés par la poursuite du plaisir et l'évitement de la peine ; une action rationnelle est une action qui vise cet objectif par les moyens les plus appropriés. Le newtonisme moral de Bentham fait de l'intérêt une véritable « force », au sens que la physique donne à ce terme.

L'utilité revêt ainsi à la fois une valeur descriptive et une valeur normative : elle rend compte de traits propres à la nature humaine tout en servant de norme pour l'action individuelle et pour l'action collective. Mais l'action collective n'est possible que suite à un processus d'agrégation des intérêts et de calcul des pertes et profits en matière de plaisirs et de peines. C'est là la tâche du législateur. Celui-ci doit également harmoniser artificiellement les intérêts individuels divergents, et c'est en cela que consiste l'art de la politique. Notons que la thèse benthamienne de l'harmonisation artificielle des intérêts s'oppose à la thèse d'une harmonie naturelle des intérêts défendue par Adam Smith dans le cadre des mécanismes d'échange de biens et de travail [51].

[50] BENTHAM, Jeremy, *An Introduction to the Principles of Morals and Legislation*, Dover Publications, Inc., Mineola, New York, 2007, I, IV, p 3. Traduction libre..

[51] A ce sujet, cf. HALEVY Elie, *La Formation du Radicalisme philosophique*, T.I, Paris, Presses Universitaires de France, 1996[2], p 25-27.

1.1.2. Le calcul utilitariste et ses règles

Les évaluations individuelles de l'utilité ne sont pas les seuls jugements de valeur pris en compte dans le système benthamien. La méthode d'agrégation et le calcul de maximisation présupposent les propositions suivantes :

- Chaque individu membre de la société compte pour un au moment où l'on recense les informations sur l'utilité.
- L'efficacité est le critère premier d'évaluation des politiques publiques ; les considérations telles que l'intégrité des personnes ou la coutume lui sont subordonnées. La maximisation de l'utilité est prioritaire. L'éthique benthamienne est une éthique conséquentialiste, pour laquelle c'est l'effet qui compte en premier lieu.
- Le « législateur », terme qui désigne la ou les personnes chargées de collecter les informations sur les utilités individuelles, d'évaluer les conséquences des différents choix possibles sur le bien-être social total et de prendre les décisions appropriées, est impartial. Bentham imagine ainsi différents dispositifs permettant de garantir l'impartialité, comme la publicité des débats et le contrôle de l'action publique par les citoyens.
- Les utilités d'individus différents sont mesurables et comparables : elles peuvent être évaluées à l'aide d'une échelle commune ; sans cela, faire la somme des utilités individuelles pour en déduire la quantité de bien-être collectif n'aurait pas de sens.

Les jugements de valeur sous-tendus par ces quatre propositions permettent ensuite d'agréger les utilités individuelles. Quel est le statut de ces jugements? Si, dans la perspective benthamienne, la légitimité d'un jugement de valeur lui est conférée par sa « réalité », au sens où il est traduisible en plaisirs et peines sensibles, peut-on en dire autant de ces considérations d'agrégation ? Celles-ci permettent de passer de l'utilité, propriété positive d'un état de l'individu, au principe d'utilité comme norme de l'action sociale : elles sont d'une importance vitale. L'éthique sociale utilitariste de Bentham ne peut être dite fondée exclusivement sur les évaluations individuelles, puisqu'il y entre nécessairement des normes qui ne sont pas déterminées par les individus membres de la société.

Parmi ces normes, figure la règle selon laquelle chaque membre de la société « compte pour un » dans le calcul des utilités au moment où le législateur recueille les données. L'utilitarisme élimine ainsi *ex ante* la priorité des meilleurs comme celle des plus défavorisés.

Ce qui n'empêche pas les décisions sociales du d'être inégalitaires *ex post* , en vertu des propensions individuelles à obtenir plus ou moins de satisfaction de la décision en question.

Etablir que la société doit maximiser le bonheur collectif, entendu comme agrégation des bonheurs individuels, revient à dire que les décisions sociales doivent aller dans le sens des attentes individuelles – étant donné l'hypothèse selon laquelle les individus recherchent généralement la maximisation de leur bien-être - . Le contenu de la décision sociale dépend ainsi exclusivement des attentes individuelles, explicites ou non. Cette stricte dépendance constitue également une norme parente de la condition dite de Pareto en théorie du choix social. C'est la souveraineté du peuple exprimée en termes d'agrégation. Cette norme n'est pas d'origine individuelle.

Il est donc inexact d'établir que *seule* la propension individuelle à rechercher le plaisir et à fuir la peine est à la base des décisions sociales. Le législateur a besoin de règles d'agrégation des données individuelles. Ce choix est déterminé par des considérations normatives qui ne dépendent pas directement des utilités individuelles. La conception benthamienne de la décision sociale n'est donc pas exclusivement basée sur les utilités individuelles, et *in fine* sur les lois psychologiques de l'action humaine.

1.1.3. Le cercle utilitariste : les dispositifs institutionnels et leurs effets sur la formulation des attentes en termes d'utilités

La doctrine benthamienne de l'évaluation sociale ne doit donc pas être comprise comme un mouvement à sens unique, des utilités individuelles réelles vers un bien-être collectif qui leur serait subordonné. Comme le soutient Christian Laval, le mouvement est en fait circulaire :

> *« Le jugement individuel est donc pratiquement déterminé par la considération des intérêts de la somme des individus, considération qui doit être au principe de l'action du législateur. On a donc affaire à un mouvement circulaire : "l'intérêt du plus grand nombre des individus" est déterminé par la sommation des intérêts individuels. Cet intérêt, qui doit être maximisé, détermine les intérêts individuels eux-mêmes par la distribution des droits et obligations et par la définition des peines et des récompenses, c'est-à-dire via le tissu des fictions productrices de plaisirs et de peines. A partir de là, peut-être comprend-on mieux pourquoi Bentham a pu regarder ses principes de la quantification et sa théorie du langage comme ses deux instruments fondamentaux. »*[52]

[52] LAVAL, Christian, « Fiction et utilité chez Jeremy Bentham », Revue du MAUSS n° 6, *Qu'est-ce que l'utilitarisme ? Une énigme dans l'histoire des idées* (1995, 2e semestre). Cf. également LAVAL, Christian, *Jeremy Bentham. Le pouvoir des fictions.*,Paris, PUF « Philosophies », 1994.

Par ailleurs, les *Fragments sur le Gouvernement*[53]ainsi que les *Sophismes Anarchiques*[54]critiquent la conception de l'individu « à l'état de nature » et la taxent de pure fiction. Pour Bentham, il n'y a pas d'individu à l'état de nature, formant ses attentes indépendamment des autres individus.

Par conséquent, tous les dispositifs au service du recueil des données individuelles, de leur agrégation et du calcul du bien-être social sont « déjà là » au moment de l'évaluation des situations individuelles et de la formulation des attentes en matière de plaisir et de peine. Ces dispositifs structurent ces attentes par la menace de sanctions et la promesse de récompenses : l'art du législateur consiste en effet à harmoniser les intérêts en régulant la propension de chacun à rechercher sa satisfaction à tout prix. Ces dispositifs sont très divers, de la sanction légale à la surveillance réciproque des gouvernants et des gouvernés, du système éducatif à la réglementation économique. Christian Laval montre que la construction même d'un langage dans lequel les intérêts individuels puissent être formulés *et* harmonisés est requise par la science de la législation. Créer une harmonie artificielle des intérêts justifie un recours systématique aux fictions du langage. Ainsi, le bien-être social n'est pas généré spontanément à partir d'intérêts individuels constitués de façon autonome. La possibilité même de dériver le bien-être collectif et la décision sociale des utilités individuelles est conditionnée par des fictions langagières déterminantes pour la formulation et l'agrégation du plaisir et de la peine.

La tâche du philosophe utilitariste consiste ainsi à élaborer une langue dans laquelle l'intérêt puisse être exprimé de la manière la plus adéquate à la réalisation du plus grand bonheur du plus grand nombre. La recherche d'un espace d'évaluation adéquat aux attentes *réelles* des membres de la société est en fait la recherche d'une langue de l'intérêt qui s'inscrive au mieux dans les attentes des sociétés présentes ou à venir.

Christian Laval évoque également « la fiction du jugement collectif composé de la somme de jugements individuels formulés dans le secret »[55], pour en signaler par ailleurs son insuffisance à exprimer l'opinion publique. Non que la démocratie soit un leurre et que les suffrages soient manipulés : le recours à la fiction est inévitable. L'institutionnalisation du vote comme paradigme de l'exercice de la souveraineté suppose la mobilisation de constructions langagières propres.

La « matière » même du choix collectif utilitariste n'est donc pas donnée, mais dépend des dispositifs institutionnels incluant le langage utilisé pour formuler les attentes. Certes, le

[53] BENTHAM, Jeremy, *A Fragment on Government,* reéd. Cambridge University Press, 1988.
[54] BENTHAM, Jeremy, *Anarchical Fallacies,* éd. 1843.
[55] LAVAL, Christian, *Jeremy Bentham. Le pouvoir des fictions.*, Paris, PUF « Philosophies », 1994.

plaisir et la peine sont des événements individuels au sens où ils sont supportés par des corps individuels, c'est-à-dire des corps singuliers et distincts d'autres corps. Mais la traduction de ces événements en revendications passe par un processus d'identification de ce corps : il faut le nommer, le rendre reconnaissable, l'inscrire dans la collectivité comme membre de celle-ci, membre qui « compte pour une voix »[56]. Plaisir et peine s'éprouvent dans une configuration sociale particulière. Ce processus rend l'individu « assignable » par son nom ou sa fonction[57]. Un individu non assignable ne peut ni revendiquer droits et réparations, ni être assigné à comparaître : la justice ne fonctionne pas sans la construction sociale de l'identité, et il en va de même pour les différentes formes, directes ou indirectes, de choix collectif[58].

Est-ce à dire que la mesure de l'utilité d'un individu dans une configuration sociale donnée est toujours contingente ? Autrement dit, que vaut le projet benthamien de refonte du langage ? En fait, ce qui permet de distinguer les fictions utiles des fictions néfastes, le langage qui sert la maximisation du bonheur collectif du langage mis au service des intérêts corporatistes, c'est, à défaut de l'adéquation, la contiguïté avec l'expérience réelle. Certaines fictions sont moins arbitraires que d'autres.

Ainsi, à l'aube de la théorie du choix social et du bien-être collectif, loin de constituer la base matérielle du choix collectif, l'individu ne se laisse appréhender qu'au travers des constructions institutionnelles et langagières.

La quête d'un langage exprimant au mieux les attentes collectives va de pair avec la traque des formules arbitraires. La Nouvelle Economie du Bien-être est initiée par l'économiste italien Vilfredo Pareto ; faisant le pari du formalisme comme langage le plus susceptible de

[56] *Ibidem,* p 46-47.

[57] BENTHAM, Jeremy, *Traité de Législation Civile et Pénale*, Tome III, ED Dumont, ch. VI.

[58] On peut se demander si la perspective benthamienne ne fait pas même du « Moi » une fiction ? Le corps individuel régi par les lois naturelles d'attraction pour le plaisir et de répulsion pour la peine est dans un premier temps la condition nécessaire du Moi auquel il sert de support. Mais en est-il la condition suffisante ? Le Moi, c'est l'identification de soi par soi, notamment par la différenciation d'avec autrui. Les corps humains sont singuliers, sans doute. L'identité génétique, si cette formule a un sens, est unique. Mais cette identité génétique n'est pas immédiatement perceptible : je n'ai jamais expérimenté mon ADN. Ce qui est perçu, ce sont les expériences de plaisir et de douleur ; or, précisément, ces expériences caractérisent le corps humain de façon générale c'est cela même qui confère au principe d'utilité sa validité, dans la perspective de Bentham – . Le problème de l'identification des individus, par eux-mêmes et par les autres individus, apparaît lorsqu'il devient nécessaire d'énoncer des jugements leur attribuant les expériences actuelles ou potentielles de plaisir et de peine. Cela devient nécessaire au moment où des formes d'organisation sociale se proposent de contribuer à la réalisation du plaisir et à l'atténuation de la peine (conformément aux attractions et répulsions spontanées des êtres humains). Logiquement, plus ces formes d'organisation sociale sont abouties – et moins elles sont corrompues par des groupes d'intérêts particuliers –, plus l'identification de l'individu comme support de plaisirs et de peines à gérer collectivement sera poussée. Toute organisation vouée à la gestion des plaisirs et des peines met en place, de manière plus ou moins complexifiée, des dispositifs d'identification : le nom, et tout ce qui est susceptible de désigner – d' « assigner » – un individu à l'exclusion des autres, comme les caractéristiques physiques, l'ADN, la profession, sont autant d'informations collectées dans cette perspective. La différenciation, qui permet de s'auto-désigner comme « Moi », intervient lors de la construction sociale des identités. Les paramètres de cette construction évoluent selon le degré de complexification des formes sociales, et surtout selon le système de représentations en vigueur.

produire des propositions proches des attentes réelles de l'être humain, elle élimine les informations individuelles dont le législateur benthamien ne peut maîtriser la portée. L'intensité du plaisir ou de la peine éprouvés dans une configuration sociale donnée dépend trop fortement des caractéristiques particulières de celui ou celle qui les éprouve pour faire l'objet d'une expression fiable. Information imparfaite, elle sera purement et simplement éliminée. En d'autres termes, le caractère arbitraire des comparaisons interpersonnelles d'utilité amène l'économiste italien Vilfredo Pareto à renoncer à l'utilité comme critère d'évaluation, lui substituant ce qu'il appelle l' « ophélimité », une version ordinale de l'utilité.

1.2. Ophélimités et préférences ordinales : l'individu dépsychologisé

Pareto sera, avec plusieurs années d'avance, le précurseur d'un véritable tournant pour l'économie du bien-être. Ce tournant est qualifié de « positiviste », car il a éliminé des modèles de l'économie du bien-être toute référence à des entités non observables, parmi lesquelles la psychologie de l'individu, en tant qu'elle détermine les expériences de plaisir et de peine.

Au début de son *Manuel d'Economie politique* paru en 1907, Pareto expose les « principes généraux » d'une économie scientifique au sens propre du terme. L'économie devrait être exclusivement fondée sur l'observation de phénomènes concrets ; ensuite l'économiste procède par abstraction, il sélectionne parmi les phénomènes ceux qui sont pertinents pour sa recherche. L'économie s'intéresse aux actions « économiques » de ces corps, c'est-à-dire aux actions réalisées en vue de l'intérêt. L'*homo oeconomicus* est une abstraction au même titre que l'*homo ethicus* ou l'*homo religiosus*[59]. L'économie politique n'a donc pas vocation d'expliquer l'ensemble des comportements humains, et les autres domaines de l'étude de l'humain n'ont pas vocation d'expliquer les actions économiques[60]. Ainsi, l'économie parétienne se constitue comme science en définissant son objet par abstraction et en limitant son champ d'investigation. Mais si l'objet diffère, la méthode s'inspire clairement de celle des sciences de la nature. Pareto compare d'ailleurs l'abstraction économique à l'abstraction de la mécanique rationnelle ou de la chimie[61].

[59] PARETO, Vilfredo, *Manuel d'Economie Politique,* 1909, rééd. Genève, Librairie Droz, 1966, VI.

[60] Contrairement aux ambitions théoriques de Gary Becker, par exemple, qui défend l'explication économique de comportements considérés traditionnellement comme ne relevant justement pas de l'économie (le mariage, l'éducation des enfants, le crime ou les discriminations).

[61] Il serait intéressant de voir quelle science économique pourrait se constituer en se calquant sur le modèle de la physique quantique!

Il est donc possible d'abstraire parmi les comportements humains ceux qui intéressent l'économie.

> *« La mécanique rationnelle, quand elle réduit les corps à de simples points matériels, l'économie pure, quand elle réduit les hommes réels à l'*homo oeconomicus*, se servent d'abstractions parfaitement semblables et imposées par des nécessités semblables. »*[62]

Pareto met alors l'accent sur un problème technique : le problème de la comparabilité des utilités cardinales de chaque individu. L'utilité est la mesure de l'effet d'une configuration économique et sociale sur la satisfaction d'un individu. Cette mesure est dite « cardinale » lorsque cet effet est mesuré en termes de magnitude. Les comparaisons d'utilité entre individus dépendent de la possibilité d'utiliser une même échelle de mesure pour comparer la magnitude de satisfaction d'individus différents. Le critère utilitariste ne tient donc que si ces comparaisons interpersonnelles d'utilité peuvent être exprimées à l'aide d'un étalon de valeur commun, étalon qui doit être trouvé par l'observation des individus et non pas décrété par un agent extérieur. Or, cela paraît impossible. Pareto nie alors que de telles comparaisons, ainsi que l'idée de « plus grand bonheur du plus grand nombre » puisse avoir une quelconque signification.

Pareto propose donc de remplacer l'utilité par l' « ophélimité », qui n'a qu'une valeur ordinale (contrairement à l'utilité, qui a une valeur cardinale). « Ophélimité » est un néologisme construit à partir du grec ancien ώφέληματος, qui signifie « utilité », « avantage », « profit »[63]. Pareto substitue « ophélimité » à « utilité » pour deux raisons. La première est exposée ici :

> *"Le mot utilité est amené à signifier en économie politique tout autre chose que ce qu'il peut signifier dans le langage courant*[64]*. C'est ainsi que la morphine n'est pas utile, au sens ordinaire du mot, puisqu'elle est nuisible au morphinomane ; elle lui est au contraire utile économiquement, puisqu'elle satisfait un de ses besoins, alors même qu'il est malsain. Bien que les anciens économistes aient déjà fait mention de cette équivoque, on l'oubliait encore parfois : aussi est-il indispensable de ne pas employer le même mot pour indiquer des choses aussi différentes. Nous avons proposé dans notre Cours de désigner l'utilité économique par le mot ophélimité, que d'autres auteurs ont adopté depuis. »*[65]

---

[62] PARETO, Vilfredo, *Manuel d'Economie Politique,* 1909, rééd. Genève, Librairie Droz, 1966, II, 21, p 17

[63] PARETO, Vilfredo, *Cours d'Economie politique,* rééd. Librairie Droz, Genève, 1964, § 5.

[64] La célèbre *Préface* de *Mademoiselle de Maupin* de Théophile Gautier témoigne de ce malentendu qui hante la réception de l'utilitarisme.

[65] PARETO, Vilfredo, *Manuel d'Economie Politique*, 1909, rééd. Genève, Librairie Droz, 1966, III, 30, p 157.

Le terme « ophélimité » élimine ainsi une ambiguïté. Une société de morphinomanes satisfaits est une société qui remplit les critères de bien-être des partisans de l'utilité. Pareto justifierait une telle position en arguant que l'utilité-ophélimité est un concept descriptif et non normatif. En ce sens, la nouvelle économie du bien-être rompt avec la philosophie utilitariste traditionnelle et son projet éthique.

L'ophélimité sert à décrire des phénomènes humains : les ressources n'ont pas d'ophélimité propre. Dans son *Cours*, Pareto signale ainsi que, si l'espèce humaine venait à disparaître, l'or ne serait plus qu'un métal rare, mou, mais il serait dénué de valeur parce que d'ophélimité. L'ophélimité est subjective : pour beaucoup d'entre nous, la morphine n'a pas d'ophélimité. L'ophélimité est un « rapport de convenance » entre un être humain particulier et un bien, et ce dans une situation particulière.

La rupture de l'ophélimité avec l'utilité n'est pas que terminologique, et ne tient pas qu'à l'autonomisation de l'économie à l'égard des considérations éthiques. En fait, elle est surtout due au passage de l'utilité cardinale à l'ophélimité ordinale. Prenons l'exemple du choix du dessert dans un restaurant par un individu donné. Le tableau suivant expose la valeur de chaque dessert pour cet individu, en termes d'utilité et en termes d'ophélimité :

	Valeur en utilité (magnitude de satisfaction lors de la consommation)[66]	Valeur en ophélimité (rang dans le classement)
Tarte aux pommes	4	1er
Mousse au chocolat	2	2e
Salade de fruits	1	3e

Tableau 1. Utilité et ophélimité

L'ophélimité désigne donc l'élément préféré, le premier d'un classement. Elle fait abstraction de nombreuses informations, mais offre l'avantage de la simplicité, et de la possibilité de la déduire de la seule observation des comportements de choix, sans recours à l'introspection ou à des comparaisons nécessitant l'introduction d'une échelle supplémentaire. Le terme « ophélimité » n'aura pas de postériorité et sera remplacé par celui de « préférence » ou de « préférence ordinale ».

[66] Cette valeur peut correspondre au prix que le consommateur est prêt à payer, ou bien à la note qu'il mettrait à chaque dessert.

Ces préférences, ou ophélimités, ne peuvent être additionnées comme les utilités ; pour déterminer une préférence collective, il suffit de les agréger ; les éléments qui dominent les autres dans le classement final seront considérés comme collectivement préférés.

Le critère utilitariste du « plus grand bonheur du plus grand nombre » est reformulé : ce sera le maximum d'ophélimité, qui s'énonce ainsi :

> *« Nous dirons que les membres d'une collectivité jouissent, dans une certaine position, du* maximum d'ophélimité, *quand il est impossible de trouver un moyen de s'éloigner très peu de cette position, de telle sorte que l'ophélimité dont jouit chacun des individus de cette collectivité augmente ou diminue. C'est-à-dire que tout petit déplacement à partir de cette position a nécessairement pour effet d'augmenter l'ophélimité dont jouissent certains individus, et de diminuer celle dont jouissent d'autres : d'être agréable aux uns, désagréables aux autres. »*[67]

Ce « maximum d'ophélimité », plus tard rebaptisé « optimum de Pareto », ne requiert que les informations suivantes : la délimitation de la collectivité considérée et les préférences des individus. Ce critère est défini négativement, par ce qu'il n'est pas, à savoir une configuration économique – ou sociale – dans laquelle il serait possible d'augmenter l'ophélimité de certains sans réduire celle d'autres. Il s'agit d'un critère minimaliste, déflationniste même. Au lieu de qualifier la situation souhaitable par ce qu'elle est, on la caractérise par ce qu'elle n'est pas. La théorie déflationniste de la vérité comme le critère parétien économise les entités : la première stipule qu'affirmer qu'une proposition est vraie revient à affirmer la proposition ; la vérité n'est pas une propriété surajoutée à la proposition vraie. Dire que *p* est vraie revient à dire que *p* est un fait. Du côté du critère parétien, le critère de valeur n'est pas une propriété surajoutée à une situation optimale au sens de Pareto. Dire qu'une situation est Pareto-optimale revient à dire qu'elle n'est pas non-optimale au sens de Pareto, et rien de plus. En d'autres termes, la société est « bonne », ou souhaitable, en vertu des seuls jugements de ses membres. Le Bien n'est pas une propriété surajoutée à la société. Ce positionnement déflationniste est qualifié par Arrow de « nominaliste » et opposé à un positionnement « platonicien » :

> *"Pour les nominalistes des temps modernes, l'hypothèse d'existence d'un idéal social appartenant à un univers platonicien est dépourvue de sens. La philosophie utilitariste de Jeremy Bentham et de ses successeurs essaya, au contraire, de fonder le Bien social sur le Bien individuel. La psychologie hédoniste, associée à cette philosophie a, par la suite, permis de soutenir que le Bien de tout individu*

[67]PARETO, Vilfredo, *Manuel d'Economie Politique,* 1909, rééd. Genève, Librairie Droz, 1966, VI, 33, p 354.

s'identifie à ses aspirations. Par la suite, le Bien social devait apparaître, dans un certain sens, comme une synthèse des aspirations individuelles. »[68]

L'ophélimité individuelle constitue la seule matière du Bien, c'est-à-dire de la fin poursuivie par l'action individuelle et les politiques collectives, d'une part, et le critère permettant de choisir une configuration économique plutôt qu'une autre, d'autre part. Pour Pareto et pour Arrow, la seule réalité du Bien réside dans la satisfaction individuelle. Les individus sont vus comme des corps réels sources d'actions diverses.

En 1932, Lionel Robbins, dans son *Essay on the Nature and Significance of Economic Science*, plaide pour l'abandon des comparaisons interpersonnelles d'utilité, dans la mesure où celles-ci sont arbitraires et superflues: quelle que soit la quantité de plaisir procurée par l'objet choisi, l'agent le choisit et c'est ce qui devrait intéresser en premier lieu l'économiste. Comparer la magnitude de la satisfaction éprouvée par le boulanger lorsqu'il vend son pain six dollars avec celle de la satisfaction du client n'est ni nécessaire ni souhaitable, affirme Robbins[69].

La nouvelle économie du bien-être dépsychologise ainsi l'individu. De fait, il n'a d'existence qu'en tant qu'il préfère et choisit. Il est juste agent ; les autres dimensions de son existence n'intéressent plus l'économie. Mais cet individu reste le support supposé d'une préférence distincte d'autres préférences. On suppose que « derrière » ces préférences se trouve un individu doué d'une faculté de choix qu'il peut exercer sans entrave. Une telle supposition procède-t-elle nécessairement du cadre conceptuel welfariste propre à la Nouvelle Economie du Bien-Etre comme à l'utilitarisme ?

[68]ARROW, Kenneth, *Social Choice and Individual Values*, New Haven, Yale University Press, 1951 (2ème édition révisée, 1961).Traduction française par l'Association de Traduction Economique de l'Université de Montpellier, Paris : Calmann-Levy, 1974, p54.

[69] ROBBINS, An Essay on the Nature and Significance of Economic Theory, London, McMillan, 1932, 1945, ch. 6.

Chapitre 2: L'individu dans la théorie du choix social arrovienne

Dans la théorie du choix social telle qu'Arrow la met en place en 1951, les préférences individuelles sont la base d'information du choix collectif. L'individu semble donc bien être à l'origine desdites préférences.

Mais le cadre conceptuel de la théorie du choix social (au sens large) est une idéalisation, et ne retient de la réalité que les aspects pertinents pour ses deux objectifs, l'évaluation du bien-être collectif et l'étude des procédures de vote. La théorie du choix social n'a donc pas de prétention explicative, et on ne peut inférer de son cadre même la thèse selon laquelle les individus détermineraient exclusivement et de manière autonome leurs préférences.

Cependant, Arrow et ses successeurs soulignent l'importance de la « souveraineté de l'agent », c'est-à-dire l'importance de reconnaître la primauté de la perspective évaluative[70] de l'agent individuel. Cette importance peut se comprendre en deux sens. Dans un premier sens, elle signifie que les agents individuels déterminent de fait leurs préférences, et par conséquent le réalisme du modèle du choix social dépend d'une description adéquate de cet état de fait. Dans un second sens, elle signifie que, même si les agents ne déterminent pas de fait leurs préférences, le droit de le faire devrait leur être reconnu. L'analyse de la contribution d'Arrow, notamment des articles sur les enjeux philosophiques de la théorie du choix social, montre que c'est dans le second sens que la primauté de la perspective évaluative de l'individu doit être comprise chez Arrow. Par conséquent, l'individu n'est explicitement pas considéré par Arrow comme la source de fait exclusive des préférences et des jugements de valeur qui comptent dans pour choix social. En revanche, l'attribution de ces préférences aux individus procède du souci de valoriser la perspective évaluative de ces individus.

Nous étudierons la place de l'individu dans le cadre conceptuel arrovien et les deux raisons pour lesquelles le théoricien du choix social peut vouloir voir l'individu comme la source des préférences, à savoir une raison épistémique et une raison éthique (2.1.). Nous montrerons ensuite que cet « individualisme » doit être nuancé par trois considérations : Arrow ne réserve pas la rationalité à l'individu, les motivations de l'individu ne sont pas vues comme nécessairement déterminées par l'intérêt individuel, et la logique atomiste de la théorie du choix social est compatible avec une description empirique des relations sociales (2.2.). Enfin,

[70] Cette formulation est de R. Arneson.

nous mettrons en évidence la motivation d'Arrow à l'origine du refus des comparaisons interpersonnelles d'utilité : plus que le caractère arbitraire de ces comparaisons, Arrow est préoccupé par la question de l'intégrité de l'individu et de l'autonomie de ses choix (2.3.).

2.1. L'individu, support des préférences

Arrow défend une conception de l'individu soucieuse de préserver l'autonomie de ce dernier. L'individu semble avoir la primauté dans la théorie du choix social : les jugements de valeur qui sont les composantes mêmes du choix collectif sont des jugements de valeur individuels (2.1.1.). Ce choix théorique est justifié par des raisons épistémiques : l'individu particulier, en tant qu'organisme perceptible, devrait être une entité observable satisfaisante pour une approche positiviste (2.1.2.) ; il se comprend également d'un point de vue « éthique » ou normatif : dans les articles postérieurs à la publication de son théorème d'impossibilité, Arrow rattachera son choix conceptuel à la tradition libérale qui a fait de l'autonomie de l'individu et du respect des choix de chacun une norme fondamentale (2.1.3.).

2.1.1. Arrow, défenseur de la primauté del'évaluation individuelle

La primauté de l'évaluation individuelle apparaît dans la conception de la démocratie exposée par Arrow (a), dans le caractère mécanique de l'agrégation des préférences individuelles (b) et dans la définition même des jugements de valeur de base comme individuels (c).

a) L'électeur souverain

La théorie du choix collectif repose sur une conception de la démocratie selon laquelle la souveraineté de l'électeur n'est respectée que si le mécanisme d'agrégation des votes (ou des préférences) n'est pas arbitraire. Dès le début de sa monographie, Arrow propose une classification des méthodes de prise de décision en collectivité, dans laquelle il oppose la démocratie à la dictature, mais aussi à la « convention ».

> *« Dans une dictature idéale il n'y a qu'une volonté qui exerce les choix ; dans une société idéale régie par la convention, il n'y a que la volonté divine ou peut-être une volonté générale exercée par tous les individus, relative aux décisions sociales, si bien que, dans l'un et l'autre cas, aucun conflit de volontés particulières ne peut apparaître. »*[71]

[71]ARROW, Kenneth, *Social Choice and Individual Values*, New Haven, Yale University Press, 1951 (2ème édition révisée, 1961).Traduction française par l'Association de Traduction Economique de l'Université de Montpellier, Paris : Calmann-Levy, 1974, Introduction, p 20.

Ainsi, la décision sociale ne peut être un choix collectif que si les volontés particulières ont pu s'exprimer. Une « démocratie » dans laquelle les décisions seraient déterminées par une « volonté générale » indépendante des volontés particulières n'est pas un régime politique fondé sur le choix collectif. Pour qu'il soit possible de parler de « démocratie » et non pas de « dictature » ou de « société régie par la convention », il faut que le « conflit de volontés particulières » puisse apparaître. Il faut souligner que cette classification est effectuée *a priori* : en qualifiant d' « idéales » la dictature comme la société régie par la convention, Arrow entend bien souligner qu'il ne s'appuie ni sur l'usage historique de ces termes, ni sur des théories empiriques. Le critère permettant de distinguer « dictature » et « démocratie » est en fait purement formel et peut être décrit en termes mathématiques, puisqu'il s'agit ni plus ni moins de la règle de décision mobilisée pour la collectivité.

Dans la mesure où le vote est le paradigme du choix collectif, la recension des « voix » individuelles peut être mise en rapport avec le choix collectif de deux manières au moins :

- Soit, comme semble le dire Arrow, les préférences individuelles constituent la matière même du choix collectif. Celui-ci est ainsi construit au moyen de l'agrégation de celles-là. Lorsque ces matériaux individuels ne sont pas « compatibles » entre eux, il y a impossibilité du choix collectif. Le collectif se confond ici avec l'agrégatif. Il n'y a pas de collectif sans l'agrégation, l'agglomération même, de données individuelles, et le collectif n'est rien de plus que cette agrégation. Agréger un élément, c'est l'ajouter à un corps déjà constitué (par exemple, en Suisse, l'agrégation bourgeoisiale est une procédure au moyen de laquelle des personnes souhaitant être naturalisées peuvent devenirs « bourgeoises » d'une commune). Le processus d'agrégation n'est pas fondateur, c'est une construction continue. Les « éléments » agrégés sont distincts et susceptibles de mener une existence indépendante de la collectivité à laquelle ils s'agrègent.
- Soit le choix collectif est un procédé visant à révéler plutôt qu'à construire la volonté collective. Cette perspective est celle de Rousseau dans le *Contrat Social*, lorsqu'il décrit le vote démocratique.

  "Hors [le] contrat primitif, la voix du plus grand nombre oblige toujours les autres; c'est une suite du contrat même. Mais on demande comment un homme peut être libre et forcé de se conformer à des volontés qui ne sont pas les siennes. Comment les opposants sont-ils libres et soumis à des lois auxquelles ils n'ont pas consenti ?

Je réponds que la question est mal posée. Le citoyen consent à toutes les lois, même à celles qu'on passe malgré lui, et même à celles qui le punissent quand il ose en violer quelqu'une. La volonté constante de tous les membres de l'Etat est la volonté générale : c'est par elle qu'ils sont citoyens et libres. Quand on propose une loi dans l'assemblée du peuple, ce qu'on leur demande n'est pas précisément s'ils approuvent la proposition ou s'ils la rejettent, mais si elle est conforme ou non à la volonté générale qui est la leur : chacun en donnant son suffrage dit son avis là-dessus, et du calcul des voix se tire la déclaration de la volonté générale. *Quand donc l'avis contraire au mien l'emporte, cela ne prouve autre chose sinon que je m'étais trompé, et que ce que j'estimais être la volonté générale ne l'était pas. Si mon avis particulier l'eût emporté, j'aurais fait autre chose que ce que j'avais voulu, c'est alors que je n'aurais pas été libre. »*[72]

Selon Rousseau, le vote n'est pas l'expression d'une préférence subjective, mais un jugement de la conformité d'une option à la volonté générale. Le résultat du vote révèle donc la volonté générale – dans la mesure où le pacte social originaire est respecté - ; autrement dit, celle-ci subsiste indépendamment de l'expression des volontés individuelles. Le choix collectif déterminé par la volonté générale doit être retrouvé et non pas généré par les « voix » des membres de la société. La liberté de l'individu ne réside pas dans le consentement personnel au choix de la société, mais dans la conformité de ce choix à une volonté générale authentique. Mais un écart trop grand entre la voix individuelle et la volonté générale est le symptôme d'un dysfonctionnement du corps politique.

Dans cette perspective rousseauiste, le collectif ne se confond pas avec l'agrégatif. L'agrégation des voix ne vise pas à produire le choix collectif ; elle n'est qu'un moyen de le reconnaître. Le collectif est « déjà là » lorsque l'agrégation a lieu. La volonté générale s'exprime par les voix individuelles, mais ne se résume pas à ces dernières.

[72] ROUSSEAU, Jean-Jacques, *Du Contrat social*, Livre IV, Chapitre III. Nous soulignons.

b) Une approche mécaniste

Arrow conçoit l'agrégation, ou le regroupement des préférences individuelles sur un modèle mécanique. Cette « mécanique » permet de concevoir en amont les individus comme des entités distinctes, séparables et autosuffisantes, que l'on regroupe par la suite. La manière dont il décrit la mécanique du choix social dans le passage suivant fait apparaître cette conception clairement :

> *« Chaque individu est libre de choisir son propre ordre [de préférence], de sorte que la société doive avoir une règle pour construire un classement social basé sur les valeurs individuelles. Une telle règle sera basée sur les classements individuels. […] Nous pouvons imaginer une machine dans laquelle chaque personne introduirait son classement de préférences pour tous les candidats possibles ; un tour de manivelle, et la machine sort un classement de préférences pour tous les candidats possibles. Les urnes et les machines à voter sont en fait une variété particulière de fonctions de choix social. »*[73]

La préférence sociale dépend exclusivement des préférences individuelles : dans le classement social devraient pouvoir se retrouver tous les classements individuels ; le classement social conserve l'intégrité des préférences individuelles. L'expérience de pensée de la « machine » montre le caractère mécanique de l'agrégation ; aucune composante autre que les ordres individuels n'entre dans l'élaboration du choix collectif ; cette expérience montre également que la société n'est pas considérée comme un « sujet » émettant des jugements de valeur au même titre que les individus. La mécanique arrovienne repose sur une description atomiste de la collectivité.

[73] ARROW, Kenneth, „Le principe de rationalité dans les décisions collectives", *Economie Appliquée*, 5 (1952) : 469-484.

c) Les jugements de valeur seraient individuels

Les préférences individuelles sont des jugements de valeur; la préférence sociale peut-elle également être considérée comme un jugement de valeur ? La réponse d'Arrow est d'abord négative :

> *« Je suis entièrement d'accord avec le fait que seuls les individus peuvent avoir des jugements de valeur – et que, par conséquent, l'ordre de préférence construit en utilisant la fonction de choix social (basée sur les préférences individuelles) n'est qu'un moyen de produire des décisions sociales et en aucun cas l'expression d'une quelconque éthique interpersonnelle. »*[74]

Un jugement de valeur semble ne pouvoir être attribué qu'à un individu. Afin de comprendre pourquoi, selon Arrow, les jugements de valeur ne peuvent être qu'individuels, nous nous référons à un article de 1967, où il propose une définition du « système de valeur individuel » définition qui se présente comme une « description complète ». Cette description est en fait une expérience de pensée qui met en scène un individu *complètement isolé*[75]. Ce Robinson n'est limité dans le choix de ses actions que par son environnement physique et ses propres qualités et compétences. Il peut ainsi choisir entre différentes actions mutuellement exclusives. Une action est définie comme la « description complète de toutes les activités qu'un individu entreprend, et deux actions alternatives sont deux descriptions quelconques qui se distinguent de manière pertinente ». Une répartition déterminée du temps de travail et de loisir constitue ainsi une « action ».[76]

La conception du jugement de valeur à laquelle songe Arrow lorsqu'il affirme que les seuls vrais jugements de valeur doivent être individuels est donc la suivante : un jugement de valeur est la règle qui permet à l'individu de déterminer, parmi un choix d'options, celle qui correspond à ce qu'il valorise le plus. Néanmoins, ce qui fait que l'individu détermine cette action plutôt qu'une autre n'est pas explicité par Arrow. Pas plus que ce qui fait que ce n'est que l'individu qui est capable d'une telle activité de juger et de choisir.

Mais les quatre « conditions raisonnables » du choix collectif sont aussi des jugements de valeur. Peut-on dire que les jugements de valeur individuels sont subordonnés à ces jugements de valeur-là ? Ce serait remettre en cause la primauté de l'individu. Arrow répond à cela en montrant que les « conditions raisonnables » du choix collectif peuvent elles-mêmes être

[74]Ibidem.
[75]ARROW. Kenneth, « Values and Collective Decision Making », P. Laslett and W.G. Runciman, eds, *Philosophy, Politics and Society*, Third Series, Oxford: Basil Blackwell, 1967, pp215-232.
[76] Ibidem

conçues comme des jugements de valeur individuels. Il donne l'exemple d'un individu, Primus :

« Supposons que le jugement de valeur fondamental de Primus est que la liberté de choix de chaque individu doit être respectée. »[77]

Primus crée alors un ordre social qui reflète les points de vue des différents membres ; cette combinaison est conforme aux conditions raisonnables du choix collectif, qui constituent l'éthique de Primus en matière de distribution. Ainsi, les jugements de valeur qui permettent de distinguer un choix collectif sont susceptibles de faire partie d'une éthique individuelle.

[77] ARROW, Kenneth, „Le principe de rationalité dans les décisions collectives", *Economie Appliquée*, 5 (1952) : 469-484

2.1.2. La justification épistémique de la primauté de l'évaluation individuelle

Arrow semble donc considérer que l'hypothèse selon laquelle l'individu est la source des évaluations et des préférences est valide. Deux raisons peuvent être avancées pour cela. La première est d'ordre épistémique et est examinée dans cette section : l'individu serait la seule entité connaissable, ou du moins la plus aisée à connaître. La seconde raison est d'ordre éthique est fait l'objet de la section 2.1.3..

A ses débuts, la théorie du choix social hérite d'une base d'information qui ramène la décision à sa source supposée, l'agent individuel. L'unité du phénomène de la décision est garantie par l'unité de l'individu. (a) Par ailleurs, la préférence est attribuée à un individu particulier, étudié dans des conditions réelles. Arrow se démarque en effet explicitement de toute approche qui, bien que reconnaissant la pertinence de l'attribution de jugements, décisions et actions aux individus particuliers dans des conditions sociales et économiques réelles, privilégierait les jugements ou les préférences exprimés dans des conditions idéales. Il oppose ainsi son approche à des doctrines qu'il taxe d' « idéalisme ». (b).

a) L'individu comme entité observable

La primauté de l'individu s'expliquerait d'abord par des raisons épistémiques. Les jugements de valeur, ou les préférences, sont relevés en observant les actions des individus. Ces évaluations individuelles constituent les faits observables qui permettent d'apparenter la méthodologie de l'économie à celle des sciences de la nature.

L'hypothèse individualiste, selon laquelle l'individu est la source de l'évaluation, du choix et de l'action fait de chaque agent une « cause » distincte de l'action étudiée par l'économiste. Cette primauté méthodologique fait donc le pari de la situation de l'origine de l'action « dans » l'agent. Non qu'elle soit nécessairement localisée dans l'agent, mais c'est l'agent que nous rechercherons lorsqu'il s'avèrera nécessaire d'attribuer cette action. Certaines actions sont exécutées directement par l'agent, par son corps : ainsi, lorsque le boulanger pétrit la pâte et enfourne le pain afin de le vendre, il réalise une action économique dont il est la source et le moyen. Mais ces cas sont peu fréquents : le propriétaire d'une boulangerie industrielle produisant une quantité journalière importante de pain fait faire ce travail par des ouvriers et ne réalise pas lui-même l'action économique dont il est la source. La seule observation des corps ne suffit donc pas à attribuer une action ou une volonté à l'agent. La primauté de l'individu n'est donc pas celle de l'individu comme être vivant et comme corps matériel

distinct des autres corps. La primauté de l'individu semble donc moins liée aux données de l'observation qu'elle le paraissait de prime abord.

b) Conditions réelles VS conditions idéales des évaluations individuelles

Le traitement que réserve Arrow aux doctrines qu'il qualifie d' « idéalistes » éclaire un peu mieux l'idée qu'il se fait des conditions d'attribution de la préférence aux individus. Dans le chapitre VI de sa monographie sur le choix collectif, Arrow reformule en effet dans la terminologie de la théorie du choix social une tendance philosophique qu'il qualifie d' « idéaliste », et qu'il attribue à Rousseau, Kant et Thomas Green. Ce terme ne désigne pas ici l'idéalisme ontologique, ou celui de la théorie de la connaissance. Il signifie que les préférences pertinentes pour le choix social sont des préférences exprimées comme si l'agent se trouvait dans des conditions idéales et ne subissait pas l'influence de causes externes défavorables au consensus.

Arrow souligne la différence entre sa conception de l'individu et la conception «idéaliste ». Cette dernière fait de l'individu une entité substantielle : en effet, le dualisme renvoie à une psychologie même minimale ou encore à l'idée de « facultés », de débats intérieurs et de motivations. En revanche, la théorie du choix social n'appréhende l'individu qu'à travers l'action, la manifestation de ses préférences ; elle ne le connaît que de l'extérieur. S'il s'avérait que les individus pouvaient avoir deux ordres de préférence, la théorie du choix social ne connaîtrait que celui qui a été révélé par les comportements et les dires de la personne concernée, dans les conditions actuelles d'expression de ces préférences.

Dans la terminologie de la théorie du choix social, l'idéalisme stipule que l'individu peut exprimer non pas *une* préférence, d'un classement des états sociaux possibles, mais deux, l'une dans les conditions réelles de son existence, l'autre dans des conditions idéales.

> *« l'idée fondamentale de [l'école idéaliste] est de distinguer la volonté individuelle telle qu'elle est à tout moment sous l'effet de causes externes variables, de la volonté générale qui est inhérente à tout et, en tout, identique à elle-même. ».*[78]

L'accent est mis sur la variabilité de la volonté individuelle, opposée à la stabilité de la volonté générale et des institutions qui en découlent. Il n'y a pas de relation nécessaire entre volonté individuelle particulière et volonté générale, qui ne « coïncideront que par hasard ». Dans la terminologie d'Arrow, la volonté générale correspond à l'ordre de préférence

[78] *Ibidem*, VI, 3, p 151.

conforme aux autres ordres de préférences individuels et à une certaine idée de ce que devrait être la société.

> *« La doctrine idéaliste peut alors se résumer en disant que tout individu a deux relations d'ordre. L'une commande ses actes quotidiens, l'autre interviendrait sous certaines conditions idéales. C'est cette dernière qui doit être considérée comme pertinente en matière de choix collectifs. »*[79]

Arrow traduit la « doctrine idéaliste » dans le langage de l'économie du bien-être : le «royaume des fins » devient la « collectivité dotée d'une fonction de bien-être collectif satisfaisante. Arrow refuse donc de s'affilier à la thèse selon laquelle nous ne devrions retenir comme base d'information les seules préférences individuelles exprimées dans des conditions hypothétiques idéales, ces conditions étant garantes d'un consensus et de l'évitement du résultat d'impossibilité. Mais avant de nous tourner vers la justification normative de la primauté de l'individu, il nous faut examiner si Arrow peut légitimement qualifier d' « idéaliste » les doctrines de Rousseau et de Kant.

[79] *Ibidem*, p 153

L' « idéalisme » de Rousseau

Dans le *Contrat Social,* la distinction entre volonté générale et volonté particulière s'articule avec la distinction entre l'individu comme citoyen et l'individu comme homme. En tant que citoyen, l'individu est un membre constitutif du corps politique uni, indivisible et souverain. La volonté générale du peuple uni et indivisible (s'il ne l'est pas, c'est qu'il n'est plus un « peuple » à proprement parler, mais un simple agrégat) est telle que générée par le pacte social. La lecture que fait Arrow de Rousseau rend effectivement compte de ce dualisme. En revanche, elle est approximative sur les conditions d'expression de la volonté générale, d'une part, et sur la source de cette volonté générale, d'autre part.

Arrow affirme que la volonté générale, en tant que relation d'ordre individuel, interviendrait « sous certaines conditions idéales » et se traduirait par l'unanimité du choix social. En fait, bien que la volonté générale soit conçue dans un premier temps comme la volonté du peuple placé dans la situation hypothétique du contrat originaire, et puisse par conséquent être qualifiée de volonté du citoyen dans des « conditions idéales », la volonté générale peut s'exprimer dans des conditions non hypothétiques et imparfaites. Le passage de conditions parfaitement favorables à l'expression de la volonté générale à des conditions très défavorables n'est pas un passage de l'hypothèse à la réalité : il s'agit d'un lent processus de dégénérescence du corps politique, qui se démembre progressivement. Il existe donc chez Rousseau des stades intermédiaires entre la pleine santé et la mort du corps politique, où l'unanimité est remplacée par la majorité, où les voix des intérêts particuliers émergent dans le discours politique mais laissent encore la place à l'expression de la volonté générale : lors des suffrages, la majorité suffit à attester de l'expression de la volonté générale, l'unanimité n'est pas nécessaire, bien que préférable (encore que l'unanimité puisse parfois témoigner de la mort du corps politique et de l'adhésion totale à la servitude).

> *"Mais quand le noeud social commence à se relâcher et l'Etat à s'affaiblir : quand les intérêts particuliers commencent à se faire sentir et les petites sociétés à influer sur la grande, l'intérêt commun s'altère et trouve des opposants : l'unanimité ne règne plus dans les voix, la volonté générale n'est plus la volonté de tous, il s'élève des contradictions, des débats, et le meilleur avis ne passe point sans disputes. »*[80]

[80] ROUSSEAU, Jean-Jacques, *Du Contrat Social,* Livre IV, Chapitre I.

« A l'autre extrémité du cercle l'unanimité revient. C'est quand les citoyens tombés dans la servitude n'ont plus ni liberté ni volonté. Alors la crainte et la flatterie changent en acclamation les suffrages ; on ne délibère plus, on adore ou l'on maudit. »[81]

L'unanimité n'atteste donc pas de la présence de la volonté générale. Les conditions d'expression de la volonté générale sont idéales au sens où elles sont satisfaisantes d'un point de vue normatif (elles permettent au corps politique d'être tel qu'il devrait être), mais ne sont pas idéales au sens où elles seraient irréalisables. Ces conditions sont une étape de la vie du corps politique, celle de sa pleine santé, et l'art du législateur consiste à identifier et rétablir ces conditions. Parmi ces conditions, on peut citer la réunion fréquente de l'assemblée souveraine, le fait d'éviter que le pouvoir exécutif ne s'accapare les prérogatives du peuple souverain et législateur, ou encore la religion civile qui favorise l'attachement du citoyen à la République.

Lorsqu'Arrow affirme que la doctrine idéaliste, comme celle de Rousseau, attribue à chaque individu deux relations d'ordre, il semble oublier que la volonté générale subsiste indépendamment des individus. Le pacte social, fondement théorique de l'Etat, opère la mutation d'une collection de volontés individuelles en une volonté unique, commune et distincte de ces volontés individuelles : la volonté générale. Cette volonté est la volonté d'un individu collectif, le Peuple. Contrairement à ce que laisse entendre Arrow lorsqu'il dit que la doctrine de Rousseau suppose que tout individu a deux ordres de préférence (l'un conforme à son intérêt particulier, l'autre à l'intérêt général), la volonté générale n'est pas l'agrégation des volontés individuelles soucieuses de l'intérêt général. Elle n'est pas la résultante d'un consensus entre des personnes différentes se trouvant dans des conditions favorables à l'intérêt général. Elle est constitutive du corps politique et subsiste en tant que telle tant que le corps politique vit – ou survit - . Elle se révèle dans les votes et les décisions des citoyens, et ceux-ci l'éprouvent dans leur cœur même lorsqu'ils lui désobéissent. Mais ils ne sont pas en eux-mêmes la source de cette volonté générale à laquelle ils adhèrent en tant que citoyens. S'ils l'étaient, la volonté générale disparaîtrait aussitôt que le choix collectif se ferait dans des conditions non-idéales ; elle serait réduite à l'état d'hypothèse ayant valeur de norme pour tester des configurations sociales et politiques déterminées.

Cette discussion critique autour de la lecture de Rousseau par Arrow ne remet cependant pas en question l'enjeu de la critique de l' « idéalisme » en philosophie politique, à savoir l'acceptation par le théoricien de l'hétérogénéité des valeurs. Elle renforce même la

[81] *Ibidem*, Chapitre II.

distinction entre la perspective de l'agrégation, caractéristique de la théorie du choix social, et celle de l'association, propre aux tenants du pacte social. En revanche, elle désamorce l'une des critiques possibles à l'égard de ces derniers, à savoir celle selon laquelle ils ne se préoccuperaient que du choix collectif dans des conditions idéales. Bien au contraire, le *Contrat Social* témoigne d'une préoccupation constante des conditions de *réalisation* de la volonté générale.

L' « idéalisme » de Kant

Arrow reformule également les impératifs kantiens dans le langage de la théorie du choix social.

Dans les *Fondements de la Métaphysique des Mœurs,* Kant distingue trois impératifs, l'impératif technique, l'impératif pragmatique et l'impératif catégorique. Ces impératifs déterminent la volonté des agents. Le contenu des impératifs technique et pragmatique est conditionné par la fin visée par l'action qu'ils déterminent : l'impératif technique détermine l'action la plus appropriée à la réalisation d'une fin, quelle qu'elle soit (il peut s'agir de la fin de l'empoisonneur comme de celle du médecin) ; l'impératif pragmatique détermine l'action la plus appropriée à la réalisation d'une fin supérieure à toute autre fin, le bonheur.

Dans le texte d'Arrow, l'impératif technique correspond à la rationalité instrumentale : la fin étant donnée, il est déterminé par l' « environnement ». Arrow assimile ici la fin et la préférence. Il suppose en effet que la préférence exprimée par un individu est fonction de l'objectif qu'il souhaite atteindre, de la fin qu'il se donne. Lorsque cet objectif est déterminé, la règle d'action ne dépend plus que de l'environnement, c'est-à-dire de l'ensemble des circonstances de l'action : le milieu naturel, les données économiques et sociales, le cadre institutionnel effectif, la psychologie de l'agent, les actions passées de l'agent et des autres agents, etc.

L'impératif pragmatique correspond chez Arrow aux « classements individuels des états sociaux », c'est-à-dire aux préférences exprimées par les individus en fonction de leurs goûts ou de leurs valeurs.

Arrow affirme ensuite que l'impératif catégorique « correspond à [son] concept d'ordre social ». Dans une société, ce concept correspond à un classement unanime de tous les états sociaux. Pour ce faire, nous devons concevoir l'ensemble de tous les êtres raisonnables comme une collectivité ; les commandements de la morale, tels que l'interdiction de tuer, résulteraient, ou pourraient résulter, des jugements de valeur et des choix de cette « collectivité ». Ces jugements de valeur ont la forme d'un « ordre », d'un classement, dérivé

de classements individuels identiques. Ils sont identiques parce qu'ils sont tels que les exprimerait une personne qui ne connaîtrait pas les désirs et les inclinations propres au monde sensible, désirs et inclinations qui seraient donc les facteurs de l'hétérogénéité sociale.

A y regarder de plus près, on se demande s'il ne s'agit pas d'une mécompréhension du caractère catégorique dudit impératif. En effet, si l'impératif catégorique était un classement, même unanime, d'états sociaux, alors les propositions suivantes : « il vaut mieux tuer une personne qu'un million de personnes », ou encore « il vaut mieux tuer Hitler que Gandhi », pourraient faire partie de ce classement. Or, l'impératif catégorique n'admet pas de telles évaluations comparatives : il décrète, de manière inconditionnée, qu'il ne faut pas tuer, et la question du nombre de tués n'entre pas en ligne de compte. La vie humaine n'a pas de prix. L'idée de « classement » est étrangère à celle de commandement moral universel et nécessaire.

Ce bref développement nous aura enseigné que la représentation de jugements de valeur sous la forme de classements d'états sociaux présuppose une perspective conséquentialiste – au moins au sens large du terme –. En philosophie morale, elle est incompatible avec une approche déontologique, qui *ne peut pas* intégrer des comparaisons d'options autres que celle prônée par le principe défendu.

Ainsi, ce qui distingue la volonté déterminée par l'impératif catégorique de la volonté individuelle exprimée dans des conditions réelles n'est pas tant le caractère idéal des conditions que la forme même de l'expression de cette volonté. L'individu retenu par la théorie du choix social n'est pas le concept d'un être raisonnable auquel des jugements de valeur inconditionnés peuvent être rattachés. Il s'agit d'un individu concret, du « Moi empirique » kantien. Cet individu appartient au monde des phénomènes et peut prétendre au titre d' « entité observable ».

2.1.3. La justification normative de la primauté de l'évaluation individuelle : souveraineté de l'agent et pluralisme

La primauté de l'individu se justifie en fait plus aisément d'un point de vue normatif ou éthique que d'un point de vue épistémique. De ce point de vue, l'individu peut de fait ne pas déterminer ses préférences, mais dans il serait préférable, voire indispensable qu'il le fasse. Cette norme correspond à la thèse de souveraineté de l'agent, qu'Arrow désigne comme

« autonomie ». L'autonomie correspond ici à la faculté de déterminer par soi-même ses préférences, en tant qu'agent, et ce en dépit des interférences des autres agents. La conséquence de la reconnaissance d'une telle souveraineté de l'agent est la nécessité de reconnaître la possibilité de motivations multiples. Cela conduit Arrow à récuser tant l'hypothèse de la maximisation de l'intérêt personnel que l'homogénéité des préférences pour les biens premiers dans la position originelle rawlsienne.

a) Critique arrovienne de l'hypothèse de maximisation de l'intérêt personnel

Par ailleurs, Arrow distingue conceptuellement les « préférences particulières » (*tastes*) des « préférences fondamentales » (*values*). Les premières sont orientées par la consommation individuelle et la satisfaction des goûts personnels. Les secondes reflètent des conceptions morales ou éthiques plus larges. Arrow estime que ce sont ces dernières qui devraient compter dans le choix collectif, même si cela est impossible à déterminer sur le plan formel. Il récuse donc ce qu'il appelle « les hypothèses individualistes »[82], selon lesquelles la consommation individuelle déterminerait exclusivement les préférences de l'agent y compris dans le contexte du choix social (le terme « individualiste » a ici une définition beaucoup plus restrictive, puisqu'il est à peu près synonyme d' « égoïste »).

On peut comprendre ce refus de faire sienne une conception réductrice des motivations humaines à l'aune de la conception de l'autonomie que défend Arrow. En effet, la faculté de déterminer par soi-même ses préférences n'a d'intérêt et de valeur que dans la mesure où le jugement de valeur ou l'appréciation qui s'exprime dans la préférence ou le vote dépend de l'agent, de ses valeurs propres ou de celles auquel il s'identifie. Formellement, l'hypothèse individualiste est insatisfaisante aussi, car elle viole la condition de non-restriction du domaine des préférences ; de plus, elle ne lève pas le résultat d'impossibilité.

[82] *Ibidem*, chapitre V.

b) Critique des préférences rawlsiennes pour les biens premiers

Dans la *Théorie de la Justice*, John Rawls fait l'hypothèse selon laquelle des agents placés dans une situation de choix hypothétique, la position originelle, auraient des préférences identiques pour des biens premiers, qui seraient des moyens polyvalents de réaliser différentes conceptions du bien ou différents modes de vie[83]. Cette position originelle est caractérisée par le voile d'ignorance, qui dissimule aux agents les informations relatives à leurs situations particulières. La position originelle est au centre de l'argumentation de Rawls en faveur de ses principes de justice.

Arrow se préoccupe des questions de choix social plutôt que des institutions justes. Il ne critique donc pas la démarche de Rawls dans l'ensemble, mais manifeste un désaccord avec l'homogénéité des préférences pour les biens premiers. Dans ses notes critiques sur Rawls[84], Arrow revendique une appréhension authentique des différences de préférences.

> *« Dans la mesure où les individus sont vraiment des individus, chacun étant une fin autonome en soi, alors ils doivent être quelque peu mystérieux et inaccessibles pour les autres. Une règle entièrement acceptable pour tous ne peut exister. Il doit y avoir, du moins à ce qu'il me semble maintenant, la possibilité d'un conflit insoluble, qui peut se manifester logiquement sous la forme de paradoxes dans le processus de décision sociale. »*

Pour Arrow, l'élaboration d'une théorie sociale et politique ne peut faire l'économie de la diversité des expériences individuelles, fut-ce au prix de l'aboutissement de la théorie elle-même.

Les jugements de valeur individuels ont ainsi la priorité dans l'élaboration du choix collectif tel que le conçoit Arrow. Cette priorité est justifiée par des arguments d'ordre épistémique et d'ordre éthique. Mais le statut des jugements de valeur qui régissent les interactions entre individus, ainsi que celui de l'évaluation sociale qui conclut le regroupement des évaluations individuelles, demeure problématique.

---

[83] RAWLS, John, *A Theory of Justice*, The Belknap Press of Harvard University Press, 1971; traduction française par Catherine Audard sur un texte révisé par l'auteur en 1975: *Théorie de la Justice*, Paris, Editions du Seuil, 1987, 1997. Préface de la traduction française par John Rawls. § 15.

[84] ARROW, Kenneth, « Some Ordinal-Utilitarian Notes on Rawls's Theory of Justice », *Journal of Philosophy*, 70, 1973, p 245-263.Traduction libre..

2.2. Une primauté de l'évaluation individuelle à nuancer

Nous avons vu, dans la section 2.1. qui précède, que l'individu est considéré par Arrow comme le support des préférences qui constitue la base d'information du choix social, et ce pour des raisons épistémiques et normatives.

Toutefois, dans les commentaires même qu'Arrow fait de son propre travail, cet individualisme est nuancé par plusieurs considérations. D'une part les collectivités peuvent aussi être les supports de préférences et de jugements (2.2.1.). D'autre part, les motivations à l'œuvre ne sont pas orientées exclusivement en faveur de la réalisation d'intérêts individuels (2.2.2.). Enfin, si le cadre formel accueille les préférences individuelles comme des données, Arrow ne nie pas la réalité des relations sociales (2.2.3.).

2.2.1. Le sens et la portée de la rationalité collective

En promouvant le concept de préférence collective, Kenneth Arrow défend l'idée selon laquelle la collectivité peut être dite rationnelle au même titre que l'individu. Elle l'est en effet dans la mesure où elle peut exprimer une préférence répondant à certaines conditions de cohérence. L'enjeu de cette thèse est important : cela signifie en effet que l'action collective n'est pas le produit d'une conjonction accidentelle d'intérêts substantiels ; il est possible de produire une décision collective déterminée par un classement d'options, et peut-être par un classement d'option lui-même déterminé par la poursuite rationnelle de fins collectives.

En défendant donc l'idée de rationalité collective, Arrow s'oppose d'une part aux travaux de James Buchanan et Gordon Tullock, représentants majeurs de la théorie des choix publics[85], et d'autre part à l'initiateur de la construction de fonctions de bien-être collectif, Abram Bergson.

[85] Dans l'*Idée de Justice,* Amartya Sen présente la théorie des choix publics comme "d'une certain manière, le rival conservateur de la théorie du choix social" (*The Idea of Justice,* Allen Lane, Penguin Books, 2009, ch. 14).

a) *L'action collective n'est pas qu'un composé d'actions individuelles : théorie du choix social VS théorie des choix publics*

James Buchanan et Gordon Tullock sont les initiateurs de la théorie des choix publics. Cette branche de l'économie est parfois à tort confondue avec la théorie du choix social. La théorie du choix social a pour objet l'étude des fonctions de bien-être social et des règles permettant d'agréger des préférences ou d'autres attributs individuels comme les utilités et les capabilités. Elle ne fait pas d'hypothèse sur les motivations à l'œuvre dans ce processus. Elle vise simplement à déterminer un choix social représentatif des préférences individuelles, quelles qu'elles soient. En revanche, la théorie des choix publics repose sur des hypothèses substantielles concernant les motivations des individus : ceux-ci sont principalement mus par la maximisation de leurs intérêts privés. L'objectif des théoriciens du choix public est donc d'expliquer la politique à l'aide de ces hypothèses, afin de remédier par exemple à l'inefficacité publique.

L'une des divergences fondamentales entre théorie du choix social et théorie des choix publics est la suivante : tandis que les premiers estiment que cela fait sens de parler de rationalité collective et d'attribuer des préférences et des choix à une collectivité, les seconds considèrent que la rationalité est une propriété réservée à l'individu. L'action collective est une production artificielle, composée d'actions individuelles affirmées comme réelles :

> *« L'action collective est conçue comme l'action des individus qui choisissent d'accomplir leurs objectifs collectivement plutôt qu'individuellement, et le gouvernement est conçu comme rien de plus que l'ensemble des processus, la machine, qui permet à une telle action collective d'avoir lieu. Cette approche fait de l'Etat une construction humaine, un artefact. »*[86]

Pour Arrow, la rationalité caractérise un choix et non un individu. Les individus expriment leurs choix sous la forme de préférences, mais tout choix n'est pas nécessairement individuel[87]. Ainsi, l'Etat ne sera pas qu'un simple dispositif d'agencement des actions individuelles, mais un système pouvant fonctionner de manière autonome.

[86] cf. BUCHANAN, J. M., TULLOCK, Gordon, *The Calculus of Consent*, Ann Arbor, University of Michigan Press, 1962, ch.1, p 13.

[87] cf. ARROW, Kenneth, *Social Choice and Individual Values*, New Haven: Yale University Press, 2ème édition révisée, 1963. Traduction française par l'Association de Traduction Economique de l'Université de Montpellier, Paris : Calmann-Levy, 1974, p 214-217.

> *"La rationalité collective dans le mécanisme de choix collectif ne constitue pas alors simplement un transfert indu de l'individu à la société, mais une attribution importante d'un système authentiquement démocratique, capable de s'adapter pleinement aux situations qui évoluent. »*[88]

Autrement dit, la rationalité est l'attribut de préférences, de jugements ou de choix énoncés et réalisés, et non pas d'entités telles que les individus ou la société.

Si le choix collectif est rationnel, il suit immédiatement l'évaluation collective, laquelle est construite en agrégeant les évaluations individuelles. Or, Buchanan et Tullock nient la possibilité de jugements de valeur et de préférences qui ne puissent pas être attribués à un individu. La société, selon eux, ne peut être un agent, et encore moins un agent rationnel. En fait, Buchanan et Tullock conçoivent l'énonciation d'un jugement de valeur sous une forme causale : s'il y a jugement de valeur, alors il doit y avoir un individu qui le produit ; la fonction de bien-être collectif, affirment-ils en substance, ne fait que produire la résurgence d'une conception organiciste de la société que les démocraties libérales devraient abandonner pour des raisons méthodologiques.

A l'inverse de la théorie des choix publics, la théorie du choix social n'accorde pas à l'individu le privilège de la rationalité. Elle ne le peut pas, parce que son objet d'étude est précisément le « choix social », c'est-à-dire un choix déterminé par une évaluation collective à la fois rationnelle et représentative des évaluations individuelles. En privant l'agent de sa substance, le formalisme arrovien ouvre la possibilité de penser la valeur de l'action sociale en tant que telle. Ce sont les jugements de valeur, les propositions, qui sont les données de la théorie du choix social, et non pas les individus. Pour Arrow, reste à déterminer s'il est sensé de parler de « jugements de valeur collectifs », et si les jugements de valeur collectifs correspondent à des faits.

[88] *Ibidem.*

b) Jugements de bien-être collectif et choix social : il y a un sens à parler d'action collective

Arrow est également en désaccord avec Abram Bergson au sujet du statut des jugements de valeur collectif[89]. A cette époque, l'économiste américain Abram Bergson est connu comme l'inventeur des fonctions de bien-être collectif[90].

Dans sa revue critique du théorème d'Arrow[91], Abram Bergson souligne que la tâche de l'économie du bien-être est de reformuler les jugements de valeur des individus. Ce sont les individus seuls qui classent les états sociaux et produisent donc des jugements de valeur pour la collectivité. L'économiste du bien-être doit les informer sur la teneur exacte des aspects économiques des différentes options, et vérifier l'adéquation de leurs choix à leurs valeurs. Abram Bergson considère donc que la donnée première et nécessaire de l'évaluation sociale est une attente individuelle. D'après cette conception, la donnée première, irréductible, serait un individu. Cet individu serait à l'origine de jugements de valeur qu'il faut ensuite simplement clarifier, reformuler et traduire en termes techniques.

En revanche, un jugement de valeur collectif produit par une mécanique d'agrégation de valeurs individuelles – le paradigme étant le vote – ne relève pas de l'économie du bien-être, mais de la théorie politique. Abram Bergson ne nie pas l'importance du résultat d'Arrow, mais il nie la portée de ce résultat pour l'économie du bien-être, qui ne peut avoir affaire qu'aux jugements de valeur des individus, ou éventuellement de groupes d'individus s'il est avéré que ceux-ci partagent les mêmes valeurs. Le problème est donc le suivant : y a-t-il un sens à parler de jugements de valeur collectif en dehors du cas spécifique des résultats d'un vote ?

La discussion à propos des critiques d'A. Bergson par K. Arrow montre très bien qu'Arrow se refuse à réduire l'évaluation sociale aux seules évaluations individuelles. Cette position est d'ailleurs fragile, car il se trouve alors dans un environnement intellectuel qui dénie toute réalité à l'entité « Société », et lui-même tient à l'assise individuelle des jugements et des actions pour la collectivité. Sa réponse à A. Bergson consiste donc à partir non de la source des jugements de valeur collectifs, mais de leurs effets, à savoir l'action collective.

> *"Lorsque Bergson essaie de découvrir les valeurs collectives dans les jugements que les individus portent sur le bien-être, je préfère situer ces valeurs dans les actions que la collectivité arrête par*

[89] *Ibidem*, p 192-193.
[90] Il est par ailleurs connu pour ses travaux sur l'économie soviétique
[91] BERGSON Abram, « On the Concept of Social Welfare », *Quarterly Journal of Economics*, 1954, 68, pp 233-252.

l'intermédiaire de ses règles de prise de décision collective. Cette position est une généralisation naturelle de la conception ordinaliste : de même qu'elle identifie, pour l'individu, valeurs et choix, de même je considère que valeurs collectives ne veut rien dire de plus que choix collectifs. »[92]

Arrow refuse de réduire l'évaluation sociale aux évaluations individuelles. De cette manière, les jugements de valeur collectifs cessent d'être subordonnés aux jugements de valeur individuels. En effet, il est possible de les connaître non par le biais des jugements individuels, mais par celui de l'action collective. Le positivisme bien compris ne peut pas plus s'accommoder de l'entité « société » que de l'entité « individu »[93] : ce qui se « voit », ce n'est pas l'individu qui juge, ce sont des énoncés exprimés et des actions qui peuvent aussi bien être le fait d'individus que de collectivités.

[92] ARROW, Kenneth, *Social Choice and Individual Values*, New Haven: Yale University Press, 2ème édition révisée, 1963. Traduction française par l'Association de Traduction Economique de l'Université de Montpellier, Paris : Calmann-Levy, 1974, p 193.

[93] On remarquera d'ailleurs que dans la conclusion du même article sur le théorème d'Arrow, A. Bergson nous invite à prendre nos distances avec le positivisme : il est vrai que, si l'on veut « sauver l'individu » qui constituait, de prime abord, la base de l'évaluation sociale d'inspiration utilitariste, on ne peut s'en tenir aux seuls faits.

2.2.2. Evaluations individuelles et attitudes sociales

Arrow a souligné avec force que la distinction entre préférences particulières (*tastes)* et préférences fondamentales (*values*) n'était pas pertinente dans le cadre conceptuel de la théorie du choix social. Mais si le choix collectif vérifie les hypothèses individualistes, le cadre conceptuel de la théorie du choix social change de sens. Si donc seules les préférences particulières comptent – hypothèse selon laquelle chaque membre de la société poursuit exclusivement son propre intérêt, c'est-à-dire la maximisation de sa consommation de biens – les conditions « raisonnables » du choix collectif requièrent une interprétation quelque peu différente. Ces conditions « raisonnables » sont les axiomes que doit respecter la règle de choix social permettant de dériver une préférence collective de l'agrégation des préférences individuelles dans le théorème d'impossibilité d'Arrow[94]. L'expression formelle ne doit pas faire oublier que ces conditions sont ce qui fait qu'une préférence « collective » ainsi dérivée est bien un choix collectif.

> *« Nous devons bien entendu reconnaître que la signification des conditions 2-5 [les conditions raisonnables du choix collectif] a changé. Les raisonnements antérieurs relatifs à leur validité supposaient que les ordres individuels représentent les préférences fondamentales plutôt que les préférences particulières. »*[95]

Examinons donc à présent les cinq conditions ou axiome du théorème d'impossibilité du choix collectif arrovien et leur interprétation lorsque les classements de préférences représentent des valeurs fondamentales plutôt que des préférences particulières.

La condition 1, de non-restriction du domaine des préférences, n'est pas respectée dans l'hypothèse où seules les préférences particulières comptent.

La condition 2, selon laquelle les préférences collectives varient dans le même sens que les préférences individuelles, sera couplée avec la condition 4 dans l'édition de 1963. La condition 4 énonce la souveraineté du citoyen. Toutes deux se fondent dans la condition dite de Pareto. Dans le cas où seules les préférences fondamentales sont prises en compte pour le choix collectif, cette condition garantit la souveraineté des citoyens. Elle manifeste l'importance de la détermination des valeurs sociales par les membres de la société seule. Dans le cas où ce sont les préférences particulières qui sont comptabilisées, on parlera plutôt de « souveraineté du consommateur ».

---

[94] Cf. Annexe 3.

[95] ARROW, Kenneth, *Social Choice and Individual Values*, New Haven: Yale University Press, 2ème édition révisée, 1963. Traduction française par l'Association de Traduction Economique de l'Université de Montpellier, Paris : Calmann-Levy, 1974, p 138.

La condition 3 stipule l'indépendance des alternatives non pertinentes. Les préférences individuelles pour des états sociaux impossibles (par exemple, le vote pour un candidat mort ou imaginaire) ne sont pas comptabilisées. De même, la règle de Borda ne fonctionne pas ici[96].

> *« La condition d'indépendance du choix à l'égard des situations extérieures implique, dans une acception plus large, que toutes les procédures de choix collectif soient des votes. »*[97]

Dans le cas de préférences particulières, Arrow remarque que l'attachement à cette condition perd de son importance. Il donne l'exemple d'une distribution de pain et de vin ; on attribue plus de pain à ceux qui préfèrent les pains, plus de vin à ceux qui préfèrent le vin. Si par la suite le stock de vin est détruit, ceux qui avaient choisi le vin sont-ils fondés à demander davantage de pain par compensation ? Lorsque nous comparons des biens de consommation, la position relative des uns par rapport aux autres peut importer.

> *« Supposons qu'il n'existe que deux biens, du pain et du vin ; une répartition que tous jugent équitable est arrêtée. Ceux qui aiment le vin obtiennent plus de vin et moins de pain que les individus qui ne l'aiment pas. Si l'on fait maintenant l'hypothèse que le stock de vin est détruit, les amateurs de vin sont-ils fondés, de ce fait, à demander plus que la part équitable de pain qui doit leur revenir ? »*[98]

Si la condition d'indépendance est respectée, il faut recalculer la répartition sans tenir compte de la disparition des stocks de vin. Mais on peut opposer à cela que les amateurs de vin se sentent lésés et exigeront donc une compensation sous la forme d'un supplément de pain. Une telle considération fait entrer une alternative non pertinente, le vin (qui ne fait plus partie des choix possibles) dans le calcul de la répartition choisie collectivement. Néanmoins, comme il s'agit de biens de consommation et de goûts, la notion d'intensité des plaisirs bannie de l'ordinalisme d'Arrow retrouve de sa pertinence. Mais les alternatives non pertinentes sont exclues seulement parce qu'elles sont en conflit avec la réalité. Dans le cas où le choix collectif est dérivé de préférences fondamentales, elles sont exclues parce qu'elles rendent arbitraire le résultat du scrutin, posant ainsi un problème de légitimité.

La condition 5, condition de non-dictature, exclut des choix collectifs les choix déterminés par les préférences d'un seul membre de la société. S'il s'agit de préférences fondamentales, cela signifie que le dictateur impose l'état social qu'il juge le meilleur. S'il s'agit de préférences particulières, cela implique que le dictateur impose à la collectivité la satisfaction de ses goûts, et la dictature s'apparente ici davantage à la tyrannie. Dans le premier cas, la

[96] Cf. Introduction, « Un peu d'Histoire », p 17 de ce travail.
[97] *Ibidem*, p 64.
[98] *Ibidem*, p 138.

condition de non-dictature est justifiée par le choix d'une conception démocratique de la vie politique et des décisions sociales ; dans le second cas, l'argument de la préservation de soi devrait suffire à asseoir cette condition.

Ainsi, si les conditions du choix collectifs semblent aussi raisonnables pour les préférences particulières que pour les préférences fondamentales, c'est pour des raisons différentes. Au moment où il écrit sa monographie, Arrow semble penser que la solution du problème du choix collectif se trouve dans les « attitudes sociales » des individus. Les attitudes peuvent être définies comme des comportements de choix déterminés par des valeurs similaires, voire identiques. Arrow essaie ainsi de retrouver des valeurs susceptibles d'être partagées par des individus tout à fait différents.

> *« Quelques valeurs peuvent donner naissance à une telle similitude d'attitudes collectives : les aspirations à la liberté, à la puissance de la nation et à l'égalité. […] puisque les moyens d'accroître la durée de vie sont, dans une large mesure, collectifs, il existe une forte pression à la conformité des attitudes, dans ce domaine particulier. »*[99]

Le fait qu'Arrow éprouve le besoin de mettre en avant la notion d' « attitude » montre le caractère quelque peu insatisfaisant du mot « préférence » pour désigner les préférences fondamentales ou valeurs.

[99] *Ibidem*, p 140.

2.2.3. Le caractère artificiel de la factorisation de la société en composantes individuelles

Dans les articles postérieurs à l'exposé de son théorème, Arrow soulignera le caractère artificiel de la « factorisation » de la société en composantes individuelles : il ne faut pas voir dans le cadre conceptuel du choix social la représentation exacte de la réalité sociale. Arrow avait déjà remarqué que le désir individuel de prolonger la durée de vie requérait des moyens collectifs.

> *« Le besoin de discuter des valeurs publiques est causé par le fait fondamental que chaque action significative implique la participation commune de multiples individus. Même la décision individuelle apparemment la plus simple implique la participation d'une société entière. »*[100]

Arrow donne également l'exemple de la propriété privée, création sociale avant d'être le fruit d'un processus d'acquisition individuelle. La propriété n'est-elle pas garantie par la structure sociale et les institutions ?

Mais les préférences individuelles ne sont-elles pas premières dans la construction de la vie sociale ? Comment Arrow peut-il affirmer cela après avoir construit une théorie basée sur les jugements de valeur individuels ? En quel sens peut-on parler de la « factorisation de l'action sociale en une séquence d'actions individuelles » ?

> *« Bien sûr, je ne voudrais pas nier qu'une telle factorisation ait lieu, mais je voudrais bel et bien souligner que la partition d'une action sociale en composantes individuelles, et l'attribution de la responsabilité aux individus correspondant à cette partition, n'est* pas *une donnée. La factorisation dans n'importe quel contexte est plutôt elle-même le résultat d'une politique sociale et par conséquent elle est issue de valeurs sociales plus anciennes et logiquement plus primitives. »*[101]

Arrow distingue en fait deux niveaux de lecture de l'action sociale ; celui de la théorie du choix social fait des jugements de valeur et des actions individuelles les fondations de la vie collective ; un autre niveau prend en compte des éléments temporels (« valeurs sociales plus anciennes ») ou du moins l'idée d'une évolution, d'une dimension historique : ce second niveau de lecture ne considère pas le choix collectif et ses conditions comme des faits donnés (inhérents à la nature humaine ?) ; au contraire, le choix collectif est le fruit d'un processus et présuppose certaines valeurs communes, indépendamment des préférences individuelles.

---

[100] ARROW. Kenneth, « Values and Collective Decision Making », P. Laslett and W.G. Runciman, eds, *Philosophy, Politics and Society*, Third Series, Oxford: Basil Blackwell, 1967, pp215-232.

[101] ARROW. Kenneth, « Values and Collective Decision Making », P. Laslett and W.G. Runciman, eds, *Philosophy, Politics and Society*, Third Series, Oxford: Basil Blackwell, 1967, pp215-232.

L'individu d'Arrow peut se définit alors dans un environnement ; il ne se distingue même pas nécessairement de cet environnement : en effet, à quel moment puis-je dire que telle ou telle qualité est mienne, indépendamment de ma situation ? Y a-t-il une identité personnelle donnée qui me soit propre ?

> *« Je présume que nous pouvons dire de manière sensée que je suis en train d'écrire cet article – et non pas quelqu'un d'autre –. Mais bien sûr même ici l'action est prise dans les interconnexions sociales. Je suis dans une conférence organisée par d'autres, où j'utilise des mots partagés communément par une culture, où j'exprime des idées utilisant une gamme étendue de concepts forgés par d'autres, et qui constituent mon éducation. »*[102]

Distinguer ce qui relève de l'individu et ce qui relève du social n'est pas si simple.

> *«J'utilise sûrement mes propres capabilités à un moment donné de ce processus. Mais comment distinguer logiquement les capabilités qui d'une manière ou d'une autre définissent la personne, de celles qui sont le résultat des actions, extérieures, de la société ? »*[103]

Le logicisme d'Arrow l'avait amené à privilégier les jugements de valeur individuels ; le même logicisme l'incite à remettre en question la distinction conceptuelle entre ce qui dépend de l'individu et ce qui dépend de son environnement (même si notre pratique quotidienne nous porte à nous distinguer les uns des autres). La définition de l'action individuelle ne va pas de soi ; un jugement de valeur correspondant à une règle qui détermine ladite action, la définition d'un jugement de valeur individuel ne va pas de soi non plus.

Dès lors, on peut se demander pourquoi Arrow privilégie l'approche « individualiste » que nous avons décrite dans le paragraphe précédent.

> *« La factorisation implicite est une factorisation particulière, exécutée de façon particulière. Elle s'est révélée extrêmement pratique, notamment dans la perspective d'économiser le flot d'informations qui déferle dans le système économique. »*[104]

La partition de l'action sociale en jugements de valeur individuels trouve sa justification dans un souci de clarté : l'économiste du bien-être sacrifie l'exactitude à la simplicité du modèle, au risque de négliger des éléments importants de la réalité. Mais elle trouve aussi sa justification dans le souci éthique de privilégier la perspective évaluative de l'agent.

Remarquons toutefois que le néopositivisme de l'économie du bien-être des années 30-40 s'accommoderait mal de l'idée qu'une doctrine éthique puisse influencer la construction de

[102]Ibidem.
[103]Ibidem.
[104]Ibidem.

modèles « scientifiques ». Comment justifier cette « position libérale », selon laquelle il est souhaitable que les individus puissent déterminer eux-mêmes leurs valeurs et leurs actions ?

> *« La seule justification rationnelle en faveur de ce qu'on peut appeler une position libérale, ou peut-être plus exactement un principe de préférence sociale limitée, est que cette justification elle-même est un jugement de valeur. En d'autres termes, le fait qu'un individu pense qu'il n'est pas correct d'influencer les conséquences en dehors d'un domaine limité peut justement faire partie de son système de valeurs. »*[105]

Les contraintes méthodologiques de l'économie du bien-être ne permettent pas à Arrow de statuer sur un système de valeurs pour la société. La position libérale ne vaut que comme jugement de valeur, et le problème du choix social peut naître : l'expérience nous montre des hommes et des femmes qui empiètent sur le domaine des uns et des autres, avec les meilleurs ou les pires effets.

La lecture attentive des textes d'Arrow sur le choix collectif a mis en évidence trois arguments en faveur de jugements de valeur proprement sociaux, irréductibles à un jugement de valeur individuel:

- La discussion avec Buchanan et Adam Bergson a montré qu'Arrow défend l'idée de jugement de valeur collectif ou social. Cette idée s'inscrit même dans la méthodologie positiviste de la nouvelle économie du bien-être, puisque les jugements de valeur collectifs peuvent être connus par l'observation des actions collectives.
- Bien que les préférences particulières et les valeurs ne soient pas distinguées *a priori* dans le cadre conceptuel de la théorie du choix social, les conditions d'un choix collectif posées par Arrow ne prennent véritablement sens que dans les hypothèses non individualistes : le choix collectif est dérivé des valeurs individuelles, c'est-à-dire des préférences individuelles qui témoignent d'un intérêt pour l'organisation sociale plus que pour la consommation propre. Le choix collectif présuppose ainsi déjà l'existence de valeurs propres à juger la société.
- La partition de la décision collective en composantes individuelles a un caractère artificiel indéniable. Elle procède d'un choix pragmatique privilégiant la simplicité du modèle, mais également d'une aspiration plus normative, aspiration à l'autonomie individuelle et à la libre détermination des préférences de chacun.

[105]Ibidem.

2.3. Comparaisons interindividuelles et intégrité individuelle

L'analyse de la justification arrovienne de la primauté de la perspective évaluative de l'agent met en évidence le fait suivant : l'attribution de ces préférences aux individus procède du souci de valoriser la perspective évaluative de ces individus plus que du souci de décrire ce qui se passe exactement au moment de cette évaluation.

Cette hypothèse de lecture est confirmée par les raisons qui motivent Arrow à exclure les comparaisons interpersonnelles d'utilité de la base d'information du choix social. A première vue, cette exclusion est motivée par le caractère arbitraire desdites comparaisons. Mais Arrow exprime aussi une préoccupation plus profonde : les comparaisons interpersonnelles tendent à réduire les individus à des qualités, des propriétés, et ignorent la faculté de ces derniers de déterminer par eux-mêmes leurs préférences.

En effet, le cadre formel de la théorie du choix social ne caractérise pas par lui-même l'individu comme doté d'une volonté substantielle et stable (2.3.1.). Les comparaisons interpersonnelles de qualités, de situations et de satisfaction permettraient d'enrichir la description de l'individu (2.3.2.). Mais cette description fait de l'individu un support incertain de qualités éclatées ; or Arrow souhaite plutôt défendre une conception de l'individu comme agent autonome et comme support unifié et intègre de qualités (2.3.3).

2.3.1. Un individu sans substance

Le formalisme de la théorie du choix social laisse une certaine liberté dans la détermination et la définition des individus. L'individu n'est connu que par sa relation d'ordre, sa préférence révélée. La définition de la collectivité considérée reste arbitraire ; elle peut donc inclure des individus fictifs, qui ne sont pas présents ou qui n'existent pas encore.

La collectivité d'individus est en fait une collection de préférences, et nous pouvons nous en rapporter à Hume :

> *« Mais, laissant de côté certains métaphysiciens de ce genre, je peux me risquer à affirmer que les autres hommes ne sont qu'un faisceau ou une collection de perceptions différentes […] »*[106]

Les préférences sont des énoncés ou des comportements qui sont perçus. Le théoricien du choix social n'a pas directement affaire à un ensemble de personnes, mais à une collection de

[106] HUME, David, *Traité de la Nature Humaine*, traduction P. Baranger et P. Saltel, Paris : Garnier-Flammarion, 1995, I, IV, VI.

préférences révélées et perceptibles. Dans cette perspective, chaque agent n'est en fait qu'une succession de perceptions.

Cette proposition est la base d'un argument en faveur de la prise en charge par les institutions politiques et sociales de « nouveaux départs », et c'est en ce sens que l'utilise Marc Fleurbaey :

> *« Une interprétation littérale des théories de l'égalité des opportunités répondrait par l'affirmative à la question mentionnée ci-dessus, à savoir si les individus devraient être tenus responsables des choix passés qu'ils regrettent, mais cela peut sembler impitoyable. Après tout, on peut soutenir les changements de préférences ou de personnalités correspondent à une succession de différents « Mois » [selves], et il n'est pas certain que notre Moi actuel devrait être tenu responsable de toutes les décisions des Mois passés. Nous ne tiendrions pas des individus pour responsables des décisions d'autres individus responsables, à moins qu'ils ne les aient manipulés, et, même si différents Mois ne sont pas différentes personnes, ils pourraient être suffisamment distincts pour qu'on puisse faire valoir des considérations similaires – en outre, les Mois ultérieurs ne peuvent être soupçonnés d'avoir manipulé les Mois précédents. »*[107]

Cet argument montre que la théorie du choix social a affaire à des préférences et non des individus substantiels et continus. Les changements de préférences dans le temps montrent que la formule <une préférence = un individu> ne peut pas fonctionner. Sur le plan formel, la théorie du choix social n'a pas besoin de présupposer la permanence de l'identité personnelle.

D'ailleurs, le cadre conceptuel d'Arrow fonctionne également pour les « interactions » entre les différentes facettes d'une même personne. J. Elster et A. Hylland montrent l'intérêt de ce formalisme pour ouvrir certaines perspectives dans la conception d'une décision sociale :

> *« L'extension de la théorie du choix social aux cas intrapersonnels peut se présenter de différentes façons. La contribution d'Ian Steedman et d'Ulrich Krause dans The Multiple Self, un ouvrage de référence pour le présent travail, en est un exemple. Ils observent qu'un individu évalue souvent une situation de choix de différents points de vue, chacun étant associé avec un ordre de préférence des options. Chacun de mes « Mois » [selves], l'égoïste, l'altruiste, le moral et le social peuvent classer différemment les alternatives, de sorte qu'un mécanisme d'agrégation commence à devenir nécessaire. »*[108]

[107]FLEURBAEY, Marc, *Fairness, Responsibility and Welfare*, Oxford University Press, 2008, 7.1.Traduction libre.

[108] ELSTER, Jon et HYLLAND, Aanund, *Foundations of Social Choice Theory*, Cambridge University Press, 1986, Introduction, p 5. Traduction libre..

Ian Steedman et Ulrich Krause proposent ainsi une interprétation intrapersonnelle du théorème d'Arrow[109], qui fait du résultat d'impossibilité le symptôme des dilemmes auxquels un individu peut être confronté lors de son existence. Soulignant la possibilité d'une psychologie individuelle à plusieurs « dimensions », ou manifestant plusieurs « personnalités », Steedman et Krause montrent que le cadre conceptuel de l'économie du bien-être peut aussi représenter cet individu à plusieurs facettes. Chacune des facettes correspond alors un ordre de préférences, de même que chaque individu correspond à un ordre de préférence dans la théorie du choix social. Remarquons que cette analogie entre l'existence individuelle et l'existence collective, entre les tensions habitant l'individu et celles propres à la Cité évoque Platon, qui retrouvait dans la tripartition de l'âme les différentes classes de la Cité[110].

Ce que Steedman et Krause montrent, c'est que les outils formels de l'économie permettent de représenter la multiplicité des dimensions de l'individu. Il n'y a pas de relation nécessaire entre le formalisme et l'hypothèse d'un individu représentable par un seul ordre de préférence complet et transitif, hypothèse qui revenait à affirmer qu'un agent rationnel sait ce qu'il veut et choisit ce qu'il veut. L'argument selon laquelle cette conception de l'individu est nécessaire à la théorie économique parce qu'elle est la seule à pouvoir être modélisée tombe : il n'y a pas de correspondance nécessaire entre un individu et un ordre de préférence.

Steedman et Krause appliquent donc le résultat d'Arrow aux décisions indiviudelles, ou plus exactement « intrapersonnelles ». Il leur faut pour cela interroger la signification d'une condition de non-dictature pour le choix individuel. Steedman et Krause assimilent sa violation à un caractère « fanatique », l'une des facettes de l'individu déterminant toutes les autres.

Sur le plan formel, le cadre conceptuel de la nouvelle économie du bien-être ne suffit donc pas à justifier la préséance de l'individu sur la société.

[109] STEEDMAN, Ian, et KRAUSE, Ulrich, « Goethe's Faust, Arrow's Possibility Theorem and the individual decision-taker », in ELSTER, Jon (ed.), *The Multiple Self*, Cambridge University Press, 1987, pp 197-231.
[110] PLATON, *République*, 441c.

2.3.2. Les comparaisons interpersonnelles : enrichir la description de l'individu

L'individu de la théorie du choix social n'est en fait connu que par son classement de préférence. Si nous voulons caractériser cet individu de manière plus précise, nous devons introduire plus d'information à propos des évaluations individuelles. Ainsi, les comparaisons interpersonnelles permettraient de mieux connaître l'individu. Celui-ci gagnerait en consistance et en « épaisseur ».

Ces comparaisons feraient ainsi appel à une description précise des états et comportements individuels. Cette description recenserait beaucoup plus d'informations, enrichissant l'individu de caractéristiques ignorées par l'approche ordinale. Il faudrait :

> *« explorer, bien au-delà des régions à considérer dans le choix individuel, le système de préférences de l'individu en présence de nombreuses situations extérieures au champ de choix. »*[111]

Chaque individu devrait comparer sa situation actuelle avec les situations qu'il pourrait occuper.

Pour mettre en œuvre ces comparaisons interpersonnelles, certains économistes du bien-être[112] ont mis en place un système de seuils permettant d'attribuer une valeur numérique à chaque degré de satisfaction ou d'insatisfaction qu'éprouve l'individu. Pour que ces seuils aient une signification quelconque, les individus doivent pouvoir partager des expériences similaires. Or, Arrow remarque qu'

> *« une faible différence du nombre de seuils de discrimination entraîne une plus grande inégalité de revenu. »*[113]

Des comparaisons interpersonnelles mal formulées ou peu précises pourraient générer des décisions sociales contraires aux aspirations collectives. Dans la mesure où la théorie du choix social tente de supprimer l'arbitraire dans les décisions telles que l'allocation des revenus, c'est gênant. Il importe donc que les individus « se comprennent ».

---

[111] ARROW, Kenneth, *Social Choice and Individual Values*, New Haven: Yale University Press, 2ème édition révisée, 1963. Traduction française par l'Association de Traduction Economique de l'Université de Montpellier, Paris : Calmann-Levy, 1974, p 204.

[112] Cf. GOODMAN et MARKOVITZ, « Social Welfare function based on individuals rankings », *American Journal of sociology*, vol. 58, 1952, pp 257-262, VAN PRAAG (B.M.S.), "The Relativity of the Welfare concept", in NUSSBAUM and SEN (Eds), *The Quality of Life*, Oxford University Press, 1993.

[113] ARROW, Kenneth, *Social Choice and Individual Values*, New Haven: Yale University Press, 2ème édition révisée, 1963. Traduction française par l'Association de Traduction Economique de l'Université de Montpellier, Paris : Calmann-Levy, 1974, n 64 p 213.

Les comparaisons interpersonnelles caractérisent ainsi les individus avec plus de précision, requérant pour cela un accord sur les critères d'évaluation de chacun. Le choix des préférences ordinales comme base d'information du choix collectif n'est justifié que s'il s'avère :

- que nous ne pouvons pas avoir accès à plus d'information (pour des raisons épistémologiques : les individus ne seraient pas fiables — ce qui revient à privilégier la parole de l'expert — , ou pour des raisons pratiques : ces comparaisons ne sont pas faisables —mais exiger de chaque individu qu'il classe *tous* les états sociaux possibles par ordre de préférence l'est-il davantage ?)
- que les individus ne se comprennent pas entre eux.

Arrow est sceptique quant à la possibilité d'un accord sur les critères d'évaluation sociale : le résultat d'impossibilité ne met-il pas en évidence ce dont chacun de nous peut faire l'expérience, à savoir les conflits insolubles ?. Les comparaisons interpersonnelles enrichiraient la base d'information de la théorie du choix social, mais à condition de déterminer une méthode permettant aux individus de se comprendre et de s'accorder sur les critères de comparaison.

2.3.3. Nous ne comparons pas des individus, mais des qualités

Plus que par le caractère arbitraire des comparaisons interindividuelles de satisfaction, Arrow est préoccupé par la question de l'intégrité de l'individu et de l'autonomie de ses choix. En effet, bien qu'il affirme que des comparaisons de situations individuelles motivées par la sympathie peuvent constituer une solution au problème du choix social (a), il reste préoccupé par la tendance de cette perspective des comparaisons interpersonnelles à ne voir dans l'individu qu'un agrégat de qualités à comparer (b).

a) Comparaisons interpersonnelles et sympathie étendue

A la fin de son article « Values and Collective Decision Making », Arrow invite les chercheurs futurs à relaxer la condition d'indépendance des alternatives non pertinentes afin de réintroduire des comparaisons interpersonnelles. La méthode suggérée dans un article postérieur est celle de la « sympathie étendue » (*extended sympathy*)[114] : celle-ci est en œuvre lorsqu'un professeur de Harvard affirme qu'il vaut mieux donner 500 dollars à un mendiant

[114] ARROW, Kenneth, "Extended Sympathy and the possibility of Social Choice", *American Review Paeres and Proceedings*, 67 (1977); pp 219-225.

bengali plutôt qu'à lui, ou lorsqu'une jeune fille cède sa place à une dame âgée dans un bus. De tels jugements ont la forme suivante : « l'action x est meilleure pour toi que pour moi » (mais ce peut être l'inverse, je peux juger qu'il vaut mieux m'attribuer une ressource). Il appartient à la personne qui compare de décider ce qui fait qu'une action est « meilleure ». Les comparaisons interpersonnelles intègrent une liste d'éléments très étendue : des « biens », des « satisfactions » ou encore des « réalisations ».

Arrow démontre ainsi que l'effacement de la distinction entre possessions matérielles et aptitudes individuelles est pertinent dans une situation de choix collectif. L'effacement de cette distinction enrichit ainsi les comparaisons interpersonnelles. Remarquons que pour formuler des jugements tels que « l'action x est meilleure (ou pire) pour moi que l'action y pour ma voisine », nous devons nous mettre à la place d'autrui.

> *« Sous cette forme, la comparaison tient compte des caractéristiques qui définissent un individu. En effet, ces caractéristiques vont de pair avec les biens en général considérés comme éléments de la fortune d'un individu. L'appropriation des instruments de production a souvent été considérée comme un élément d'un état social ; pourquoi n'en serait-il pas de même pour les aptitudes à utiliser ces instruments, ou encore pour l'intelligence qui donne naissance à ces qualifications ? Les individus, lorsqu'ils comparent entre eux leurs différents niveaux de bien-être, non seulement tiennent compte de l'appropriation des biens matériels mais également des qualités de rayonnement ou d'adresse qui y sont attachées. »*[115]

La comparaison interpersonnelle implique donc un intérêt pour autrui – contrairement aux « hypothèses individualistes » longuement développées auparavant – ainsi qu'une caractérisation très fine des individus et de leur situation économique et sociale. La conception de l'individu est enrichie, il cesse en effet d'être identifié par son seul classement de préférences ; en revanche, l'idée que les individus puissent changer conserve sa pertinence. Ces comparaisons mettent également en lumière les relations entre individus, contrairement à l'approche en termes de préférences ordinales non comparables. En effet, elles présupposent des rapports entre les individus (je ne puis me comparer qu'avec des êtres que je connais et qui m'*intéressent*). Par ailleurs, en attirant l'attention sur les aptitudes et les qualités des individus permettant de faire usage d'une ressource, ce type de comparaison peut mettre en évidence ce qui, dans les rapports interindividuels, peut favoriser ou non l'usage d'un bien : les Juifs du Ghetto de Varsovie ne pouvaient pas aller à l'extérieur du ghetto, même s'ils avaient un vélo et des jambes valides. Cet exemple dramatique montre qu'il est parfaitement

[115] ARROW, Kenneth, *Social Choice and Individual Values*, New Haven: Yale University Press, 2ème édition révisée, 1963. Traduction française par l'Association de Traduction Economique de l'Université de Montpellier, Paris : Calmann-Levy, 1974, p 208.

pertinent d'utiliser les comparaisons interpersonnelles pour décrire la domination et ses effets sur le bien-être des individus.

b) *Comparaisons interpersonnelles, incommensurabilité des individus et autonomie*

Revenons à notre problématique de la place de l'individu dans l'évaluation sociale : afin de comprendre l'enjeu de cette réintroduction des comparaisons interpersonnelles par le biais de la sympathie étendue, il nous faut répondre à deux questions. Ces comparaisons ont-elles l'individu pour source première ? L'individu est-il la donnée exclusive de ces comparaisons ?

En réponse à la première question, on peut noter que des individus peuvent être à l'origine de jugements de valeur tels que : « il est préférable de donner cet argent à x plutôt qu'à y, étant donné les privations dont soufre x ». Mais de tels jugements, de tels classements de préférences, peuvent également être le fait d'institutions, le résultat d'un vote, bref, de tout agent susceptible de déterminer la répartition des ressources entre x et y.

Concernant la seconde question, si l'individu est bel et bien ce dont la situation et les attentes constituent l'élément premier de l'évaluation et la base d'information de l'évaluation sociale, alors nous revenons ici à ce que défendait l'économie du bien-être d'avant l'ordinalisme : le bien-être social sera déterminé par l'agrégation de données individuelles, définies grâce à une méthode certes plus élaborée, dont l'intérêt réside dans le fait qu'elle fait explicitement appel aux préoccupations altruistes des membres de la société. Mais Arrow ne l'entend pas vraiment ainsi : il suggère que, dans la théorie du choix social revisitée par les comparaisons interpersonnelles inspirées par la sympathie étendue, ce qui est agrégé, ce ne sont pas des individus identifiés à travers des attentes bien distinctes, mais des *qualités.*

Et, afin de souligner que l'individu n'est pas, selon lui, un support de qualités mais davantage une somme desdites qualités, Arrow cite Blaise Pascal à la fin de l'article « Extended Sympathy and Social Choice» :

> *« Qu'est-ce que le Moi ?*
>
> *Un homme qui se met à la fenêtre pour voir les passants ; si je passe par là, puis-je dire qu'il s'est mis là pour me voir ? Non ; car il ne pense pas à moi en particulier ; mais celui qui aime quelqu'un à cause de sa beauté, l'aime-t-il ? Non ; car la petite vérole, qui tuera la beauté sans tuer la personne, fera qu'il ne l'aimera plus.*
>
> *Et si on m'aime pour mon jugement, pour ma mémoire, m'aime-t-on ? moi ? Non, car je puis perdre ces qualités sans me perdre moi-même. Où est donc ce moi, s'il n'est ni dans le corps, ni dans l'âme ? et comment aimer le corps ou l'âme, sinon pour ces qualités, qui ne sont point ce qui fait le moi,*

puisqu'elles sont périssables ? car aimerait-on la substance de l'âme d'une personne, abstraitement, et quelques qualités qui y fussent ? Cela ne se peut, et serait injuste. On n'aime donc jamais personne, mais seulement des qualités.

Qu'on ne se moque donc plus de ceux qui se font honorer pour des charges et des offices, car on n'aime personne que pour des qualités empruntées. »[116]

La thèse de Pascal est que le milieu social n'a pas accès au « Moi » (s'il en est un), mais à des qualités, qualités que l'on peut perdre.

Qu'apporte Pascal à cette réflexion sur la place de l'individu dans le processus de choix social ? Si les comparaisons des situations et satisfactions entre personnes ont un sens, elles peuvent se référer à des comparaisons de situations faisables ou simplement concevables. Ainsi, il est possible de classer les situations non seulement en fonction des ressources dont jouissent chacun des membres, mais aussi en fonction de leurs dispositions, des qualités qu'ils ont acquises dans l'un ou l'autre des états sociaux. Ainsi, il serait possible, nous dit Arrow, de rassembler les différences entre personnes en une liste de qualités : en d'autres termes, nous ne comparons pas des individus, mais des qualités. Ces comparaisons interpersonnelles n'ont donc pas pour élément premier l'individu, mais les qualités : je compare la beauté, la mémoire, et non les personnes belles, celles qui ont de la mémoire…

Mais que serait un individu qui ne serait qu'une liste de qualités ?

"[…] Réduire l'individu à une liste spécifique de qualités revient à nier son individualité, au sens propre du terme. La dernière ligne de la citation de Pascal devrait éveiller un doute. D'une certaine manière, que je ne parviens pas à bien expliciter et que je ne suis pas sûr de vouloir défendre, l'autonomie des individus, un élément d'incommensurabilité mutuelle entre les personnes, semble niés par la possibilité de comparaisons interpersonnelles. Je ne doute pas que ce soit cette sorte de sentiment qui m'ait rendu si réticent à abandonner l'ordinalisme pur, en dépit de mon désir de trouver une base pour une théorie de la justice»[117]

Il faut ici distinguer le point de vue de l'évaluation sociale de celui de la personne dont la situation est évaluée. L'individu est un *datum* parmi d'autres *data* possibles pour l'évaluation sociale. L'information première pour la décision collective peut être une information émanant de l'individu, mais également une qualité, ou une privation de qualité. Les évaluations sociales fondées sur les comparaisons s'intéressent de prime abord aux qualités. En revanche, du point de vue de la personne dont la situation est évaluée, il est possible de dire « Moi ».

[116] PASCAL, Blaise, *Pensées*, Lafuma 688.

[117] ARROW, Kenneth, « Extended Sympathy and the Possibility of Social Choice », *American Economic Review Papers and Proceedings* 67, 1977, pp 219-225. Traduction libre.

Non pas que l'identité personnelle soit vierge de tout questionnement ! Mais l'identification, la possibilité de distinguer un Moi et un Autrui, identification à l'œuvre dans la vie quotidienne et dans le langage courant, dès le plus jeune âge, n'est pas remise en cause par une théorie du choix social qui prendrait appui sur des comparaisons de qualités et non d'individus. Cette distinction permet peut-être de lever le doute qui habite Arrow au moment où il écrit ces quelques lignes. En revanche, elle soulève certaines questions : jusqu'ici, nous n'avons pas porté notre attention sur le fait que la théorie du choix social, en se présentant comme un processus de construction de jugements de bien-être collectifs, ne distinguait pas l'information sur les attentes exprimée par les membres de la société (dans le cas du vote) et celle découverte par une analyse d'expert étudiant la situation économique des membres de la société (le travail de l'économiste du bien-être). Or, en distinguant le point de vue « externe » sur l'individu et le point de vue « interne », il apparait que l'expression par l'individu de ses propres attentes peut jouer un rôle structurant pour ce dernier. D'autre part, certaines attentes peuvent correspondre à des finalités de type identitaires, où c'est la singularisation de l'individu qui est en jeu. La théorie du choix social considère toutes les attentes, économiques ou identitaires, de manière indifférenciée : peut-être est-ce une erreur ?

Arrow ne s'est pas engagé plus loin dans cette voie, que suivront John Harsanyi et Amartya Sen, jugeant qu'il était difficile de construire une théorie du choix social sur les comparaisons interpersonnelles. Cette analyse a décelé dans les écrits d'Arrow une tension entre les possibilités du formalisme et le désir de sauver l'intégrité de l'individu. Nous prenons ces possibilités comme une invitation à la réflexion, aussi curieuses que puissent paraître les implications : les enjeux soulevés par les questions de choix collectif et d'identité personnelle, ou collective, sont trop importants pour que nous puissions négliger ne fût-ce qu'une seule piste. Quant à l'intégrité de l'individu, reste à clarifier son sens. L'hypothèse choisie sera déterminante pour la signification de la théorie du choix social, car il nous appartiendra de décider si le choix social est le produit des individus autonomes et autosuffisants, ou bien s'il produit lui-même les conditions de développement de l'individualité.

Chapitre 3 : Evaluation individuelle et évaluation sociale dans l'approche par les capabilités

Avant de commencer l'analyse de la conception de l'individu présupposée par les développements de la théorie du choix social qu'a proposés Amartya Sen, il convient de faire le point sur sa méthodologie philosophique. La méthode d'élaboration des concepts et des arguments philosophiques de Sen est largement inspirée par la philosophie du « langage ordinaire ». Un concept ou un argument ne peuvent être acceptés s'ils heurtent trop les conceptions ordinaires et le langage courant. Ses critiques des conceptions concurrentes de la pauvreté ou de la justice font fréquemment appel au « sens commun » ; ce choix méthodologique supporte le passage du « rationnel » au « raisonnable », concepts empruntés à Rawls mais largement retravaillés par Sen. La méthode de recherche de Sen présente les caractéristiques suivantes :

- Il n'y a pas de restriction *a priori* des critères pertinents en matière d'évaluation : les critères émergent de la description des faits, description qui se doit d'être le plus riche possible ;
- L'élucidation des problèmes d'évaluation passe par la comparaison d'états sociaux, réels ou virtuels. L'état social préféré sera retenu pour l'évaluation, parce qu'il a passé avec succès l'examen de la comparaison et non pas parce qu'il se conforme à un critère déterminé *a priori*.

In fine, c'est notre « bon sens », le sens partagé de ce qui est acceptable et de ce qui ne l'est pas, qui permet d'exclure un critère, une conception du monde ou encore un état social particulier.

De ce point de vue, la démarche de Sen est conséquentialiste, mais il défend une conception riche de la conséquence. Cela signifie que la description des conséquences peut et doit inclure les valeurs en jeu, les points de vue des acteurs impliqués, la manière dont le fait est advenu, ce que le fait rend possible ou non pour chacun de nous.

A présent, rappelons notre problème: dans l'approche prônée par Amartya Sen, les individus sont-ils la source exclusive, ou du moins la source principale, de l'évaluation collective ? Peut-on taxer Sen d'individualisme méthodologique ? Ce problème a deux dimensions, séparables en théorie mais fortement intriquées dans les faits :

- Une dimension épistémologique : il s'agit de déterminer si les individus sont à l'origine de l'évaluation sociale, quel que soit le cadre de cette évaluation (élections politiques, revendications sociales ou encore échanges économiques).
- Une dimension éthique : les individus devraient-ils être à l'origine de l'évaluation sociale, même s'ils subissent des contraintes les empêchant d'exercer leur jugement de manière effective ?

Dans cette partie, nous montrerons que la réponse à apporter est la suivante : la contribution de Sen à la théorie du choix social et à l'économie normative (incluant donc l'approche par les capabilités) présuppose un individu dont la capacité de préférer, de juger et de choisir n'est pas donnée, mais dépend de conditions sociales de réalisation. En particulier, les capabilités visent à rendre possible cette capacité de préférer, de juger et de choisir.

Pour ce faire, nous nous attacherons à dégager la compréhension de l'individu dans l'ensemble de la contribution d'Amartya Sen (en philosophie comme en économie), dont nous présupposons la cohérence[118]. Nous y étudierons l'individu de trois points de vue différents, chacun représentant un enjeu distinct et constituant une étape de notre démonstration. Le premier point de vue est celui de l'expert, du tiers extérieur, réel ou fictif, qui s'attache à décrire aussi précisément que possible la source réelle des évaluations comptabilisées comme évaluations individuelles par la théorie du choix social ; cette source peut en fait mieux être décrite en termes de « position » qu'en termes de « subjectivité » (3.1.). Le deuxième point de vue est celui de l'individu lui-même ; ce point de vue « interne » permet de mettre en évidence le caractère « construit » de l'individualité à travers l'histoire personnelle (3.2.). Le troisième point de vue est celui de la collectivité sur les individus qui en font partie ; point de vue obtenu par agrégation ou recoupement, conformément aux exigences du choix social, il définit la place de l'individu dans le processus de choix collectif et la manière dont les institutions et les processus sociaux peuvent influencer cette place (3.3.).

[118] Cette cohérence est revendiquée par Sen lui-même : ainsi, parmi ses derniers livres, *L'Idée de Justice* présente ses différentes contributions, tant en théorie du choix social qu'en économie du développement, en histoire des sciences ou en philosophie morale, comme les pièces d'une même approche de la justice. Cela dit, il convient de préciser que ses contributions en économie et en philosophie ne sont pas toujours lues et travaillées par les mêmes personnes, et la cohérence ne se dégage donc pas à première vue.

3.1. Des préférences individuelles aux évaluations positionnelles

Du point de vue de l'expert ou du tiers, quelle est la source des évaluations recensées par le choix social et les jugements de bien-être collectif ? En économie normative, ces évaluations prennent la forme de la description d'attributs individuels, description sur laquelle repose la formulation des attentes en termes d'avantage. Cette description s'est d'abord faite en termes d'utilités et de préférences. Sen a inauguré une nouvelle « métrique » de l'avantage individuel, la capabilité.

Le fait que ces attributs soient de prime abord conçus comme individuels, y compris chez Sen, implique-t-il que ce dernier pourrait faire sienne la thèse selon laquelle un individu substantiel, doté d'une psychologie singulière et de facultés propres, serait à la source des évaluations de l'avantage ?

Afin d'apporter des éléments de réponse, il nous faut raisonner ainsi : s'il existe au moins un individu qui ne peut pas évaluer son état et celui de la société dans laquelle il vit, alors on ne peut taxer la théorie du choix social réinterprétée par Sen d' « individualiste ». Il faudra alors dissocier le point de vue propre à l'individu et le point de vue de ce que nous avons appelé l' « expert »[119]. Comment, du point de vue de l'expert, décrire alors ce qui est à la source des évaluations recensées dans l'agrégation des préférences, des utilités ou des capabilités ?

Notre hypothèse est que cette source peut se décrire à l'aide du concept senien de « position ». Le concept d'objectivité positionnelle introduit par Sen dans sa contribution en philosophie permet de caractériser la place, les limites et l'importance de l'évaluation individuelle pour le choix social. Cette reconstitution comprend quatre étapes. Premièrement, nous verrons que les attributs individuels pertinents pour l'évaluation sociale selon Sen incluent tant des caractéristiques objectives que subjectives (3.1.1). Deuxièmement, nous rappellerons que l'approche par les capabilités est une solution au problème de la distorsion entre évaluation individuelle et subjective, d'une part, et condition objective de l'agent, d'autre part (3.1.2). Troisièmement, nous proposerons une caractérisation plus fine du positionnement de Sen vis-à-vis de l'individualisme méthodologique, à l'aide des concepts de néo-holisme et d'infra-individualisme (3.1.3). Quatrièmement, nous montrerons en quoi l'objectivité positionnelle caractérise de la façon la plus adéquate l'attribution d'une évaluation à un agent (3.1.4).

[119] L'auteur de ces lignes regrette le caractère bureaucratique de l'expression.

3.1.1. Les attributs objectifs et subjectifs de l'individu dans l'évaluation de l'avantage

L'évaluation individuelle, lorsqu'elle a lieu, non seulement s'exprime de différentes manières, mais porte sur des contenus divers qui représentent les aspects variés de l'existence humaine. Sen déploie plusieurs concepts pour mesurer la qualité de l'existence d'un individu. Un tableau permet de recenser et de classer ces différents concepts[120] :

[120] Tableau élaboré à partir de : SEN, Amartya, *On Ethics and Economics*, 1987 (The Royer Lectures, John. M. Letiche (ed.) (University of California), Berkeley), Oxford, Blackwell Publishers, 1987 [Trad. Française S. Marnat, *Ethique et Economie*, Paris, PUF, 1993, p 56-57] et de SEN, Amartya, *Rationality and Freedom*, Belknap Press of Harvard University Press, 2002 [Trad. Française M.-P. d'Iribane-Jaawane, Odile Jacob, 2005, p 487].

Finalité de l'agent / Modalité de l'aspect évalué	**Bien-être** (« well-being ») : « ce qu'une personne accomplit et les possibilités qui s'offrent à elle du point de vue de son avantage personnel (qui n'équivaut pas nécessairement à la seule satisfaction des désirs) »	**Action** : « [cet aspect] tient compte également d'autres objectifs et valeurs, qui peuvent dépasser largement la poursuite du bien-être personnel »
Accomplissement : « Ce que l'individu finit par accomplir ou obtenir. »	FONCTIONNEMENT Ex : Nadia mange à sa faim	ACTION ACCOMPLIE Ex : Les femmes ont le droit de vote, conformément aux valeurs et objectifs de Nadia
Liberté-opportunité : « options ou opportunités dont jouit l'individu »	CAPABILITE (*capability*) Ex : Nadia peut manger à sa faim si elle le souhaite	LIBERTE D'ACTION Ex : Nadia peut militer en faveur du suffrage féminin, y compris en faisant une grève de la faim
Liberté-processus : procédure de décision impliquant le contrôle de la personne elle-même	PREOCCUPATION PERSONNELLE DU PROCESSUS : « manière dont les individus aiment à faire leurs choix » Ex : Nadia peut choisir elle-même de manger à sa faim ou non, sans contrainte sociale (heures de repas, régime imposé pour raisons esthétiques ou religieuses) ou physique (alimentation par sonde)	PREOCCUPATION SYSTEMIQUE DU PROCESSUS : « vision des individus au sujet des règles et des institutions sociales » Ex : Conformément aux valeurs de Nadia, les institutions sociales favorisent la participation des femmes à la vie politique.

Tableau 2

Chaque dimension représentée dans ce tableau peut être comprise dans la description d'un état social soumis à l'évaluation et au choix.

Ce tableau permet de cartographier les concepts évaluatifs de Sen. Certes, les distinctions sont toujours floues, puisqu'elles sont dégagées de l'observation d'une réalité complexe. Ainsi, le souhait de Nadia de voir les femmes avoir le droit de vote concerne à la fois ses engagements en faveur de l'égalité des hommes et des femmes et son propre bien-être (étant une femme,

elle souhaite que ses intérêts soient représentés). De même, une action accomplie peut procurer de la joie et améliorer ainsi le bien-être d'un individu.

On remarquera le caractère hétérogène des attributs individuels pris en compte dans les jugements de valeur pertinents pour le choix collectif. Les aspects « bien-être » et « action » évoquent à la fois les écoles déontologiques et conséquentialistes en philosophie morale. L'attention portée aux différentes modalités de réalisation (opportunité et accomplissement, capabilités et réalisations) évoque une description du bien-être et de l'épanouissement de type néo-aristotélicien. Cet œcuménisme méthodologique fait écho à une position pluraliste en philosophie politique et morale.

Mais cet œcuménisme soulève une difficulté d'ordre épistémologique : certains des attributs individuels retenus pour l'évaluation sociale sont « subjectifs », au sens où ils dépendent exclusivement de l'individu particulier élevé : par exemple, le plaisir éprouvé par Nadia lorsqu'elle apprend que les femmes ont obtenu le droit de vote dans son pays est une caractéristique subjective. D'autres sont « objectifs » : ainsi, le droit de vote de Nadia peut être caractérisé indépendamment de l'existence de celle-ci[121].

Les deux colonnes du tableau (la colonne « bien-être » et la colonne « action ») se distinguent par le rapport de l'évaluation individuelle à l'avantage personnel. Mais, en y regardant de plus près, nous constatons que chaque colonne suppose deux sens différents de l'expression « évaluation individuelle ». Dans la première colonne, l'individu, plus précisément son bien-être, est l'objet de l'évaluation, ce sur quoi celle-ci porte. Mais rien ne nous dit qu'il effectue lui-même cette évaluation : les capabilités et les fonctionnements de Nadia peuvent aussi bien être décrites par un observateur extérieur. En revanche, la seconde colonne suppose l'expression d'un jugement de valeur par l'individu concerné lui-même. L'individu est à proprement parler le sujet de l'évaluation, au sens grammatical du terme « sujet ». En revanche, et c'est ce qui nous oblige à écouter la « voix » de l'évaluateur, ce qu'il évalue peut être tout à fait indépendant de sa propre situation.

[121] Pour une typologie des théories économiques de la justice en fonction des caractéristiques individuelles retenues pour l'évaluation sociale, cf. FLEURBAEY, Marc, *Théories économiques de la justice*, Paris, Economica, 1996, p 18. Fleurbaey qualifie les attributs individuels retenus par Sen de « complets ».

3.1.2. Le problème de la distorsion des évaluations individuelles : illusion objective et capabilités

Les travaux de Sen sur la pauvreté et les inégalités de genre ont mis en évidence la distorsion entre évaluation subjective et évaluation objective des situations individuelles. Le problème est que l'évaluation individuelle peut n'être que le produit d'une morale traditionnelle qui va à l'encontre des intérêts fondamentaux de l'individu concerné (a). Sen a suggéré de penser cette difficulté en termes d' « illusion objective », concept qu'il attribue à Marx (b). Cette difficulté a motivé le refus de l'individualisme méthodologique dans l'approche des capabilités (c).

a) Morale traditionnelle et arguments cognitifs

Lorsque les évaluations individuelles reflètent la morale traditionnelle, elles peuvent aller à l'encontre d'une description objective de la situation de ce même individu. Sen a mis en évidence ce problème dans ses travaux sur l'inégalité des sexes dans les milieux conservateurs.

> *« La question de la perception est capitale pour comprendre l'inégalité des sexes dans les sociétés traditionnelles, et si l'on veut remettre en question la morale traditionnelle sur le plan de l'éthique il faut s'appuyer sur des arguments cognitifs »*[122]

En effet, des travaux empiriques ont montré que les femmes souffrent d'inégalités objectives dans de nombreux pays en voie de développement : leur longévité est inférieure à celle des hommes, alors qu'il semble qu'elles aient un avantage biologique de ce point de vue. Cela qui modifie le ratio des sexes en leur défaveur. Ainsi, le rapport du PNUD sur le développement humain en 2010 présente un indice de mesure d'inégalité des genres s'appuyant sur la santé de reproduction, la participation politique et l'éducation, ainsi que l'insertion dans la vie active. Aucun pays n'atteint l'indice 0, représentant une parfaite égalité des genres, et de nombreux pays affichent des résultats préoccupants[123]. Ces données constituent les « arguments cognitifs » que l'on oppose à la « morale traditionnelle » : cette expression désigne ce qui détermine l'évaluation de leur propre situation par les femmes concernées elles-mêmes. On peut ainsi confronter les données du tableau 4 et celles du tableau 9[124] du

[122]SEN, Amartya, *On Ethics and Economics*, 1987 (The Royer Lectures, John.M. Letiche (ed.) (University of California), Berkeley), Oxford, Blackwell Publishers, 1987 [Trad.Française S. Marnat, *Ethique et Economie*, Paris, PUF, 1993, n2 p 22].

[123] Cf. : http://hdr.undp.org/en/media/HDR_2010_FR_Tableau4.pdf, consulté le 15 novembre 2010.

[124] Cf. : http://hdr.undp.org/en/media/HDR_2010_FR_Tableau9.pdf, consulté le 15 novembre 2010.

rapport 2010 du PNUD sur le développement humain, et découvrir d'apparentes incohérences : par exemple, le taux de mortalité maternelle au Niger est de 1800 pour 100 000 naissances vivantes, et celui du Togo est de 510 pour 100 000 naissances vivantes : une Nigérienne a donc trois à quatre fois plus de chances de mourir en couches qu'une Togolaise et devrait exprimer un mécontentement face à cet état de fait. Mais 82% des Nigériens et Nigériennes se disent satisfaits de leur santé, et 94% des Nigériennes s'estiment « traitées avec respect ». En revanche, seulement 40% des Togolais et Togolaises sont contents de leur état de santé, et seulement 55% des Togolaises pensent être « traitées avec respect ». Dans ce cas précis, les expressions d'insatisfaction ne constituent pas un outil d'évaluation fiable[125].

[125] Bien entendu, nous souhaiterions en savoir davantage afin de circonscrire les raisons de cette distorsion (par exemple, il serait intéressant de savoir à quel type d'expérience correspond l'expression « ne pas être traité avec respect » pour les personnes interrogées). Il faut également ajouter que les statistiques de la perception reflètent une satisfaction d'ensemble pour les pays satisfaisant au mieux les critères de développement humain du PNUD : les « voix » des agents ne sont pas toujours biaisées.

b) Le prisme marxien

L'article de Sen, « Gender and Cooperative Conflict », fut motivé par des résultats de cet ordre, ainsi que par des études de terrain qui avaient suscité de multiples questionnements chez les spécialistes du développement. Cet article montre que les prétentions des femmes à une part équitable du revenu familial (sous forme de nourriture, par exemple) sont limitées par le fait qu'elles ne s'en perçoivent pas comme les tributaires légitimes, notamment lorsqu'elles n'ont pas un travail salarié. Mais Sen ne s'arrête pas là.

L'illusion objective

Sen refuse de s'en tenir aux préférences révélées de ces femmes pour évaluer leur bien-être : ces femmes pourraient être victimes d'une « illusion objective », concept emprunté à Marx pour caractériser une perception biaisée de l'intérêt d'un individu ou d'une classe, biaisée en raison même de la position sociale et économique de cet individu ou de cette classe. Amartya Sen utilise le concept d'"illusion objective" pour désigner l'illusion de faible morbidité des femmes originaires de régions indiennes où l'analphabétisme prévaut, comme l'Uttar Pradesh[126].Une croyance peut être positionnellement objective tout en étant erronée : positionnellement objective, parce qu'une description précise de la position économique, sociale, politique de l'agent permet de comprendre que la croyance est probable, voire inévitable lorsque l'on occupe cette position ; erronée, dans la mesure où un diagnostic émanant d'une position caractérisée par des capabilités (*capabilities*) et des libertés plus étendues ou qualitativement plus satisfaisantes permet de mettre en évidence la fausseté de la croyance en question.

Ce concept d'illusion objective est issu de la relecture de l'œuvre de Marx par G. A. Cohen, auteur du livre *Karl Marx's Theory of History : a Defence*[127], à l'origine du courant marxiste analytique. L'ambition de ce livre est de défendre un matérialisme historique reformulé selon les exigences de la philosophie analytique. En annexe de l'ouvrage, se trouve un article publié quelques années plus tôt dans lequel Cohen rappelle que le projet de Marx en écrivant le *Capital* était de mettre à nu la réalité qui sous-tendait et déterminait les relations de production capitalistes[128]. En effet, le travailleur comme l'entrepreneur capitaliste pris dans ces relations n'ont pas l'aptitude de discerner cette réalité : s'ils l'avaient, le système économique, social et politique capitaliste ne pourrait perdurer.

[126]La dernière version de son argument se trouve dans: *Rationality and Freedom*, Harvard University Press, 2002, ch. 15.
[127] Oxford Clarendon Press, 1978.
[128] « Karl Marx and the withering away of social sciences » *Philosophy and Public Affairs,* 1, no 2, hiver 1972.

De fait, les diagnostics d'illusions se multiplient tout au long du *Capital.* Ainsi, dans la section consacrée au salaire, Marx montre que le salaire est la forme phénoménale, l'apparence qui masque la réalité du rapport entre valeur et prix de la force de travail.

> *"[Toutes les formes phénoménales] se reproduisent d'une façon immédiatement spontanée, comme idées reçues et formes de pensée courantes, alors que [l'arrière-plan caché] ne peut d'abord être découvert que par la science. L'économie politique classique touche de près le fond des choses, mais sans le formuler consciemment.»*[129]

Pour l'ouvrier exploité, dont les 12 heures de travail sont payées la valeur de 6 heures (il fournit donc 6 heures de travail en plus), ces 12 heures sont objectivement le seul moyen de gagner le salaire de 6 heures.

> *"Par suite, toute variation dans la grandeur de l'équivalent qu'il reçoit lui apparaît nécessairement comme une variation dans la valeur ou le prix de ses 12 heures de travail. »*[130]

L'ouvrier croit objectivement que la valeur de son travail ou de sa compétence a diminué sur le marché, alors que le diagnostic de l'économiste marxiste est que cette variation résulte de facteurs indépendants de la valeur réelle dudit travail : le coût du pain, l'offre ou la demande…

Du point de vue du capitaliste qui exploite la force de travail de l'ouvrier, son profit ne s'explique pas par l'exploitation, mais par le fait que le prix de vente de sa marchandise est plus élevé que son coût de production ou d'acquisition. Si l'illusion propre à la production capitaliste se dissipait, la capitaliste verrait que ce qui fait la valeur de sa marchandise est le travail des ouvriers qui y est « incorporé ». Cohen remarque d'ailleurs que le capitaliste paierait le travail des ouvriers à sa véritable valeur s'il n'était pas victime de cette illusion : en effet, il n'y a rien de naturel dans la recherche de son avantage propre, aux dépens de celui des autres. Mais le capitaliste ne ferait alors pas de profit, n'accumulerait pas de capital, ne serait donc pas un capitaliste, et la société ne le serait pas davantage. La dissipation de l'illusion au niveau de l'ensemble de la société correspond à un changement radical de mode de production[131].

---

[129] MARX, Karl, *Das Kapital*, 1867, vol 1. Trad. Fr. *Le Capital,* PUF Quadrige 2009[3], chap. XVII.
[130] *Ibidem.*
[131] En d'autres termes, le problème de l'exploitation n'est pas un problème moral.

Le lien entre illusion et persistance du système capitaliste est réaffirmé plusieurs fois dans le *Capital*.

> *"Le profit, tel qu'il se présente ici, est donc la même chose que la plus-value : cette mystification découle nécessairement du mode de production capitaliste. Comme on ne peut plus distinguer, dans la formation apparente du coût de production, entre capital constant et capital variable, l'origine de la mutation de la valeur, qui se produit au cours du processus de production, doit être transférée de la partie variable du capital au capital total. Et parce que le prix de la force de travail apparaît à l'un des pôles sous la forme modifiée du salaire, la plus-value doit revêtir, au pôle opposé, la forme modifiée du profit. »*[132]

L'illusion dont sont victimes les personnes prises dans le système de production capitaliste consiste en ce que le profit est produit par l'investissement ; la réalité, c'est qu'il est le produit de la force de travail des ouvriers. Dire que ceux-ci sont exploités, c'est dire que le principe selon lequel chacun devrait recevoir une rétribution identique ou équivalente à la valeur réelle du produit de sa force de travail est violé. Et cela a lieu alors que ce principe est reconnu explicitement par la société capitaliste. Le concept d'illusion objective peut donc se comprendre ainsi. La difficulté consiste à déterminer les conditions de révélation de la réalité des rapports de production.

Le concept d'illusion objective établi par Sen à partir des travaux de Cohen sur Marx est une catégorie générale qui inclut les phénomènes décrits dans le *Capital*, mais pas seulement. Ainsi, l'illusion de l'immobilité de la planète Terre est un autre cas d'illusion objective, bien qu'elle ne soit pas liée à un mode de production. Par ailleurs, l'usage que fait Sen de ce concept fait apparaître des points de divergence avec celui qu'en fait Marx. D'une part, les illusions objectives décrites par Sen ne sont pas liées de façon nécessaire au mode de production capitaliste ou à la structure sociale et politique qui le supporte. D'autre part, Sen rattache une illusion objective donnée à une position sociale donnée, tandis que Marx proposait une analyse en termes de classes économiques et sociales définies par leur rôle dans le système de production.

Néanmoins, les cas d'illusion concernant les travailleurs que nous avons relevés dans le *Capital* font état d'un défaut de perception de légitimité à prétendre à l'équivalent réel du produit de leur travail : le parallèle avec le diagnostic des préférences adaptatives des femmes vivant dans un environnement conservateur est significatif. Le présupposé commun à ces

[132] MARX, Karl, *Das Kapital*, 1867, vol 3. Trad.Fr *Le Capital*, Paris, Gallimard, 1963, chap. I.

deux diagnostics est que l'énonciation de certaines préférences ou les comportements de choix de personnes souffrant de privations, ne peuvent être compris qu'en se référant à l'organisation économique et sociale à laquelle elles appartiennent.

La technologie sociale

Appliquant cette grille de lecture marxienne à l'analyse de la répartition des ressources à l'intérieur de la famille, Sen décrit l'organisation économico-sociale de la famille en termes de « technologie sociale »[133] : l'économie familiale s'adosse toujours à une organisation sociale, de même que la « machinerie » de l'industrie du XIXe siècle s'adosse à une organisation sociale particulière. Marx traite ainsi des interactions entre la mécanisation de la production et la modification des rapports entre les membres d'une famille : l'apparition de la machine à coudre mobilise les forces de travail et le temps des jeunes femmes, avec des effets sur l'organisation familiale. La prospérité d'un ménage ne dépend pas que du travail salarié traditionnellement masculin, mais aussi du travail domestique non rémunéré dévolu aux femmes : la division sexuée du travail repose sur une organisation sociale qui attribue aux femmes les tâches relatives à la sphère domestique et aux hommes les emplois rémunérés et socialement reconnus comme du « travail »[134]. Cette grille d'analyse permet à Sen de mettre en lumière l'importance du travail domestique pour la prospérité familiale, alors que ce travail reste ordinairement invisible, matériellement (il se déroule dans l'espace caché du foyer), économiquement (il n'est pas rémunéré) et socialement (il n'est pas valorisé socialement). Cette grille permet aussi de comprendre comment les inégalités des sexes perdurent : ces inégalités sont des effets directs de cette division du travail : les femmes ont moins de ressources et moins de possibilités de s'épanouir dans une activité qu'elles auraient choisi. Elles en sont aussi des effets indirects : l'absence de reconnaissance sociale pour les travaux domestiques fausse leur perception de la légitimité de leurs demandes.

Ce que Sen appelle « technologie sociale » dans les articles d'économie familiale est l'« assemblage de divers processus en un tout social »[135] ou la « combinaison du procès de production social »[136] dans le *Capital.* Il s'agit de décrire la relation entre le progrès technologique et les rapports sociaux et modes de vie des personnes. Marx évoque ainsi le passage de l'agriculture traditionnelle à l'agriculture industrielle : ce passage rationalise la

[133] MARX, Karl, *Das Kapital*, 1867, vol 1. Trad.Fr. *Le Capital,* PUF Quadrige 2009[3], chap. XIII.

[134] Le tableau du peintre flamand Jan Van Eyck, *Les Epoux Arnolfini* (1434, National Gallery, Londres), illustre cette division : Arnolfini se trouve du côté de la fenêtre, du monde du travail extérieur, tandis que son épouse pose à côté du lit, montrant ainsi que son destin est lié à la sphère domestique.

[135] Traduction de Sophie Marnat

[136] Traduction collective sous la direction de Jean-Pierre Lefebvre

production, mais remplace également le paysan par l'ouvrier agricole salarié, éparpille les travailleurs en accroissant les étendues cultivées, fait disparaître la culture paysanne, rompt le lien entre le paysan et la terre et dégrade la santé des ouvriers. Le passage de l'outil à la machine révolutionne donc l'ensemble de l'organisation sociale.

> *"Le caractère coopératif du procès de travail devient donc maintenant une nécessité technique dictée par la nature du moyen de travail lui-même. »*[137]

En substance, ce n'est pas la machine qui est au service des hommes, mais les hommes qui sont au service de la machine. Cette structuration des rapports sociaux autour de la technologie va jusqu'à la religion : dans le premier chapitre du *Capital,* Marx explique que la religion chrétienne est appropriée à la manière abstraite dont la production capitaliste se rapporte au travail. Marx affirme ainsi l'équivalence entre le « procès social d'existence » et le « procès de production matérielle ».

Le lien entre forces de production et structure économique et sociale ainsi établis, Sen se sert des travaux de Marx et de Cohen pour formuler sa thèse : le travail des femmes au foyer, non reconnu et non rémunéré, participe à la production des richesses dont bénéficie la famille. Mais la comparaison s'arrête là, car la perspective de Sen est normative : il veut signifier qu'il ne peut y avoir de production de richesses et de perception de revenu sans le travail accompli par les femmes, au foyer, et dénoncer le manque de reconnaissance dont elles sont les victimes.[138] Il s'agit non seulement de rendre compte de la réalité « ménagère » et prosaïque des conditions sociales de production de richesses, mais également de penser une autre technologie sociale, avec par exemple une répartition alternative des travaux dévolus traditionnellement aux femmes. En mobilisant le concept de « technologie sociale », Sen montre que cette division sexuée du travail n'a rien de naturel, mais n'est qu'une forme sociale possible et peu souhaitable, étant donnés ses effets réels sur le bien-être et l'épanouissement des femmes.

Cette analyse critique montre ainsi que la « voix », c'est-à-dire l'expression de la préférence, à laquelle la théorie du choix social identifie en principe l'évaluation individuelle, n'est pas un critère d'évaluation complètement satisfaisant de la situation d'un individu. On ne peut

[137] MARX, Karl, *op. cit.*

[138] Bien sûr, l'objectif n'est pas de revaloriser le travail au foyer pour encourager les femmes à y rester: son invisibilité fait toujours courir le risque d'une dévalorisation.

évaluer la situation individuelle de ces personnes sans se référer aux relations qu'elles entretiennent avec les membres de leur famille ou à leur place dans les institutions sociales du lieu où elles vivent.

c) Capabilités et individualisme méthodologique

Les descriptions de situations de domination, d'inhibition à choisir[139] ou d'engagements désintéressés envers autrui remettent en cause l'hypothèse selon laquelle l'individu serait de fait à l'origine des évaluations de l'avantage. Cette hypothèse relève de l'individualisme méthodologique, selon lequel, pour expliquer un phénomène collectif, il faut identifier les comportements des individus supposés à l'origine de ce phénomène.

> *« L'individualisme méthodologique est un paradigme qui affirme qu'une explication sociologique est satisfaisante si elle a identifié les comportements individuels qui sous-tendent un phénomène collectif. »*[140]

Nous voyons d'ores et déjà que, tandis que l'individualisme méthodologique vise à expliquer un phénomène, l'approche par les capabilités vise à l'évaluer.

Sen est donc amené à opposer explicitement l'approche par les capabilités à l'individualisme méthodologique :

> *« Il y a certainement eu des écoles de pensée fondées sur les pensées, les choix et les actions de l'individu, détaché de la société dans laquelle il existe. Non seulement l'approche par les capabilités ne présuppose pas un tel détachement, mais son intérêt pour l'aptitude des personnes à vivre le genre de vie qu'elles ont raison de valoriser intègre les influences sociales à la fois en termes de ce que les personnes valorisent (par exemple, « prendre part à la vie de la communauté ») et en termes d'influences opérant sur leurs valeurs (par exemple, la pertinence du raisonnement public dans l'évaluation individuelle) ».*[141]

L'individualisme méthodologique doit postuler une certaine indépendance des comportements individuels pour être pertinent. Il définira ainsi les phénomènes collectifs comme résultant de la seule agrégation, ou superposition, des comportements individuels[142]. Sen oppose à cette

---

[139] Sen évoque les femmes issues de milieu sexistes qui peinent à exercer des droits pourtant institutionnellement reconnus (tel que celui de travailler hors du foyer ou de porter les vêtements désirés), ou encore la difficulté pour certains chômeurs à faire valoir leurs droits en raison de la honte et de la stigmatisation dont ils pourraient souffrir.

[140] Cf. KEUCHEYAN, Razmig, « Individu et individualisme dans les sciences sociales », *in* LUDWIG, Pascal et PRADEU, Thomas (éds), *L'individu. Perspectives contemporaines.*, Paris : Vrin, 2008.

[141] SEN, Amartya, *The Idea of Justice*, London: Penguin Books, 2009, 11, p 244. [Traduction libre.].

[142] Nous nous appuyons sur l'article de R. Keucheyan, *op. cit.*. La théorie du choix social telle que l'expose Arrow pourrait être taxée d'individualisme méthodologique si elle se proposait comme *modèle d'explication* de phénomènes collectifs réels. En nous référant aux propres textes d'Arrow, nous défendons au contraire une *interprétation normative* de sa théorie : le cadre conceptuel de la théorie du choix social expose ce qui devrait être et non ce qui est : on ne peut sérieusement soutenir que le résultat d'impossibilité d'Arrow est *la preuve de*

approche les « influences sociales », que l'on peut classer en fonction de leur place exacte dans l'évaluation sociale et dans la qualité de leur impact sur le bien-être (*well-being*) et l'épanouissement de l'individu :

l'échec de la démocratie. Néanmoins, l'ambiguïté demeure dans la littérature assez hétérogène et de qualité variable qui existe à propos de l'interprétation dudit théorème.

	Influences sociales positives pour l'individu	**Influences sociales négatives pour l'individu**
Valeurs influencées	*Formation et identification des valeurs par la participation à la vie collective.* *Ex : le rôle du débat public et de l'exercice de la citoyenneté dans la construction des valeurs individuelles*	*Inculcation de valeurs qui nient la pertinence de l'épanouissement de la personne.* *Ex : l'acceptation de la condition des femmes des sociétés sexistes par les femmes elles-mêmes, le refus des chômeurs à percevoir leurs droits*
Influences sur ce que les personnes valorisent (au sens où elles ne recherchent pas qu'un intérêt strictement indépendant de liens sociaux)	*Cas où les personnes réalisent leur bonheur en prenant part à l'existence collective* *Ex : Faire de la participation à la vie politique un des buts d'une existence épanouie*	*Cas où les personnes valorisent le bien-être d'un groupe qui les opprime au détriment de leur propre bien-être* *Ex : les femmes des sociétés sexistes qui approuvent une division du travail et une organisation économique qui leur est défavorable*

Tableau 3

Sen ne nie certes pas que l'individu reste pour lui un déterminant majeur : c'est même une « réalité manifeste ». La posture de Sen peut être comprise en se référant à son pragmatisme philosophique : c'est parce que nos pratiques, nos comportements, notre langage, donnent sens à l'idée que l'individu choisit et agit par lui-même (la grammaire attribue des actions aux sujets), que nous devons accepter cette idée. Sen croit au sens commun. Dès lors, quelle méthode d'approche des jugements individuels permet de concilier la reconnaissance de l'existence des influences sociales avec le privilège individuel du jugement ?

3.1.3. Eléments de méthode : néo-holisme et infra-individualisme

Dans cette section, nous verrons que l'approche de Sen peut être caractérisée de néo-holiste et d'infra-individualiste. Nous avons vu que Sen refuse de considérer l'individu comme indépendant *a priori* de toute influence sociale : en ce sens, il n'est pas atomiste. L'évaluation des situations individuelles par les capabilités met en évidence les interactions sociales et les relations de dépendances entre individus.

Le néo-holisme est un concept élaboré par le philosophe américain Philip Pettit. A l'instar de Sen, Pettit tient la psychologie ordinaire [*folk psychology*] pour valide[143] : nous attribuons quotidiennement des pensées, des choix et des actions aux individus, notre langage et nos pratiques témoignent de ce fait. C'est parce qu'il juge que « nous sommes tels que nous nous apparaissons à nous-mêmes », que « nous ne sommes pas des pions ou des jouets manipulés par des forces collectives », que Pettit s'oppose à une approche collectiviste :

> *« Nous sommes tels que nous apparaissons à nous-mêmes. Nous sommes des centres de pensée, de sentiment et d'action capables de chercher, avec quelque succès, à vivre en conformité avec ce que nous nous accordons généralement à considérer comme les exigences de la raison et de la morale. Nous ne sommes pas des pions ou des jouets manipulés par des forces collectives. […].*
>
> *Le collectivisme est une doctrine exotique et extrémiste, et sous prétexte que je la rejette, on ne doit pas penser que j'accepte toutes les positions qui portent le nom d'individualiste. »*[144]

Pettit distingue individualisme et atomisme. L'individualisme s'oppose au collectivisme, tandis que l'atomisme s'oppose à l'holisme. Pettit réserve donc le terme « individualisme » aux doctrines qui font leur la psychologie du sens commun consistant à attribuer à des individus des pensées, des intentions, des actions. Le terme « collectivisme » est réservé aux doctrines qui stipulent que cette psychologie du sens commun est illusoire ou insuffisante pour caractériser les jugements, choix et actions. La psychologie du sens commun n'est que la manifestation de lois collectives cachées (à découvrir). L'opposition entre individualisme et collectivisme est qualifiée de « verticale » par Pettit, puisqu'il s'agit de déterminer quelles forces déterminent l'action collective, à savoir les forces collectives ou les forces individuelles.

Dans cette analyse de Pettit, la distinction entre atomisme et holisme concerne l'origine et les conditions des jugements, choix et actions individuels plutôt que leur attribution à un agent déterminé. Une doctrine atomiste stipule que l'origine de ces jugements, choix et actions réside exclusivement dans l'individu, indépendamment de toute relation sociale, chaque

[143] PETTIT, Philip, *Penser en société. Essais de métaphysique sociale et de méthodologie.*, trad. de l'anglais par A. Bouvier, B. Guillarme, P. Livet, A. Ogien, Paris : PUF, 2004.
[144] *Idem*, Introduction, p 3 et p 6.

individu étant un « atome », une entité distincte des autres et se suffisant à elle-même[145]. En revanche, une doctrine holiste conçoit, comme son étymologie le laisse deviner, la société comme un « tout » : chaque individu est donc pensé comme appartenant à une société, faisant partie d'un « tout », et ce même si cela fait sens de lui attribuer des jugements, choix et actions. L'opposition entre l'atomisme et l'holisme est qualifiée d'« horizontale » par Pettit, dans la mesure où elle porte sur l'existence, la nature et l'importance des relations des individus entre eux. L'atomisme nie le caractère constitutif des relations entre individu pour l'analyse des faits sociaux. L'holisme peut être défini ici comme la prise en compte de la non-séparabilité des individus et de la complexité des relations sociales dans l'analyse des faits sociaux. Pettit défend l'idée selon laquelle cette position n'est pas incompatible avec une posture individualiste, c'est-à-dire une posture qui reconnaît les individus comme des centres de pensée et des sources réelles de jugements, de choix et d'actions[146]. Afin d'éviter les confusions avec le holisme méthodologique traditionnel, qui intègre le collectivisme, nous parlons de « néo-holisme »[147].

Amartya Sen refuse également de faire sienne la position collectiviste ; il estime que l'analyse des capabilités des groupes, des collectivités ou des communautés n'est que faiblement pertinente pour l'évaluation sociale[148]. Sans se faire l'avocat d'une abstraction qui porterait le nom d'Individu, il n'adhère pas davantage à l'idée d'une Société abstraite pensée indépendamment de ses membres :

> *« La question essentielle fut posée avec une portée et une clarté admirables par Karl Marx il y a plus d'un siècle et demi : 'Ce que nous devons éviter par-dessus tout est le rétablissement de la 'Société' comme une abstraction vis-à-vis de l'individu'. La présence d'individus qui pensent, choisissent et agissent – une réalité manifeste dans le monde – ne rend pas une approche méthodologiquement individualiste. C'est l'invocation illégitime de toute présomption d'indépendance des pensées et des*

[145] Et, *a fortiori,* ces jugements, choix et actions seront attribués à l'individu. On ne peut pas concevoir de doctrine à la fois atomiste et collectiviste (mais on peut concevoir, comme le défend Pettit, une doctrine holiste et individualiste).

[146] Pettit caractérise également son approche d'"anti-singulariste": selon lui, les individus auxquels nous attribuons pensées et actions ne sont pas nécessairement des individus "singuliers", il peut s'agir d'"individus collectifs". Ainsi, certains groupes, tels que les Etats, les associations, peuvent et doivent être considérés comme des individus lorsqu'il est impossible d'obtenir une décision collective cohérente à partir des jugements des membres. Il n'est pas du tout certain que Sen suivrait Pettit ici : dans le chapitre qu'il consacre à la préférence collective[146], il refuse d'attribuer à la société des préférences et une rationalité propre sans passer par le jugement d'un individu singulier qui se prononcerait au nom de la société en question.

[147] Nous faisons ainsi nôtre la terminologie de Razmig KEUCHEYAN, *op. cit.*.

[148] SEN, Amartya, *The Idea of Justice*, London: Penguin Books, 2009, 11, p 245.

actions des personnes à l'égard de la société autour d'elles qui introduirait la bête redoutée dans le salon. » [149]

A la logique substantialiste qui caractérise aussi bien l'individualisme méthodologique que le collectivisme holiste, Sen privilégie une logique de la relation qui s'intéresse de prime abord aux liens entre les individus.

Pour Amartya Sen, la perspective du groupe non seulement cesse de refléter les attentes individuelles que l'expert s'efforce d'identifier, mais surtout nie le fait qu'un même individu puisse appartenir simultanément à des groupes distincts. Cela prive l'individu de la possibilité de cultiver plusieurs appartenances identitaires à la fois, ce qui est un facteur de singularisation individuelle et d'autonomie, ici comprise comme distance critique.

La perspective de Sen peut également être qualifiée d'infra-individualiste, dans la mesure où l'avantage individuel est en partie évalué en termes de fonctionnements. Or, chaque fonctionnement est un état ou une activité qui ne peut à lui seul constituer un individu. Par ailleurs, certains fonctionnements sont communs à plusieurs individus. Ainsi, une même personne pourra, ou non, appartenir au groupe des handicapés, des femmes, etc. La perspective des fonctionnements ne se superpose ni à celle des individus, ni à celle d'un groupe déterminé : ainsi, la classe sociale n'est une catégorie d'analyse pertinente que dans la mesure où elle correspond à la privation d'un fonctionnement donné. Le fonctionnement existe en deçà de l'individualité. La perspective des fonctionnements permet donc de qualifier l'approche de Sen d'infra-individualiste.

[149] *Idem*, p 244.

3.1.4. L'objectivité positionnelle

Dans les sections qui précèdent, nous avons établi que la pensée de Sen ne peut faire sienne la thèse selon laquelle l'individu est la source exclusive des évaluations que recense le choix social. A présent, il nous faut déterminer quelle est cette source, et comment elle peut être décrite. Pour ce faire, nous proposons de recourir au concept de « position » élaboré dans les écrits de Sen sur l' « objectivité positionnelle »[150]. En effet, Sen distingue les évaluations subjectives des évaluations « positionnellement objectives » (et plus généralement les jugements, qu'ils soient de fait ou de valeur, peuvent être dits positionnellement objectifs selon Sen). Une évaluation subjective est une évaluation que l'on attribue à un « sujet », c'est-à-dire à un être doté d'un point de vue et de tendances mentales singulières et ne pouvant être décrits en termes de paramètres extérieurs à ce sujet. Une évaluation positionnellement objective est une évaluation que l'on rapporte dans la mesure du possible aux paramètres de la position qu'occupe ce sujet. Sen définit ainsi les paramètres de position, qu'il distingue des paramètres de localisation:

> *« Les paramètres de position ne sont pas nécessairement des paramètres de localisation (liés à un emplacement dans l'espace), et peuvent inclure toute condition qui (1) peut influencer l'observation et (2) peut s'appliquer à différentes personnes sur le plan paramétrique. Divers types d'exemples de paramètres de position (dans ce sens large) incluent par exemple le fait d'être myope ou daltonien ou d'avoir une vue normale ; de connaître ou non une langue en particulier ; d'avoir ou non la connaissance de certains concepts ; d'être capable ou non de compter. »*[151]

Ces paramètres incluant « tout ce qui peut influencer » l'observation et le jugement, ils sont susceptibles d'englober une série importante de facteurs. Sen évoque ici les aptitudes physiques et cognitives. La suite de l'article cite aussi les facteurs politiques, économiques et sociaux.

La source des jugements de valeur dits « individuels » est donc définie en relation avec son contexte ; si l'on entend par individu un être caractérisé par son unité et son unicité (le fait qu'il se différencie d'autres êtres, par des frontières ou par sa singularité), alors ce n'est pas l'individu qui est à l'origine des jugements de valeur pris en compte dans l'évaluation sociale ; en effet, la position à partir desquels ces jugements sont énoncés n'est pas unie (elle est influencée par des paramètres divers) et encore moins unique (elle peut être partagée par d'autres femmes et d'autres hommes).

[150] SEN, Amartya, « Positional Objectivity », *Philosophy & Public Affairs*, Vol. 22, No. 2, 1993, pp. 126-145.
[151] *Ibidem.*

Un jugement de valeur subjectif est un jugement de valeur qui manifeste les différences de morale entre les individus : en ce sens il est arbitraire. Un jugement de valeur positionnellement objectif est un jugement de valeur qui est relatif à la position de l'évaluateur. L'objectivité positionnelle lie ainsi l'évaluation, dont l'évaluation morale, à la connaissance. Des choix apparemment peu judicieux peuvent être compris (sans être nécessairement approuvés) en ayant accès aux croyances de l'agent. L'évaluation est qualifiée d' « objective » parce que les caractéristiques de la position à partir de laquelle elle est énoncée sont objectivement descriptibles[152].

La distinction entre objectivité positionnelle et subjectivité repose sur le fait que l'individu n'est pas nécessairement le support de la position sociale. En parlant de « subjectivité », en qualifiant un énoncé de « subjectif », nous renvoyons cet énoncé à un individu qui en serait la source, nous faisons de cet énoncé le propre d'un sujet ou de ses opérations mentales. En parlant d' « objectivité positionnelle » pour qualifier les évaluations à la base du choix social, nous nous référons aux caractéristiques de la position qu'occupe l'individu qui exprime le jugement de valeur, et il est logiquement possible qu'un autre individu puisse occuper exactement la même position.

Cette distinction entre la position en relation avec son contexte et l'individu qui subsiste par lui-même implique deux choses pour l'évaluation dite « individuelle » :

- En théorie, un même individu peut occuper plusieurs positions différentes et énoncer donc plusieurs jugements de valeur différents sur un même objet ; et, en théorie, différents individus peuvent occuper une même position et énoncer un même jugement de valeur.
- Il existe une relation de dépendance entre le jugement de valeur et les circonstances dans lesquelles il est produit, entre l'évaluation de la société et son contexte. Il n'y a pas nécessairement d'évaluation sociale formulée à partir de « nulle part », à partir d'un point de vue décontextualisé[153].

La première implication ouvre des perspectives pour les raisonnements moraux, outre le fait qu'elle permet au spécialiste des sciences sociales de comprendre les comportements qu'il observe.

[152] Cf. aussi SEN, Amartya, "Rights and Agency", in *Philosophy and Public Affairs*, 11, 1, 1982, p 3-39 [Trad.Française S. Marnat, *Ethique et Economie*, Paris: PUF, 1993, "Les droits et la question de l'agent», p 150].
[153]Cf. SEN, Amartya, *Ibidem.*

La seconde implication met en difficulté la liberté de l'agent. En effet, de la nature de la relation entre position et jugement dépend le caractère plus ou moins déterministe de la philosophie sociale de Sen. La position influence-t-elle ou détermine-t-elle l'action individuelle ? Si elle la détermine, alors il n'y a plus à proprement parler de subjectivité, ni de liberté de l'agent, si l'on entend par liberté d'agent le pouvoir de déterminer par soi ses choix et ses actions. Sen aborde » la question du déterminisme au conditionnel :

> *« Selon une vision déterministe, on peut montrer que les observations et croyances de toute personne peuvent être entièrement expliquées par une caractérisation adéquate des paramètres positionnels qui influencent son observation et sa compréhension. Si ces paramètres étaient tous spécifiés en tant que partie de l'identification positionnelle, ces observations et croyances seraient positionnellement objectives dans cette situation. Il semblerait donc que chaque vision ou opinion pourrait devenir « positionnellement objective » à l'aide d'une caractérisation très soigneuse des paramètres de position. »*[154]

« Formellement », nous dit Sen, il est possible d'inclure des caractéristiques subjectives dans les paramètres de position : ainsi, des « tendances mentales » ou l' « inexpérience » peuvent être qualifiées de positionnellement objectives. Mais Sen peut contourner cette difficulté en mettant en avant la variété infinie des positions possibles. S'il existe des positions similaires, en revanche il est empiriquement peu probable d'avoir affaire à deux positions identiques. Or, la question du déterminisme peut se formuler ainsi : si deux individus occupaient *exactement* la même position, leurs choix et leurs actions seraient-ils *toujours* les mêmes ? Dans la mesure où deux personnes n'occuperont jamais deux positions strictement identiques dans les faits, la question du déterminisme devient oiseuse ; autrement dit, les variations interindividuelles impliquent un résidu d'opacité de l'évaluation individuelle. La singularité des destins individuels semble devoir plus à la singularité des situations qu'à la volonté subjective. Le cas de deux individus occupant exactement la même position est logiquement possible. Mais cette possibilité logique ne modifie pas le caractère hautement improbable du partage de la même position par deux individus différents. Ainsi, Sen évite de produire une définition complète de chaque position individuelle, se contentant de descriptions partielles mais *suffisantes* pour l'évaluation sociale.

Cependant, Sen souligne que la subjectivité est une « idée sociale importante »[155]. Il déplace en fait le problème : ce qui doit nous intéresser, c'est le contexte d'analyse :

---

[154]SEN, Amartya, *Rationality and Freedom*, Belknap Press of Harvard University Press, 2002 [Trad.Française M.-P. d'Iribane-Jaawane, Odile Jacob, 2005, p 373]

[155] SEN, Amartya, *Ibidem.*

> *« Dans le contexte de l'analyse de préjudices sociaux systématiques, partagée par beaucoup de gens occupant une place similaire dans la communauté, il serait sans doute utile de considérer un phénomène qui possède d'évidentes caractéristiques subjectives comme étant également positionnellement objectif, vu d'une position élaborée de manière appropriée, car cela nous aiderait à nous concentrer sur des liens de causalité qui ont des rôles explicatifs importants. »*[156]

La qualification de « subjectif » ou d' « objectif » dépend du contexte et des buts que l'on se donne. C'est la fécondité heuristique de l'objectivité positionnelle qui justifie son utilisation. Mais Sen ne nie pas qu'un phénomène puisse avoir d'« évidentes » caractéristiques subjectives : le subjectivisme est une « idée sociale importante » ; « le diagnostic d'arbitraire subjectif demeurerait pertinent ». Le concept de subjectivité est opératoire dans d'autres contextes que dans celui de l'« analyse des préjudices sociaux », peut-être lorsque la responsabilité juridique ou morale est en jeu.

Les capabilités et les fonctionnements peuvent être pensés comme caractérisés en fonction des positions occupées par des individus. Les capabilités représenteraient ainsi les potentialités d'activité à partir d'une position décrite comme un ensemble de relations avec le contexte naturel, physiologique, social, économique, culturel.

[156] SEN, Amartya, *op. cit.*, p 374.

3.2. La construction de l'individualité

Dans cette section, le point de vue de la première personne permet de reconstituer les processus d'individuation tels que Sen les thématise. Il apparaît ainsi que la contribution de Sen dans son ensemble ne fait pas de l'individualité une donnée, mais l'aboutissement des processus que nous décrirons dans les sous-sections suivantes.

Nous distinguerons ainsi trois formes d'individuation présentes dans les écrits de Sen : l'individuation par l'identification des privations propres (3.2.1.), l'individuation par la formation des préférences et des valeurs (3.2.2.) et l'individuation par l'acquisition de l'autonomie (3.2.3.).

3.2.1. L'individuation par l'identification des privations

Dans l'approche des capabilités telle que la développe Sen, le traitement des privations dites « naturelles » (comme les handicaps) n'est pas différencié du traitement des privations dites « sociales » (comme le revenu). Pourquoi ? En fait, il semble que l'évaluation sociale développée par Sen s'attache à mettre en évidence la singularité des privations et des attentes de chaque individu.

C'est par ailleurs ici que se situe le désaccord entre les partisans d'une définition de la privation en termes de besoins, ou « besoins de base » (*Basic needs approach*)[157], et ceux de l'approche par les capabilités (*capabilities approach),* parmi lesquels figure Sen. L'analyse de ce désaccord permet de comprendre l'importance de la singularité des privations individuelles pour l'approche des capabilités.

L'approche en termes de « besoins de base » a inspiré l'action de la Banque Mondiale à la fin des années 70 et au début des années 80, action qui s'adressait aux pays pauvres. L'idée était que certains besoins de base devaient être financés en priorité pour les pays pauvres[158]. Sen a formulé plusieurs objections contre cette démarche[159]. Parmi ces objections, deux d'entre

[157] Cf. STREETEN Paul, avec BURKI S. J., UL HAQ M., HICKS N., STEWART F., *First Things First*, Oxford: Oxford University Press for the World Bank, 1981; STEWART, Frances, *Basic Needs in Developing Countries*, Baltimore, Md: John Hopkins University Press, 1985; WIGGINS, David, "Claims of Need", p 1-57 in *Needs, Values, Truth*, Oxford: Blackwell, 1987; READER, Soran, "Does a Basic Needs Approach need Capabilities", *The Journal of Political Philosophy*, Vol. 14, number 3, 2006, pp 337-350.

[158] Indépendamment de la question que nous traitons ici, nous pensons que la manière dont cette approche a été présentée pose problème pour une approche normative, dans la mesure où le fait de distinguer le traitement réservé aux « pays pauvres » de celui des « pays riches » entraîne l'adoption de politiques publiques insuffisamment exigeantes.

[159] SEN, Amartya, *Resources, Values and Development*, Oxford: Basil Blackwell, 1984, pp 513-515.

elles nous intéressent tout particulièrement, car elles sont liées à la question de l'individuation par la description de privations singulières :

(1) Les besoins sont définis indépendamment des contingences sociales et individuelles, tandis que les capabilités prennent en compte la spécificité de la position sociale de l'individu et de ses caractéristiques propres.

(2) Le besoin est un concept trop « passif » ; en revanche, les capabilités sont définies comme des libertés réelles et reconnaissent l'importance de l'exercice de la qualité d'agent.

La philosophe Soran Reader a discuté les objections de Sen à l'approche en termes de besoins. Elle souligne que cette dernière, tout comme l'approche par les capabilités, ne définit pas la pauvreté comme une privation de ressources, mais de fonctionnements.

> *"Un être humain peut avoir besoin de [verbe] [quelque chose] : faire de la philosophie, se reposer dans le silence, perdre des tissus adipeux, conserver toutes ses dents, parler sa langue natale, entretenir une mémoire, être accueilli par une communauté, être laissé seul, partager la terre, avoir une idée, subir une punition, pouvoir pleurer un bon coup. »*[160]

Soran Reader distingue ensuite les besoins dispositionnels et les besoins occurrents. Les besoins dispositionnels sont propres à l'être qui éprouve le besoin. Tout être vivant a des besoins dispositionnels : les êtres humains riches aussi bien que les êtres humains pauvres ont par nature le besoin de s'alimenter. Les besoins occurrents correspondent à l'absence de satisfaction des besoins dispositionnels. Par exemple, un être humain riche ne souffre pas du besoin de nourriture, tandis qu'un être humain pauvre est susceptible d'en souffrir. La définition des besoins dispositionnels nous fournit le critère du besoin : ce critère est fonction du fait que nous sommes des êtres humains. Reader souligne que « être humain » ne se résume pas aux aspects biologiques de la nature humaine : alimenter une personne de force, ou avec une nourriture qu'elle ne souhaite pas ingérer, ce n'est pas satisfaire un besoin. Reader laisse la définition de ce que sont les besoins humains à la « raison pratique publique » de la communauté au sein de laquelle vit l'être qui participe à cette raison.

L'argumentation de Reader est pertinente, mais on peut se demander si le terme « besoin » est approprié pour qualifier la description du désavantage qu'elle défend. Dans l'usage courant, le terme « besoin » tend à désigner des manques qui dépendent de la constitution de l'être humain indépendamment de toute définition politique. Pourquoi Reader s'attache-t-elle alors à conserver le concept de « besoin » ? Peut-être est-ce parce que ce concept permet de tracer

[160] READER, Soran, *op. cit.* Traduction libre..

une limite entre le nécessaire et le superflu, même si, au lieu d'être définie en termes biologiques, cette limite est laissée à l'appréciation de la raison pratique publique., l'objection mérite réflexion. Soran Reader est consciente de la difficulté et soutient que cette tâche doit être laissée à la communauté dans laquelle vit l'individu, sans apporter de précisions supplémentaires (s'agit-il des institutions de l'Etat-Nation, d'autres structures sociales, de la communauté religieuse, de la famille ?), si ce n'est qu'elle laisse entendre que cette communauté est organisée démocratiquement. Les défenseurs du besoin partagent dès lors avec ceux des capabilités les problèmes de choix social. Soran Reader avance un argument supplémentaire contre l'objection de paternalisme : si la définition de l'identité d'un être humain comprend la liberté comme composante propre, alors la détermination des besoins prendra en compte cette liberté. Elle postule dès lors implicitement que toute communauté décidant des besoins de ses membres en fonction de leurs identités reconnaît la liberté comme composante de ces identités. En d'autres termes, l'approche en termes de besoin dans sa variante « libérale » stipule que la liberté est une propriété de l'être humain.

L'approche par les capabilités qualifie de « libre » ou de « non-libre » la détermination des fins et le mode de réalisation de ces dernières par un être humain, ou un groupe d'êtres humains ; elle n'attribue pas la liberté aux êtres humains en tant qu'entités. Paradoxalement[161], en dehors de la liberté, S. Reader estime que la définition du contenu des besoins doit être laissée à la raison publique pratique de la communauté, et variera sans doute selon les communautés, leur culture ou leur niveau de développement. En revanche, la distinction entre le nécessaire et le non-nécessaire, propre au besoin, demeure pour chaque liste communautaire de besoins. Afin de montrer que l'approche en termes de besoins n'est pas restrictive, Soran Reader rappelle que la délibération rationnelle et publique peut étendre la liste des besoins en fonction du degré de développement de la communauté : si les besoins vitaux sont satisfaits, celle-ci peut également pourvoir, par exemple, aux besoins de musique de ses membres. Mais, quelques paragraphes plus haut, elle avance que l'un des avantages du besoin par rapport à la capabilité est son réalisme : évoquant les besoins vitaux, elle montre que les théoriciens du besoin sont plus que tous ancrés dans la réalité matérielle des conditions d'existence des êtres humains.

A travers les différents exemples d'identités définies par la délibération publique, nous observons un glissement sémantique d' « être vivant » à « être humain » puis à « musicien ». Une société plus avancée peut satisfaire le besoin de faire de la musique, tandis qu'une société

[161] Le paradoxe en question est, nous semble-t-il, dû à ce que la définition de la raison publique pratique n'a pas été clairement établie dans cet article. Il n'est donc pas insurmontable.

plus pauvre pourvoira d'abord au besoin de s'alimenter. Mais comment être certain que cette société classera la satisfaction du besoin de s'alimenter en priorité par rapport à celui de jouer d'un instrument de musique ? Si la majorité de ses membres peut s'alimenter parfaitement grâce à une économie qui ne dépend pas des décisions sociales, elle classera peut-être comme prioritaire le besoin de musique, au détriment d'une minorité qui meurt de faim. Elle établira donc que l'identité de ses membres est une identité de musicien et non d'être subsistant par la nutrition. En revanche, les capabilités étant des ensembles de fonctionnements, une politique publique favorisant les capabilités vise à étendre cet ensemble pour chaque membre. Or, le fonctionnement « jouer de la musique » ne peut être inclus dans un ensemble capabilité qui ne comprend pas aussi le fonctionnement « manger à sa faim » (un minimum d'énergie est requis pour frapper les touches d'un piano). L'étendue de l'ensemble capabilité dépend plus fortement du pouvoir de s'alimenter que de celui de jouer de la musique. Cet exemple souligne un autre effet fâcheux de l'approche en termes de besoin : si, dans un même monde, coexistent une société qui a su satisfaire le besoin de s'alimenter et une autre qui n'a pas pu le faire, la première société pourra être amenée à définir comme « besoin », comme nécessaire donc, le fait de jouer de la musique (au moins pour certains de ses membres), tandis que la seconde le jugera superflu, préférant définir comme nécessaire celui de s'alimenter. Cela implique qu'un certain confort de vie paraîtra nécessaire à des communautés riches et superflu à des communautés pauvres.

L'approche par les capabilités voit donc la liberté dans la réalisation d'une fin, que ce soit le bien-être de l'agent, ou ce qu'il détermine librement par une démarche rationnelle de réflexion, et non dans la définition d'une identité. Elle accorde ainsi toute son attention à la singularité des privations et des attentes, afin de déterminer ce qui manque à l'agent pour réaliser les fins qu'il a choisies, va choisir ou choisirait si sa position était meilleure. Parmi ce qui manque figure parfois une conscience claire de ce manque et de cette fin : l'accroissement des capabilités correspond donc à un moment donné à cette prise de conscience, moment où l'évaluation des privations peut commencer à se faire à la première personne.

Les travaux de Sen en économie de la famille mettent également en évidence son intérêt pour la singularité des privations. Sen refuse de considérer la famille comme un seul agent (les « ménages ») ou comme un ensemble d'individus dont les attentes sont identiques. Ce choix méthodologique s'appuie sur l'observation de différences de répartition des ressources entre les membres d'une même famille, différences qui ne se justifient pas par une différence de

besoins. Ces ressources peuvent être aussi bien des aliments que des ressources éducatives, du temps de loisir (dont sont privés les membres de la famille qui doivent assurer la majorité ou la totalité des tâches ménagères), du pouvoir de décision.

De façon significative, Sen commence par identifier les privations en termes de santé et d'espérance de vie pour mener sa critique des autres approches de l'économie familiale. Il rapporte ainsi une étude comparative de la malnutrition des enfants selon leur sexe dans deux villages indiens voisins ; les petites filles du village bénéficiant d'une distribution *publique* de nourriture ne souffrent pas d'inégalité de répartition et donc des pathologies liées à la malnutrition[162]. Ce sont donc les privations liées à la survie de l'organisme humain qui incitent dans un premier temps le théoricien à s'intéresser aux attentes individuelles, y compris à l'intérieur d'une collectivité telle que la famille. On ne peut mettre de côté la singularité de ces attentes :

> *« La spécification des soi-disant « besoins » en calories et autres nutriments n'est guère fondée scientifiquement, en raison de facteurs tels que les variations interpersonnelles de métabolisme, le rôle des maladies parasitaires qui affectent l'absorption des aliments, et l'existence de multiples équilibres pour la même personne. En outre, les représentations des besoins communément utilisées semblent avoir intégré des partis pris systématiques. Dans les diagrammes des besoins spécifiés par le très distingué Comité d'Expert de la FAO et de l'OMS, diagramme largement utilisé, l'énergie consommée dans les tâches ménagères est supposée très faible, puisque les tâches ménagères sont considérées comme une activité sédentaire. On se demande où de si agréables tâches ménagères ont été observées. Peut-être dans un appartement de Manhattan plein de gadgets, bien préparé au vingt-et-unième siècle. »*[163]

L'analyse des privations suscite l'intérêt du chercheur pour l'histoire individuelle. Chaque organisme humain a des attentes singulières mises en évidence par les comparaisons interpersonnelles et l'évaluation sociale. Mais chaque organisme humain se singularise également par son histoire personnelle, et il appartient au théoricien de mesurer l'adéquation de cet organisme individuel en évolution avec son environnement.

> *« Une fonction de bien-être (welfare) familial beaucoup plus claire serait donc requise pour rapporter la collection des niveaux de bien-être (well-being) non-assimilés des membres de la famille à une mesure agrégée pour la famille dans son ensemble. Bien sûr, cela impliquera un « mini-problème de choix social ». Bien que cela puisse soulever des difficultés propres, le problème ne peut pas être évité en l'absence de l'une ou l'autre des hypothèses précédemment examinées. »*

[162] SEN, Amartya, "Economics and the Family", *Resources, Values and Development*, Harvard University Press, 1984, ch. 16, p 380-381. Traduction libre.
[163] *Ibidem*, p 382.

Les inégalités de bien-être au sein de la famille suffisent à invalider les approches de la famille comme un tout : le problème du choix social, qui confronte les données individuelles à la nécessité d'une décision collective, peut et doit apparaître déjà au niveau de la famille. Dans certains cas, il en va de l'existence même de certains de ses membres.

Cette perspective identifie l'individu humain à son organisme. Cette « individuation du sens commun » nous paraît naturelle, mais il est important de noter que l'identification de l'individu à un organisme n'est pas valable pour l'ensemble du vivant.

3.2.2. L'individuation par la formation des valeurs

L'identification des privations constitue un premier mode d'individuation dans les écrits de Sen. Mais Sen fait également droit à la formation des valeurs. Cette formation est un processus individuel.

Dans ce processus de constitution, le raisonnement joue un rôle central. Sen utilise la notion d' « examen critique » [*critical scrutiny*] afin de circonscrire deux facteurs d'individuation : la vie sociale et le caractère processuel.

> *« Une personne n'est pas seulement une entité qui peut apprécier sa propre consommation, vivre au jour le jour, se réjouir de son propre bien-être et avoir des objectifs, c'est aussi une entité qui peut examiner ses propres valeurs et objectifs et choisir en fonction de ceux-ci. […] Nous pouvons être ou ne pas être touchés par des préoccupations morales et des motivations sociales, mais aucune de ces attitudes n'est interdite par le fait de se poser des questions, par la formation de nos valeurs et par le fait de réviser nos objectifs, si nécessaire, à cette lumière. »*[164]

Quelles que soient les fins que la personne se donne, elle reste considérée comme capable de poursuivre un examen rationnel de ces dernières sur le long terme. « Le fait de se poser des questions », « la formation de nos valeurs », « le fait de réviser nos objectifs » se réfèrent à un exercice continu du raisonnement, à l'idée selon laquelle la rationalité est d'abord une démarche, démarche qui comprend la confrontation avec les raisons d'autrui : se poser des questions, entreprendre un dialogue avec soi, c'est répéter ce que nous avons d'abord fait avec les autres.

L'évaluation individuelle est une démarche.

> *« La conception de la rationalité développée dans cette introduction et dans certains chapitres (notamment les chapitres 3 à 7) considère la rationalité comme une démarche, et non comme une belle formule ou comme une doctrine essentialiste. La rationalité inclut l'utilisation du raisonnement*

[164]SEN, Amartya, *Rationality and Freedom*, Belknap Press of Harvard University Press, 2002 [Trad.Française M.-P. d'Iribane-Jaawane, Odile Jacob, 2005, p 38]

pour comprendre et évaluer les buts et les valeurs en vue de faire des choix systématiques. La croyance selon laquelle la poursuite de buts spécifiés à l'avance est censée représenter l'essence de la rationalité est contrecarrée, ce qui implique le rejet du statut supposé permanent de l'intérêt personnel comme maître exclusif du comportement rationnel. »[165]

Sen refuse de faire de la rationalité un ensemble d'axiomes, contrairement à ce que pouvait proposer Arrow dans son théorème d'impossibilité[166] : Arrow définissait les évaluations rationnelles comme des préférences représentées par des relations d'ordre conformes aux axiomes de complétude et de transitivité. La rationalité que propose Sen n'est pas une propriété logique (« une belle formule »), ni par ailleurs une propriété ontologique (« une doctrine essentialiste »). La rationalité n'est pas un attribut de l'individu au sens ontologique du terme, c'est-à-dire de l'individu comme être réel, singulier et existant par soi, qui serait rationnel par nature. Le critère permettant de distinguer une évaluation comme rationnelle ou non est l' « utilisation du raisonnement » : Sen s'inscrit ici dans une approche pragmatiste de la rationalité, et Hilary Putnam a pu noter la parenté entre les arguments de l'économiste indien en faveur de sa conception de l'évaluation sociale et l'approche défendue par la tradition pragmatiste en philosophie morale[167]. Il n'y a en effet pas dans l'approche de Sen de définition de la rationalité antérieure à la pratique rationnelle, pratique que l'on peut décrire. Or, la formation des préférences et des valeurs – et donc le processus d'individuation – fait partie de cette pratique rationnelle.

La rationalité comme pratique est inscrite dans la vie sociale et politique. Nous avons souligné le rôle de l'éducation, des échanges et de la participation politique dans la formation des jugements de valeur individuels.

« Le rôle du débat public dans la formation des préférences et des valeurs est central »[168]

A la question de savoir si une conception arrêtée de l'individu est une condition préalable de la démocratie, Sen répond en renversant les termes du problème. C'est la pratique démocratique, englobant dialogues, discussions et débats, qui est une condition préalable de l'individuation : autrement dit, les échanges de raisons avec une collectivité humaine qui soit

[165] *Ibidem*, p 45.

[166] « AXIOME DE COMPLETUDE : Pour tout x et y, soit xRy, soit yRx. »

« AXIOME DE TRANSITIVITE : Pour tout x, y, z : xRy et yRz impliquent xRz »

(ARROW, Kenneth, *Social Choice and Individual Values*, New Haven, Yale University Press, 1951 (2ème édition révisée, 1961).Traduction française par l'Association de Traduction Economique de l'Université de Montpellier, Paris : Calmann-Levy, 1974, Chapitre 1, p 38.)

[167] Cf. PUTNAM, Hilary, *The Collapse of the Fact/Value Dichotomy*, Harvard University Press, 2004, I, 3 et II, 5. Traduit de l'anglais (Etats-Unis) par M. Caveribère et J.-P. Cometti, *Fait/Valeur : la fin d'un dogme et autres essais*, Paris-Tel Aviv : Editions de l'Eclat, 2004.

[168] SEN, Amartya, *Op. cit.*, p 221.

la plus étendue possible contribuent de façon déterminante à l'élaboration de jugements de valeur individuels susceptibles de résister à un examen critique. Cette défense instrumentale de la démocratie n'est pas sans évoquer la « justification épistémologique de la démocratie » que l'on trouve chez les pragmatistes classiques et que Putnam réactualise, justification selon laquelle les pratiques qui caractérisent la démocratie, telles que l'expression libre de la pensée et de l'opinion, la confrontation des vues, la critique possible de ce qui fait autorité, le raisonnement en collectivité et le débat, favorisent la progression des savoirs[169]. Les pratiques démocratiques contribuent au développement de la rationalité et donc à l'élaboration de discours raisonnés et d'évaluations sociales rationnelles.

Ajoutons que, pour Sen, juger et choisir par soi-même peuvent être des plaisirs en soi : la qualité d'agent constitue fréquemment l'un des facteurs déterminants d'une existence réussie.

[169] Cf. PUTNAM, Hilary, *Renewing Philosophy*, Cambridge (Massachussetts) : Harvard University Press, 1992, p 180, et TIERCELIN, Claudine, *Hilary Putnam, l'héritage pragmatiste*, Paris: Presses Universitaires de France, "Philosophies", 2002, p 89.

3.2.3. L'autonomie par la révision des préférences

Sen développé une conception et une formulation de l'autonomie dans les termes de la théorie du choix social : l'autonomie est en effet comprise comme le fait de statuer sur la valeur ses propres préférences et de les réviser. L'autonomie correspond donc à une démarche réflexive[170].

Etymologiquement, est autonome celui qui se donne ses propres lois. La tradition kantienne définit l'autonomie comme le propre de tout être raisonnable ; Rawls conservera le concept de « personne raisonnable » autonome, dans la mesure où cette existence doit être postulée pour que le système de Rawls et l'exigence de justice aient un sens[171] :

> *« Les personnes [sont] considérées comme libres et égales parce qu'elles possèdent à un degré suffisant les deux facultés constituant la personnalité morale, le sens de la justice et la conception du bien. »*[172]

L'autonomie de la personne rawlsienne équivaut ainsi à la possession de facultés déterminées. La procédure d'évaluation sociale défendue par Rawls requiert donc l'existence, postulée ou avérée, d'une catégorie[173] ayant les propriétés qui définissent ce qu'est la personne raisonnable (dont l'autonomie).

Cette définition rawlsienne permet de mettre en évidence les spécificités de la conception de l'autonomie dans l'approche de Sen.

L'originalité tient à la définition formelle de l'autonomie qu'il propose.

> *« En effet, la pluralité des préférences peut être liée de près à la question de l'autonomie d'une personne. L'autonomie peut avoir des implications variées en attirant une plus grande attention et un plus grand respect pour la possibilité de classements de préférences alternatifs qu'on puisse associer à la même personne. Car l'autonomie d'une personne peut être pertinente pour l'évaluation de l'opportunité de plusieurs manières. D'abord, on peut montrer qu'une personne peut avoir quelque chose à dire sur le statut de ses propres préférences […]. En second lieu, la personne peut retenir la liberté de réviser ses préférences quand et comme elle le désire (et quand et comme elle est capable d'y parvenir) […]. Troisièmement, qu'une personne soit ou non capable de réviser ses préférences,*

---

[170] SEN, Amartya, *Rationality and Freedom*, Belknap Press of Harvard University Press, 2002 [Trad.Française M.-P. d'Iribane-Jaawane, Odile Jacob, 2005, p 480]

[171] Cet aspect de sa théorie est moins présent dans la *Théorie de la Justice.* C'est à partir des années 80 qu'il prend de l'importance.

[172] RAWLS, John, *Political Liberalism*, New York: Columbia University Press, 1994, I, 5. Traduit de l'américain par C. Audard, *Libéralisme politique*, Paris: Presses Universitaires de France, 1995, p 61. Nous soulignons.

[173] Sur la distinction entre catégorisation et caractérisation, cf. SEN, Amartya, *The Idea of Justice,* London : Penguin Books, 2009, p 44.

elle peut avoir de bonnes raisons de ne pas apprécier que les autres considèrent ses préférences comme « données » […]. »[174]

La possibilité d'associer plusieurs classements de préférences alternatifs à une même personne permet d'échapper à la réduction de l'individu à la préférence qui menaçait la théorie du choix social. Les comparaisons entre les différents classements qui semblent pertinents pour l'agent peuvent être comprises comme des processus de *constitution* de l'individualité.

L'autonomie est également conçue comme « liberté de réviser ses préférences » ; Sen rapporte ainsi l'autonomie à l'histoire personnelle et à l'activité ; l'autonomie n'est pas le propre des personnes dites « autonomes », mais plutôt la caractérisation de la pratique qui consiste à réviser ses préférences et à les confronter à un examen raisonné.

Le choix d'un certain type d'existence et la connaissance des fonctionnements qui la constituent rendent indispensables la liberté de réviser ses préférences. En effet, la conception de la vie bonne d'un individu, ainsi que les informations dont il dispose pour la réaliser, évoluent ; il lui faut donc pouvoir réviser ses préférences afin que celles-ci correspondent à ces évolutions. Mais le pouvoir de réviser ses préférences, même s'il n'est pas exercé, permet également d'éviter que « les autres [ne] considèrent ses préférences comme « données » ». Considérer les préférences comme « données » c'est réduire l'individu à un classement de préférences et occulter la distance critique qu'il peut prendre à l'égard de celles-ci.

En conservant la représentation formelle du choix, Sen montre que les présupposés de la théorie économique ne sont pas surdéterminés par un mode d'expression qui entretient une proximité certaine avec les sciences dures et peut être garant institutionnellement d'une certaine rigueur. Une théorie économique peut à la fois se servir de la formalisation et se libérer d'une conception atomiste et mécanique du monde. Sen met ainsi en place le concept de « métaclassement » qui offre une illustration formelle du concept philosophique d'autonomie. On parlera de métaclassements de préférences (classements de classements de préférences) et de métaclassements d'opportunités ou métaopportunité[175]. Une métaopportunité est l'ensemble des différents ensembles d'opportunités disponibles, chaque ensemble d'opportunités se référant à une dimension, voire à une identité de l'individu. La valeur d'une métaopportunité est fonction des métaclassements de préférence, ceux-ci

[174] *Ibidem,* p 482.
[175] *Ibidem,* p 520.

peuvent être incomplets. Une conception de l'autonomie comme démarche réflexive sous-tend les concepts techniques de métaclassement et de métaopportunité.

L'autonomie n'est donc pas une propriété, mais un certain type d'activité, plus précisément la démarche qui consiste à statuer sur les différentes options et les différents classements d'options qui nous paraissent accessibles et raisonnables, à les adapter progressivement à ce que nous jugeons être la conception de la vie bonne, et enfin à s'approprier cette démarche même. Sen fait de l'autonomie un exercice de comparaison raisonnée et non une faculté conçue *a priori.*

Si on considère que la philosophie morale et sociale de Sen repose sur l'idée que l'individu est le fruit d'un processus d'individuation, alors on peut éviter la contradiction entre la reconnaissance de la « qualité d'agent » et de l'autonomie des évaluations individuelles, d'une part, et l'assertion selon laquelle certaines évaluations énoncées par l'individu sont biaisées par sa position sociale. En effet, le processus d'individuation part de ce qui est déjà là, de la position de l'agent (encore sujet d'observation) pour aboutir à la possibilité de statuer sur ce que sont les fins propres à l'individu et les moyens les plus adéquats pour y parvenir. Ce processus est pour ainsi dire inachevé : il est constitué principalement de comparaisons, comparaisons entre des positions sociales différentes (des descriptions d'états sociaux incluant la personne concernée et la manière dont ces états sociaux ont été construits). Le point de vue de la troisième personne est une étape de ce processus. La capabilité est le critère d'évaluation sociale qui convient à cette conception de l'individu. La réalisation est la modalité d'existence individuelle que Sen privilégie pour l'évaluation sociale.

Se pose dès lors la question du rapport entre les hypothèses empiriques nécessaires pour attester de l'importance des processus d'individuation et le cadre analytique de la théorie du choix social.

3.3. La place de l'individu dans le choix social senien

Dans cette section, nous analyserons la place de l'individu dans la contribution de Sen consacrée à la théorie du choix social. Il apparaît que, tant dans sa contribution formelle (le théorème d'impossibilité du libéral parétien que nous allons présenter), que dans l'interprétation qu'il en livre, Sen s'est toujours soucié de la mise en relation de la réalisation des intérêts individus avec l'environnement social.

Nous analyserons dans un premier temps le traitement que réserve Sen à un paradoxe de la théorie du choix social, le conflit Pareto-liberté (3.3.1.). Dans un deuxième temps, nous

mettrons en évidence le concept de « liberté indirecte » que Sen développe à partir du cadre même de la théorie du choix social ; ce concept permet de penser la réalisation des fins individuelles à l'aide de l'environnement social (3.3.2.). Dans un dernier temps, nous mettrons en relation la description des évaluations en termes positionnels avec la question de la détermination d'une évaluation collective transpositionnelle (3.3.3.).

3.3.1. Le conflit Pareto-liberté : libertés individuelles et préférence sociale

En 1970, Sen publie l'un des plus importants résultats de la théorie du choix social, le théorème d'impossibilité du libéral parétien[176] Ce théorème démontre que trois conditions de choix collectif apparemment raisonnables, la condition de non-restriction du domaine des préférences, la condition de Pareto et la condition de liberté minimale, que nous allons présenter ci-après, sont incompatibles : il n'existe pas de règle de choix social satisfaisant ces trois conditions simultanément.

Les deux premières conditions sont les mêmes que celle du théorème d'Arrow. La condition de domaine non-restreint (U) stipule que l'on ne peut retreindre *a priori* le domaine des préférences possibles de chaque individu : logiquement, chaque individu peut avoir n'importe quelle préférence, dès lors que celle-ci est cohérente (transitive), complète (elle se prononce sur chaque alternative) et réflexive, autrement dit, dès lors qu'elle correspond à sa propre définition.

La condition dite de Pareto (P) énonce, dans sa version faible, que si tous les membres d'une société préfèrent strictement une option à une autre, alors la société préfèrera également cette option[177].

---

[176] Le résultat a été présenté pour la première fois dans: SEN, Amartya, "The Impossibility of a Paretian Liberal" in *Journal of Political Economy*, 78, 1970), pp 152-157. Cf. Annexe 4.

[177] Le principe de Pareto fort exige que si au moins un individu préfère strictement une option à une autre, les autres étant indifférent, alors la société préfère cette option.

Le principe de Pareto-indifférence stipule que, si tous les membres de la société sont indifférents entre deux options, alors la société est aussi indifférente.

Pour X = {x, y, z…} l'ensemble des alternatives envisageables, R la relation de préférence large (« préfère ou est indifférent à »), P la relation de préférence stricte (« préfère ») et I la relation d'indifférence (« est indifférent à »), et une société composée de n individus, un individu i pouvant exprimer des relations de préférence R_i, P_i et I_i :

PRINCIPE DE PARETO FAIBLE : $\forall x, y \in X, [\forall i, xP_iy] \Rightarrow xPy$

PRINCIPE DE PARETO FORT : $\forall x, y \in X, [\forall i, xR_iy] \Rightarrow xRy$; et $\forall x, y \in X, [\forall i, xR_iy$ et $\exists i, xP_iy] \Rightarrow xPy$

PRINCIPE DE PARETO-INDIFFERENCE : $\forall x, y \in X, [\forall i, xI_iy] \Rightarrow xIy$.

D'après FLEURBAEY, Marc, *Théories économiques de la justice*, Paris : Economica, 1996.

La condition de Liberté Minimale (LM) suppose qu'au moins deux membres de la société ont une sphère personnelle dans laquelle leurs préférences individuelles sont déterminantes[178].

Les conditions sont définies de la manière la plus faible possible, afin que le résultat soit le plus général possible. Par exemple, il est improbable qu'une société qui n'accorde qu'à deux de ses membres le droit d'être décisif pour leur sphère personnelle puisse être qualifiée de « libérale ». Mais, si le résultat est valide pour une société où seuls deux individus jouissent de ce droit, alors il l'est *a fortiori* pour une société où chacun des membres jouit de ce droit.

Sen illustre ce théorème par l'exemple, bien connu des spécialistes, du prude, du libertin et de la lecture de l'*Amant de Lady Chatterley*[179]. Nous choisirons ici un exemple qui, à notre sens, montre avec plus d'acuité l'intérêt de cette démonstration pour penser les difficultés des démocraties libérales contemporaines. En effet, le conflit Pareto-liberté n'est pas un simple casse-tête pour occuper les heures perdues des mathématiciens de la décision. Il se rapporte à des problèmes d'organisation de la vie collective bien réels. Lorsqu'une société choisit d'interdire la réédition de *Mein Kampf*, la consommation et la vente de stupéfiants ou encore la prostitution, elle interfère effectivement avec les choix de certains de ses membres. Supposons donc que la société doive choisir de rééditer ou non *Mein Kampf*[180]. Les deux membres de la société qui jouissent du pouvoir de déterminer leur sphère personnelle sont un néonazi (N) et un antifasciste (A). Le néonazi aimerait pouvoir lire ce livre, mais apprécierait encore davantage que l'antifasciste soit obligé de le lire : ainsi, pense-t-il, il sera ou « converti », ou choqué, ces deux perspectives lui étant agréables. L'antifasciste ne voudrait sans doute pas lire *Mein Kampf*, mais serait encore plus ennuyé de savoir que l'influençable néonazi puisse avoir accès à cet ouvrage sulfureux. Leurs préférences sont donc les suivantes :

Ordre de préférences du <u>Néonazi (N)</u>		**Ordre de préférences de l'<u>Antifasciste (A)</u>**	
« pour moi » (N)	**« pour lui » (A)**	**« pour lui » (N)**	**« pour moi » (A)**
1)Lire Mein Kampf	1)Lire Mein Kampf	1)Ne pas lire Mein K.	1)Ne pas lire Mein Kampf

[178] Un version plus forte de cette condition, la condition de Liberté L, stipule que *chaque* membre de la société est déterminant (décisif) pour sa sphère personnelle.

[179] L'*Amant de Lady Chatterley* est un roman de D. H. Lawrence, écrit en 1928. Il qui décrit des relations sexuelles entre une aristocrate et son garde-chasse et a fait scandale à l'époque de sa parution.

Dans l'exemple de Sen, le prude non seulement refuse de lire le livre mais veut aussi empêcher le libertin de le faire (et serait même prêt à lire le livre pour arriver à ses fins). Le libertin, quant à lui, a envie de lire ce roman, mais trouve surtout particulièrement jouissif d'obliger le prude à le faire (même si pour cela il devait lui se priver de cette lecture).

[180] *Mein Kampf* ferait partie des meilleures ventes des libraires dans le monde arabe. L'ouvrage est vendu sans préface ni avertissement, et la lecture se fait hors contexte historique : l'ouvrage est « réinterprété » à l'aune du conflit israélo-palestinien. Ce fait montre que le conflit Pareto-liberté n'est pas qu'un paradoxe amusant, mais, comme bien des résultats de la théorie du choix social, manifeste une difficulté réelle de la vie collective.

2)Ne pas lire Mein Kampf 3)Lire Mein Kampf 4)Ne pas lire Mein Kampf	2)Lire Mein Kampf 3)Ne pas lire Mein K. 4)Ne pas lire Mein K.	2)Ne pas lire Mein K. 3)Lire Mein Kampf 4)Lire Main Kampf	2)Lire Mein Kampf 3)Ne pas lire Mein Kampf 4)Lire Mein Kampf

Tableau 4

Les ordres de préférence qui devraient être pris en compte dans le choix social selon a condition de Pareto sont marqués en rouge. Les ordres de préférence qui devraient être pris en compte dans le choix social selon la condition de Liberté Minimale sont marqués en vert.

D'après la condition de Pareto, la situation où N ne lit pas *Mein Kampf* et A le lit domine celle où A ne lit pas ce livre tandis que N le lit (en rouge dans le tableau ci-dessus). Mais la condition de liberté minimale commande que chacun des protagonistes soit décisif pour ses lectures (qui font partie de leurs sphères personnelles respectives), et son respect implique que SN lira *Mein Kampf* tandis que A s'en abstiendra (en vert dans le tableau). Les deux conditions sont donc incompatibles.

Ce résultat peut être interprété comme une remise en cause du parétianisme, la mise en évidence des limites du libéralisme lui-même ou encore la preuve de la nécessité de filtrer en amont les préférences afin d'en exclure les plus antisociales. Sen ne tranche pas en faveur de l'une ou l'autre des interprétations de ce résultat. C'est la diversité des champs d'application et des possibilités de réflexion du résultat d'impossibilité qui intéresse Sen, plus qu'une solution elle-même.

La notion de « préférence collective » constitue donc l'un des enjeux des critiques du conflit Pareto-liberté ; ces critiques ne portent certes pas sur les aspects formels du problème, mais sur sa portée en philosophie politique. L'intérêt de ces critiques est de montrer la divergence de vue entre certaines acceptions en philosophie politique et les conceptions de Sen qu'anticipe le cadre conceptuel de la théorie du choix social.

a) La critique libertarienne de la liberté comme préférence décisive

Ainsi, Robert Nozick met à profit le résultat d'impossibilité pour critiquer la conception des droits et de la liberté individuelle sous-tendue par le « conflit Pareto-liberté ». Nozick a été un

des principaux représentants du courant libertarien de droite. Le libertarisme (*libertarianism*) soutient que chaque personne est pleinement propriétaire d'elle-même (dans la lignée de Locke). Le libertarisme de droite dérive de ce postulat de la propriété de soi le droit de s'approprier les ressources naturelles sans le consentement des autres personnes[181], tandis que le libertarisme de gauche exige soit ce consentement, soit un partage égalitaire de tout ou partie de ces ressources[182].

De cette très brève présentation des thèses libertariennes, on comprend aisément l'importance que ce courant accorde aux droits individuels et à une définition appropriée de ces derniers. Nozick n'a donc pas ignoré le résultat de Sen et y consacre une section de son livre *Anarchie, Etat et Utopie*. Pour Nozick, le conflit Pareto-liberté repose sur une définition erronée des droits :

> *« Les problèmes commencent lorsqu'on assimile le droit d'un individu à choisir entre des solutions différentes au droit à décider de la hiérarchie relative de ces choix à l'intérieur d'une hiérarchie existant dans la société.*
>
> *Une vision plus adéquate des droits individuels se présente comme suit : les droits individuels sont « co-possibles » ; chaque personne peut exercer son droit comme elle l'entend. L'exercice de ces droits fixe certaines caractéristiques du monde. A l'intérieur des contraintes de ces caractéristiques fixes, un choix peut être opéré par un mécanisme de choix social basé sur une préférence sociale ; s'il reste encore des choix à faire ! »*[183]

Les droits individuels fixent la frontière entre ce qui peut faire l'objet d'une préférence collective et ce qui ne le peut pas. La marge de manœuvre de la collectivité peut s'en trouver considérablement réduite et l'on comprend l'exclamation : « s'il reste encore des choix à faire ! ». La théorie libertarienne postule le statut moral d'individus dotés *a priori* de droits d'appropriation et de propriété, droits qui définissent la sphère personnelle de cet individu. Pour que cela soit possible, l'individu des libertariens doit faire preuve d'intégrité, d'autonomie, de séparation et de singularité, ces qualités étant comprises dans le concept d'individu. Le caractère volontaire assumé des actes d'appropriation et de transfert manifesterait l'intégrité et l'autonomie de l'individu ; la distinction entre les droits d'un individu et ceux d'un autre, distinction exprimée par l'expérience de pensée de l'état de nature par laquelle commence *Anarchy, State and Utopia*, correspond à la qualité de séparation : les

[181] Nozick, par exemple, exige seulement le respect d'une « clause lockéenne » qu'il reformule, et selon laquelle l'appropriation des ressources par une personne est légitime si elle ne porte pas préjudice à une autre personne.
[182] Les principaux représentants du courant libertarien de gauche sont M. Otsuka, H. Steiner, P. Vallentyne.
[183] NOZICK, Robert, *Anarchy, State and Utopia*, Oxford: Blackwell, 1974 [Traduit de l'anglais (Etats-Unis) par E. d'Auzac de Lamartine, *Anarchie, Etat et Utopie*, Paris : PUF, 2003, p 206-207].

droits tracent des frontières entre les individus d'une part, et entre le domaine des décisions individuelles et celui des décisions collectives d'autre part.

En conclusion, la formulation même du conflit Pareto-liberté est incompatible avec les principes de départ du libertarisme. Le libertarisme conçoit les droits individuels comme des co-possibilités définies indépendamment des préférences collectives[184]. Le conflit Pareto-liberté définit les droits comme des « préférences décisives », c'est-à-dire des préférences qui, du fait qu'elles ont pour objet l'état de la sphère privée individuelle, peuvent déterminer l'ordre de préférence collective dans les cas où il y a incompatibilité avec cette dernière. Autrement dit, la conception libertarienne établit en amont les limites de la sphère privée. En revanche, la conception des droits avancée par Sen dans ce théorème laisse la possibilité d'une préférence collective qui prendrait pour objet la sphère privée. Si cette préférence est compatible avec les préférences de l'individu concerné, c'est bienvenu. Si ce n'est pas le cas, il y a *conflit.* Mais il ne peut y avoir conflit que parce que, précisément, l'exercice des droits individuels n'a pas de priorité sur le principe de Pareto. Dans la perspective libertarienne, n tel conflit ne peut survenir, parce que la légitimité même d'une préférence collective violant les contraintes fixées à l'avance par les droits individuels est exclue.

[184] Cela dit, il n'est pas nécessaire d'être libertarien pour concevoir les droits de cette manière. C'est même en fait la conception la plus proche de l'usage de ce terme dans le langage courant.

b) La critique libérale de la liberté comme préférence décisive

Le philosophe et théoricien de la justice Brian Barry énonce une critique similaire à celle de Nozick, bien que ne s'adossant pas à la doctrine libertarienne (Barry n'étant pas un libertarien). Pour Brian Barry, l'incompatibilité des conditions de liberté et de Pareto viendrait d'une compréhension erronée de la philosophie politique de la liberté. Il faut selon lui distinguer deux niveaux, celui de l'allocation des droits et celui de l'exercice des droits. Le niveau de l'allocation des droits suppose une doctrine sur qui devrait avoir le droit de décider qui lit quoi (si nous reprenons notre exemple), tandis que le niveau de l'exercice des droits définit qui lit effectivement quoi dans tel ou tel état social.

> *« En résumé, si une fonction de bien-être (welfare) social nous dit ce qui constitue le meilleur état du monde, il ne peut y avoir aucun conflit entre une fonction de bien-être social, quel que soit son contenu, et le principe selon lequel il devrait y avoir une sphère protégée dans laquelle les personnes devraient être légalement libres de faire ce qu'elles choisissent de faire. Car tous deux ont en effet un contenu très différent : la fonction de bien-être social porte sur ce qui est « socialement meilleur », le principe libéral sur ce que les personnes devraient être capables de faire sans coercition légale. »*[185]

Barry comprend ainsi le principe de liberté comme un principe régissant des institutions, et la limite entre ce qui peut faire l'objet d'un classement définissant ce qui est bien pour la société et la sphère personnelle où s'exerce exclusivement la liberté individuelle est une limite légale. L'Etat pourrait ainsi garantir la liberté de N de lire *Mein Kampf*[186], et celle de A de ne pas lire ce livre. La sphère individuelle jouit d'une reconnaissance institutionnelle et n'est pas concernée par le problème du choix social.

[185] BARRY, Brian, "Lady Chatterley's Lover and Doctor Fischer's Bomb Party", in ELSTER, Jon et HYLLAND, Aanund (éds), *The Foundations of Social Choice Theory*, Cambridge University Press, 1986. Traduction libre..

[186] A condition que cette lecture n'entrave pas à long terme la liberté d'autrui (on peut raisonnablement douter du libéralisme de SN). Cf. MILL, John Stuart, *On Liberty,* traduit de l'anglais par Dupont White, *De la liberté,* Paris : Gallimard, 1990.

c) *Répondre aux critiques : liberté en exercice, droits-buts et préférence décisive*

Sen se serait-il trompé en faisant du principe de liberté une simple condition de choix social, en refusant de lui accorder une prééminence ou un statut d'exception ? Pourtant, Nozick et Barry estiment qu'il faut restreindre le domaine du choix collectif, au lieu d'essayer d'intégrer les préférences individuelles au choix collectif (ces préférences constituent le matériau même du choix collectif). La perspective du choix social est toute autre : elle ne s'intéresse pas aux garanties institutionnelles ou aux droits individuels pour eux-mêmes (des « non-sens sur pilotis », disait Bentham), mais aux états sociaux effectifs. C'est une perspective conséquentialiste, qui mesure la liberté à ses effets. Les états sociaux parmi lesquels doivent se prononcer des individus comme N ou A sont des états sociaux dont la description inclut non seulement la satisfaction ou l'insatisfaction de chacun, mais également ce que serait *effectivement* une sphère personnelle protégée. La liberté comme objet de choix collectif est une liberté en exercice ; est dite « libre » une action volontaire effectuée dans un état social décrit aussi précisément que possible. La liberté n'est pas conçue indépendamment de son contexte d'application ; elle est garantie par le choix collectif lui-même et non par une procédure définie indépendamment des attentes réelles des membres de la société.

A la liberté en exercice impliquée par la condition de liberté LM correspond le concept de « droit-but » que développe Sen, étroitement lié à celui de capabilité dynamique (*capability*).

« On appellera « système de droits-buts » un système moral dans lequel le respect et la violation des droits sont compris parmi les buts, intégrés à l'évaluation des situations puis appliqués au choix des actions par des liens de conséquence. »[187]

Le système des droits-buts inclut dans la description des états sociaux disponibles deux types d'information supplémentaires :

- l'existence, ou non, de droits institutionnellement reconnus : la garantie de la liberté d'expression, par exemple ;

- la manière dont ces droits modifient l'état du monde : ainsi, la liberté d'expression implique la publication de journaux dénonçant l'absence d'action du gouvernement lorsqu'éclate une famine, par exemple. Cet aspect de la description peut inclure des informations telles que les moyens financiers dont disposent les journalistes, la manière dont ils accèdent aux régions

[187]SEN, Amartya, "Rights and Agency", in *Philosophy and Public Affairs*, 11, 1, 1982, p 3-39 [Trad.Française S. Marnat, *Ethique et Economie*, Paris: PUF, 1993, "Les droits et la question de l'agent», p 130].

touchées par la famine, leur dépendance matérielle à l'égard du pouvoir politique…La description n'est pas close dès l'instant que le droit est énoncé.

Sen défend ainsi une conception positive du droit, au sens où le droit est une capacité réelle de modifier un état du monde en vue d'une fin. Le droit individuel tel que le conçoit Nozick correspond à une exigence pour autrui de ne pas modifier ce qui appartient à la « sphère personnelle ».

Néanmoins, et c'est une des leçons du conflit Pareto-liberté, s'en tenir au cadre conceptuel originel de la théorie du choix social, c'est admettre que les attentes en matière de liberté individuelle ne coïncideront qu'accidentellement avec celles en matière de bien-être collectif. On ne peut être sûr que le « compatriote » d'A partage ses idées.

Le formalisme de la théorie du choix social mettrait ainsi en évidence une difficulté propre à l'existence collective, dont la solution ne pourrait être déterminée qu'empiriquement : l'éducation citoyenne, l'entretien de la sociabilité, la diminution des inégalités sociales, l'homogénéisation des comportements pourraient figurer dans la liste des solutions à cette difficulté.

3.3.2. La liberté individuelle comme réalisation sociale

Sen a tenté d'utiliser les outils conceptuels de la théorie du choix social pour penser les conditions sociales de réalisation du bien-être et des finalités individuelles. D'une part, il s'est soucié du rôle de la société dans l'amélioration des opportunités ou capabilités individuelles (a). D'autre part, il a montré que le cadre conceptuel de la théorie du choix social permet de penser la liberté sans le contrôle direct de l'individu sur la réalisation de son action (b).

a) Responsabilité individuelle, responsabilité sociale et rareté

L'approche par les capabilités vise à mettre en place des politiques publiques d'amélioration de l'avantage individuel. En d'autres termes, les institutions et même l'organisation sociale en général sont déterminantes pour l'amélioration des capabilités. Comment penser alors la division du travail entre responsabilité sociale et responsabilité individuelle dans l'approche des capabilités ? Quelques pistes nous sont données ici, non par Sen, mais par l'économiste Vivian C. Walsh qui, avant Sen, a eu l'idée d'évaluer la privation économique en termes de privation de liberté.

Dans son ouvrage sur la rareté[188], Walsh consacre plusieurs pages à décrire comment la réflexion qui précède la décision individuelle fait apparaître la rareté. Son exemple est celui d'un étudiant de philosophie qui doit choisir entre acheter les œuvres de Descartes ou chauffer sa chambre. En outre, Walsh ajoute que ce type de réflexion parle à chacun d'entre nous, et nous rend sensibles à ce que peuvent éprouver les personnages d'un roman lorsqu'ils ont affaire à la rareté au moment de décider :

> *« Toute la grande littérature est remplie de vivantes descriptions de telles évaluations et décisions. »*[189]

Voyons Maupassant :

> *« Lorsqu'il fut sur le trottoir, il demeura un instant immobile, se demandant ce qu'il allait faire. On était au 28 juin, et il lui restait juste en poche trois francs quarante pour finir le mois. Cela représentait deux dîners sans déjeuners, ou deux déjeuners sans dîners, au choix. Il réfléchit que les repas du matin étant de vingt-deux sous, au lieu de trente que coûtaient ceux du soir, il lui resterait, en se contentant des déjeuners, un franc vingt centimes de boni, ce qui représentait encore deux collations au pain et au saucisson, plus deux bocks sur le boulevard. C'était là sa grande dépense et son grand plaisir des nuits […]. »*[190]

L'évaluation individuelle est une comparaison qui détermine tout choix rationnel, du plus ordinaire (s'offrir un bock le soir) au plus important, en termes de conséquences. Mais l'évaluation individuelle s'accompagne d'un sentiment de privation : privation d'argent, comme dans le cas de Bel-Ami (choisir le repas du soir, c'est renoncer à celui de midi), privation de temps, d'énergie…; cette rareté est ce qu'il y a d'« économique » dans l'expérience humaine du choix.

> *« Ce qui est économique dans toute expérience réside dans ses effets sur la fin ou les fins qu'une personne est plus ou moins consciemment en train de poursuivre. Il n'y a rien d'économique dans la fin en tant que telle ; rien n'est économique en soi. »*[191]

L'économiste devrait ainsi s'intéresser à l'aptitude de chacun à réaliser ses buts, étant donnée la rareté relative des ressources au sens large du terme (argent, biens, mais aussi capabilités physiques, temps, relations sociales…)..

La rareté contraint l'agent à faire un choix qu'il ne juge pas toujours bon, et ce sans en être responsable. L'enjeu de cette réflexion sur la rareté est la mise en évidence des limites de la responsabilité individuelle et de la nécessité d'une responsabilité sociale pour la réalisation

[188] WALSH, Vivian C., *Scarcity and Evil,* Prentice-Hall, 1961.
[189] WALSH, Vivian C., *Scarcity and Evil,* Prentice-Hall, 1961, p 22. Traduction libre.
[190] MAUPASSANT, Guy de, *Bel-Ami.*
[191] WALSH, Vivian C., *Scarcity and Evil,* Prentice-Hall, 1961, p 23-24. Traduction libre..

des buts raisonnables. Walsh souligne ainsi que nous attribuons parfois de la responsabilité là où il n'y a que de la rareté[192]. Or, l'allocation de ressources de nature diverses (éventuellement en se servant de compensations) est l'objet privilégié du choix collectif tel que le thématisent les économistes du bien-être et les théoriciens de la justice. La rareté est un souci collectif.

Où s'arrête la responsabilité sociale ? Cette question concerne aussi le conflit Pareto-liberté : la société est-elle responsable du fait que le sympathisant nazi lise *Mein Kampf* ? On peut distinguer au moins trois conceptions des limites de la responsabilité sociale :

- La responsabilité sociale peut être définie par son objet : ainsi, on distinguera les privations naturelles des privations sociales. La société n'est responsable que des privations sociales, dans la mesure où elle les aurait causées. Cette conception de la responsabilité présuppose une dichotomie entre avantages naturels et avantages sociaux. La volonté de Sen de prendre en compte à la fois des dimensions naturelles et des dimensions sociales dans l'évaluation des capabilités nous incite à penser qu'il ne serait pas favorable à cette division du travail entre responsabilité individuelle et responsabilité sociale.
- La responsabilité sociale peut être définie par sa structure : la société est responsable de ce qui peut être formulé en termes de droits, indépendamment de considérations particulières. Ainsi, les privations dont la société serait responsable sont des privations générales. Cette conception caractérise les approches que Sen qualifie dans *L'Idée de Justice* de transcendantales ou d'institutionnalistes .Ayant adopté une approche comparative, Sen ne peut souscrire à cette conception de la responsabilité sociale.
- La responsabilité sociale peut aussi englober ce que la société *peut faire.* Cette conception caractérise les approches conséquentialistes de la justice sociale. En effet, alors que les priorités définies par le contrat social limitent le champ d'intervention des institutions, tout est susceptible d'être un facteur de bien-être individuel.

[192] *Ibidem*, p 50.

b) Contrôle direct et liberté indirecte

Lorsque nous parlons de « liberté de choix », nous identifions liberté et pouvoir de décision. Selon Sen, la théorie du choix social propose une conception moins restrictive de la liberté : sa structure permet de parler de « liberté indirecte ». De quoi s'agit-il ? Lorsque l'individu a le pouvoir de contrôler totalement sa sphère personnelle, cette notion n'a pas lieu d'être. Mais Sen veut montrer que ce contrôle est en fait assez rare, voire tout à fait inadéquat pour invoquer la liberté de l'agent dans certaines situations.

Il réfute donc la restriction de la liberté individuelle au pouvoir (exercé) de décision par l'exposition d'un cas[193]. Par conviction, Edmond refuse de faire souffrir les animaux. Blessé dans un accident, il se trouve confronté à l'alternative suivante : soit il choisit un traitement A très efficace, mais mis au point à l'aide d'expériences qui ont fait souffrir des animaux, soit il se contente d'un traitement B qui occasionne beaucoup d'effets secondaires désagréables, mais qui n'a pas requis d'expériences sur les animaux. Edmond choisit le traitement B : on parle ici de « liberté directe », ou liberté de choix.

A présent, Sen nous demande d'imaginer qu'Edmond est inconscient suite à son accident : il appartient donc au médecin et à l'épouse d'Edmond (qui connaît l'engagement de ce dernier en faveur des animaux) de décider du traitement. Sen soutient que ces derniers serviraient mieux le bien-être d'Edmond en choisissant le traitement A, mais qu'ils le priveraient alors de sa « liberté indirecte ». Autrement dit, même s'il est incapable de contrôler les décisions concernant sa sphère personnelle, Edmond reste libre dans la mesure où il existe des personnes capables et désireuses de réaliser ce qu'il aurait lui-même voulu.

> *« Pour caractériser la liberté, la théorie du choix social compare la situation obtenue avec ce que la personne aurait choisi, même si elle n'effectue pas elle-même le choix. Mais cette méthode laisse de côté un aspect qui peut être important pour la liberté : le fait de savoir si ce que la personne obtient a été effectivement choisi par elle, et n'est pas simplement ce qu'elle aurait choisi (pas nécessairement en effectuant elle-même le choix). C'est là une lacune, et bien qu'elle ne remette nullement en cause le conflit Pareto-liberté […], elle peut être importante dans un traitement plus général de la liberté (par opposition au traitement de certaines de ses conséquences). La seule manière de combler cette lacune est d'enrichir la description des états sociaux pour y incorporer l'agent du choix. Cela implique que l'on s'éloigne du format existant de la théorie du choix social, dans lequel les personnes*

[193]SEN, Amartya, « Liberty and Social Choice », *The Journal of Philosophy,* LXXX, 1, janvier 1983, p 5-28.Traduit de l'anglais par MARNAT, Sophie, *Ethique et Economie,* "Liberté et choix social", Paris: PUF, 1993, p 175-176.

choisissent entre différents états sociaux sans que la description du choix soit incorporée dans ces états eux-mêmes […]. »[194]

La description de la liberté peut inclure trois facteurs : la préférence de l'agent, l'activité de choisir et le résultat, conforme à la préférence[195]. La théorie du choix social ne retient dans sa description des états sociaux que le premier et le troisième facteur : elle ne se demande pas *comment* ni *par qui* ont été réalisés les buts de l'agent. Ainsi, l'action sociale peut contribuer à la réalisation des fins individuelles en lieu et place de l'action individuelle: l'expérience de la rareté traduit ainsi la nécessité de compléter l'activité individuelle par des ressources sociales.

Néanmoins, l'éthique qui sous-tend l'approche par les capabilités intègre également des considérations propres à la tradition libérale : ainsi, le fait que la personne réalise elle-même ses buts peut être une information qui la détermine à préférer l'état social où c'est possible. Nous pourrions donc développer l'exemple de Sen : Edmond est certes inconscient, mais il est possible de le réanimer et de lui administrer ensuite le traitement choisi, sans diminuer ses chances de guérison. Dans ce cas, il est probable que son épouse et son médecin choisiraient de le réanimer avant de lui demander quel traitement il préfère, et ce *même si tous deux sont certains qu'il choisira le traitement B*. Et peut-être Edmond se sentirait-il lésé s'ils avaient procédé autrement, *même si le résultat ne change pas.*

En incluant l'agent du choix dans la description des états sociaux, Sen signifie qu'il appartient à la société de garantir les libertés individuelles, à la fois au sens conséquentialiste du terme (la liberté comme réalisation des fins désirées, ou liberté-opportunité) et au sens de la tradition libérale (la liberté comme processus de choix dans lequel la personne concernée est également l'agent du choix). Mais, contrairement aux théories du contrat social, cette garantie ne prend pas la forme d'une structure institutionnelle spécifique, mais se vérifie par la modification de la situation sociale existante en faveur d'une augmentation des activités individuelles libres.

Philip Pettit rapproche la conception de la liberté de l'approche par les capabilités et la tradition républicaniste, qu'il présente ainsi :

« La conception républicaine de la liberté est socialement plus radicale que l'approche dominante dans la mesure où elle nie que la non-interférence soit suffisante pour la liberté (Pettit 1997, Ch. 2). Cet aspect signifie que nous devons considérer ceux qui dépendent de la bonne volonté des autres pour jouir d'une vie normale sans entrave comme non libres, même s'ils ont assez de chance pour s'attirer

[194] SEN, Amartya, *op. cit.*, p 177.

[195] Cf. PETTIT, Philip, "Capability and Freedom: a Defence of Sen", *Economics and Philosophy*, 17, 2001, pour une formulation de cet ordre.

cette bonne volonté. Selon ma lecture, Sen soutient ce radicalisme social dans la mesure où il accorde de l'importance aux capabilités de fonctionner et pas seulement des perspectives de fonctionnement, lorsqu'il s'agit de déterminer la qualité de vie d'une personne. Il ne suffit pas qu'il soit probable que la femme sera bien traitée par son puissant mari, l'employé par son puissant employeur, le pauvre du village par le puissant propriétaire terrien. Il faudra également que ceux qui se trouvent dans de telles positions bénéficient de ce traitement pour des raisons qui soient indépendantes de la bonne volonté des puissants.» [196]

Le républicanisme, au sens où Pettit entend ce terme, promeut la liberté comme non-domination. Jouir de ce type de liberté, c'est vivre au sein d'institutions qui garantissent que personne n'a la capacité d'interférer avec nos choix sur une base arbitraire, et ce même si les préférences des dominants concordent accidentellement avec le contenu des préférences des dominés. La liberté comme non-domination, comme la capabilité, concerne les possibilités des personnes qui jouissent de l'une ou de l'autre, et pas seulement ce qu'elles réalisent effectivement. De même que Pettit préfère décrire la privation de liberté en termes de domination (possible) plutôt que d'interférence (actuelle), de même Sen préfère décrire l'avantage en terme de capabilités (un avantage réellement possible) qu'en termes de fonctionnements (un avantage réalisé).

Mais le parallèle s'arrête ici : en effet, Sen justifie la perspective des capabilités par l'importance qu'il accorde au choix de l'agent de réaliser telle ou telle capabilités. Tandis que Pettit valorise davantage la non-domination parce qu'elle offre une sécurité plus grande contre la privation de liberté que l'absence d'interférence. Sen refuse donc de souscrire à la lecture que fait Pettit de l'approche par les capabilités. On peut donner de quatre raisons à ce refus de prise de position en faveur du républicanisme :

- La première raison est d'ordre méthodologique. Sen souligne que la distinction entre interférence potentielle et interférence actuelle est significative[197]. L'approche républicaniste se concentre sur les interférences potentielles. Mais la théorie du choix social (dont Sen fait la promotion) évalue des états sociaux effectifs, dont la description inclut d'abord interférences actuelles. On peut certes enrichir cette description en notant tous les faits susceptibles de signaler des interférences potentielles. Mais cela ne va pas sans difficulté : ainsi, il faut recenser des faits qui nous permettent d'affirmer qu'un mari pourrait entraver les actions de son épouse ; nous ne pouvons pas affirmer *a priori*, indépendamment d'une bonne connaissance de

[196]Ibidem.

[197] SEN, Amartya, *The Idea of Justice*, London: Allen Lane (Penguin Books), 2009, p 308.

l'environnement social, du droit en vigueur ou encore du passé de ce ménage, que le mari « domine » son épouse. En choisissant une méthode d'évaluation sociale empirique et comparative, Sen refuse de cristalliser son approche autour de la seule question de la domination.

- La seconde raison tient au fait que Sen devrait refuser de faire sienne la dichotomie du naturel et du social. Consacrer la non-domination comme priorité sociale et politique, c'est faire abstraction de l'importance des entraves «naturelles » pour l'activité humaine. L'exemple que donne Sen est tout à fait significatif de ce point de vue : afin de mettre en évidence des situations de domination, il expose le cas d'une personne handicapée qui ne peut se déplacer elle-même[198]. Soit elle ne reçoit aucune aide (cas 1), soit elle est *toujours* aidée par des volontaires ou par des personnes envoyées par les services sociaux du lieu où elle vit (cas 2), soit elle est aidée par des employés qu'elle paie elle-même (cas 3). Le cas 1 est indiscutablement une privation de capabilité. Le cas 2 rend cette handicapée dépendante de la bonne volonté d'autres personnes (qui choisissent de l'aider directement ou de continuer à financer par leurs impôts des services sociaux) : on peut parler de domination, puisque ces autres personnes *pourraient* interférer avec la liberté de mouvement de cette handicapée. Mais elles ne le font *jamais* : il n'y a donc pas de différence, en termes de capabilités, entre le cas 2 et le cas 3, où cette personne handicapée ne subit ni domination potentielle, ni domination actuelle. Et il n'y a pas de différence, en termes de non-domination, entre le cas 1 et le cas 2. Néanmoins, le cas 2 semble infiniment préférable au cas 1 : autrement dit, la privation de capabilité requiert ici une action prioritaire à celle qui consiste à éliminer la domination potentielle. C'est la comparaison entre les états sociaux, et non pas l'existence ou non de domination, qui permet de déterminer l'action sociale.
- La troisième raison est que Sen ne fait pas de philosophie politique au sens strict du terme. Il ne souhaite pas mener une réflexion sur la nature des institutions ; celles-ci ne l'intéressent que dans la mesure où elles produisent des effets sur l'existence des personnes. Or, affirmer la priorité de la non-domination, cela revient à définir au moins partiellement ce que devraient être les institutions politiques.
- La quatrième raison est que Sen souhaite laisser aux membres des sociétés le soin de hiérarchiser eux-mêmes les valeurs qui détermineront le choix collectif. On ne peut

[198] *Ibidem,* p 306.

> être à la fois un théoricien du choix social et défendre « sa » vision du monde valable pour toute société donnée. Une fois le travail de clarification des valeurs effectué, l'expert, économiste ou philosophe, se retire et laisse se prononcer la collectivité[199].

S'il appartient à l'action collective de garantir la liberté individuelle, il lui appartient aussi de choisir le type de liberté qu'elle mettra en œuvre.

[199] Ce positionnement de Sen est particulièrement clair concernant la question de la liste des capabilités, sur laquelle nous revenons dans le Chapitre 6 de ce travail.

3.3.3. Un accord basé sur les évaluations positionnelles ? A propos des jugements transpositionnels

Nous avons vu dans la première partie qu'il est possible de mettre en relation les évaluations que comptabilise la théorie du choix social avec la description des positions qu'occupent les individus. Cette hypothèse de lecture permet de mieux comprendre la place de l'individu dans la contribution de Sen. Mais, puisqu'il est question de choix collectif, il reste à déterminer comment l'on peut passer de ces évaluations positionnelles à un jugement de valeur valable pour la collectivité. Dans cette dernière section, nous proposons d'étudier le concept de jugement transpositionnel à cette fin.

Il est possible de parler d'évaluation transpositionnelle lorsqu'il apparaît que des jugements de valeur énoncés à partir de positions différentes se recoupent[200]. Souvenons-nous : l'illusion objective est causée par « l'absence d'accès à un autre examen positionnel »[201]. De même, des jugements de valeur qui heurtent le sens commun pourraient ne pas résister à un examen critique effectué à partir d'une autre position.

Un jugement de valeur transpositionnel est donc le fruit de comparaisons successives entre des jugements de valeur positionnels. Ce processus d'élaboration distingue les valeurs transpositionnelles de valeurs universelles posées *a priori.* Ainsi, des propositions telles que les Droits de l'Homme peuvent être acceptées si elles passent avec succès l'examen transpositionnel, et non parce qu'elles seraient déduites de la connaissance de la nature de la raison humaine, par exemple.

Nous avons vu que Sen estime que les débats méthodologiques fondés sur la dichotomie individu-société, à l'instar du débat entre l'individualisme méthodologique et le holisme classique, sont de faux débats. Sen privilégie la logique de la relation qui fait défaut dans ce débat : ce qui devrait intéresser la philosophie morale et politique, ce sont les relations qu'entretiennent entre elles ces entités que sont les individus et les sociétés : la relation d'appartenance d'un individu à un groupe, à une société ou à l' « humanité » ne va pas de soi, mais elle est première dans la compréhension des valeurs de cet individu.

[200] Ceci évoque le "consensus par recoupement" rawlsien ; mais ce dernier est un consensus entre personnes raisonnables et ayant chacune une conception du bien, tandis que l'évaluation transpositionnelle est un accord entre des individus en tant qu'ils occupent des positions sociales et économiques différentes : cet accord ne fait pas nécessairement appel à la volonté de vivre ensemble, mais plutôt à l'existence parfois avérée de coïncidences de valeurs en dépit de l'hétérogénéité sociale.

[201]SEN, Amartya, *Rationality and Freedom*, Belknap Press of Harvard University Press, 2002 [Trad. Française M.-P. d'Iribane-Jaawane, Odile Jacob, 2005, p 370]

C’est sur cette base que Sen critique l’approche « nationale » de l’évaluation sociale (par exemple, par la mesure du PIB). Les statistiques élaborées à l’échelle d’une Nation occultent l’hétérogénéité économique et sociale : parler de « pays riche » ou de « pays pauvre » est une absurdité.

Conclusion de la Ière Partie.

Cette investigation historique de la théorie du choix social a eu pour but de déterminer s'il est nécessaire de postuler une individualité substantielle qui serait à la source des préférences, intérêts et jugements qui constituent la base d'information ou la matière du choix social tel que le comprennent Arrow et Sen. Le Chapitre 1 a montré qu'en substituant les préférences aux utilités, la Nouvelle Economie du Bien-Etre a, en un sens, dépsychologisé l'individu. Dans le Chapitre 2, l'analyse des justifications qu'apporte Arrow à sa propre contribution formelle met en évidence une tension entre les possibilités du formalisme et son désir de sauver l'intégrité de l'individu. Enfin, le Chapitre 3 a fait le pari de la cohérence d'ensemble des travaux de Sen[202], afin d'éclairer le statut de l'individu dans sa propre approche du choix social par ses contributions sur les capabilités et l'objectivité positionnelle. L'individualité y est apparue comme construite, dépendante des relations sociales, mais susceptible de se voir attribuer en propre des préférences. Conceptuellement, la théorie du choix social n'a pas besoin de considérer l'individu comme une donnée naturelle et nécessaire, source de toute évaluation, de tout désir et de tout choix *réel*, mais peut, à l'instar de Sen, y voir l'aboutissement d'un processus social d'individuation.

L'une des caractéristiques distinctives de la théorie du choix social, caractéristique qu'elle partage avec l'utilitarisme et qui la distingue des théories du contrat social ou des théories de la démocratie délibérative, est l'*agrégation.* Il n'y a pas de théorie du choix social sans agrégation d'évaluations distinctes, que ces évaluations aient la forme d'utilités, de préférences ou de capabilités. L'individu, entité distincte et séparable d'autres individus, semble dès lors une excellente base pour une telle approche. *Mais cela ne signifie pas qu'il soit la seule.* Ce qui rend nécessaire l'agrégation, c'est l'hétérogénéité des composantes, la pluralité des évaluations, mais ces composantes et ces évaluations ne sont pas nécessairement individuelles : il peut s'agir de corporations, de lobbies, de classes sociales, de tribus, de familles, d'Etats-Nations, de régions…[203]

[202] Lui-même la présente comme cohérente dans son ouvrage *L'Idée de Justice*, et Arrow a également signalé cette cohérence lors du discours de remise du Prix Nobel d'économie à Sen.

[203] Nous reviendrons sur le cas des "individus collectifs" ; cf. Ph. Pettit, *Penser en société*, PUF, 2004, trad. A. Bouvier, B. Guillarme, A. Ogien, P. Livet.

IIe partie : Nature et valeur des préférences

Les préférences constituent la base d'information du choix collectif dans la plupart des travaux de la théorie du choix social, généralement décrite comme l'étude de l'agrégation des préférences[204]. Ces préférences sont vues comme subjectives, et sont assimilées aux choix dans les approches qui s'appuient sur le concept de « préférence révélée » que nous présenterons dans le prochain chapitre. En ce sens, elles semblent à des lieues de l'évaluation de l'avantage en termes de capabilités. Afin de déterminer dans quelle mesure le concept de préférence peut être associé à l'approche par les capabilités (ce qui fera l'objet de la IIIe partie), il nous faut opérer un travail de clarification du concept de préférence utilisé en économie normative en général et en théorie du choix social en particulier.

Or, l'interprétation du concept de « préférence » y reste ambiguë, comme le rappellent Marc Fleurbaey et Philippe Mongin :

> *« Enfin, parce qu'ils emploient les méthodes de l'économiste théoricien, et héritent de concepts techniques qu'ils n'analysent pas toujours complètement, les auteurs contemporains [de la théorie du choix social] en répercutent certaines ambiguïtés caractéristiques. En particulier, même si la critique du « welfarisme » a permis d'en mieux délimiter le rôle, la notion de préférence apparaît à la fois omniprésente et mystérieuse: elle intervient tantôt dans la prise de décision et le choix, tantôt dans ces actes purement idéels que sont l'évaluation et le classement; elle décrit tantôt le bien de l'individu de façon quasi objective, tantôt son bien-être ou son bonheur, tantôt, seulement, ses goûts. »*[205]

Les critiques adressées aux approches de l'évaluation sociale ou de la philosophie morale basées sur les préférences visent généralement le caractère « capricieux » de ces dernières. Les préférences seraient déterminées par les goûts de l'individu. Ces goûts ne se discutent pas, ne se critiquent pas. Les préférences seraient donc potentiellement arbitraires, tant sur le plan cognitif que sur le plan normatif. Elles seraient arbitraires sur le plan cognitif dans la mesure où elles pourraient être déterminées par des croyances erronées ou des perceptions biaisées de ce qui est faisable et de ce qui ne l'est pas, comme dans le cas des préférences adaptatives. Et elles seraient aussi arbitraires sur le plan normatif dans la mesure où rien

[204] Pour un exemple de formalisation des préférences en théorie du choix social, on se reportera à l'Annexe 2.

[205] FLEURBAEY, Marc et MONGIN, Philippe, « Théorie du choix social et économie normative », *in* CANTO-SPERBER Monique (éd.), *Dictionnaire de philosophie morale,* Paris, P.U.F., 1996.

n'interdit les préférences antisociales ou perverses de compter dans le calcul du bien-être collectif comme dans le choix social.

Par conséquent, l'intérêt d'étudier les préférences à partir du cadre conceptuel formel de la théorie du choix social est de montrer que ce cadre est compatible avec d'autres interprétations. On pourrait alors déterminer les plus appropriées au contexte d'application de la théorie du choix social. Autrement dit, il faudrait, comme l'affirme Elizabeth Anderson, concevoir la préférence comme un outil au lieu d'en faire un usage dogmatique.

> *« La critique d'un concept n'est pas un rejet de celui-ci, mais une exploration de ses diverses significations et limitations. Une manière d'exposer les limitations d'un concept consiste à introduire de nouveaux concepts qui ont des significations différentes mais peuvent de façon plausible concurrencer les concepts critiqués pour certains des usages auxquels ces derniers servent. L'introduction de tels nouveaux concepts nous donne des choix de façons de penser que nous n'envisagions pas clairement auparavant. Avant d'envisager en effet ces alternatives, notre usage du concept en question est* dogmatique. *Nous le déployons automatiquement, sans poser de question, parce qu'il semble être comme le cadre conceptuel inévitable au sein duquel l'investigation doit être menée. En envisageant des alternatives, nous convertissons les dogmes en* outils, *c'est-à-dire des idées que nous pouvons* choisir *d'utiliser ou non, selon que l'usage de ces idées convient bien ou non à notre objet de recherche. »*[206]

L'usage dogmatique du concept de « préférence » utilisé en théorie du choix social et en économie du bien-être consiste à l'appliquer au problème traité sans considérer des solutions alternatives. La proposition d'Anderson peut être spécifiée et adaptée aux différents usages du concept de préférence, à savoir l'usage explicatif et l'usage normatif. Notre travail est concerné en premier lieu par l'usage normatif. La question est alors la suivante : le concept formel de préférence présuppose-t-il nécessairement une interprétation aussi réductrice et une compréhension aussi pauvre de l'évaluation et des choix ?

L'usage dogmatique de la préférence ne se limite pas à l'utiliser pour décrire des phénomènes qui gagneraient à être décrits à l'aide d'autres concepts ; il consiste également à assigner à la notion de préférence une interprétation unique qui reflète une conception appauvrie de la complexité des motivations humaines et du bien-être. Nous pouvons « préférer » une chose à une autre pour des motifs bien plus variés que le plaisir ou la recherche de la satisfaction égoïste. L'usage normatif de la préférence doit être évalué à l'aune de la capacité de cette dernière à décrire ce qui est bon pour l'agent à qui cette préférence est attribuée. La difficulté

[206]ANDERSON, Elizabeth, "Unstrapping the straitjacket of 'preference': a comment on Amartya Sen's contributions to philosophy and economics."*Economics and Philosophy*, 17, 2001, pp 21-38.

est que la préférence de la théorie du choix social et de l'économie du bien-être est d'abord un concept formel défini par des propriétés mathématiques. En revanche, ce qui est « bon pour l'agent » relève d'une compréhension substantielle, constituée soit sur des faits empiriques, soit sur une anthropologie plus compréhensive. Il y a ici des choix philosophiques à faire en matière d'interprétation de cette structure formelle, la structure étant elle-même partiellement contraignante.

Lors de cette analyse, il apparaîtra que le cadre formel autorise des interprétations très diverses de la préférence. Le risque est de se perdre. Nous serons donc amenés à faire des choix interprétatifs. Ces choix interprétatifs seront déterminés par les deux objectifs poursuivis dans cette IIe partie :

1) Le premier objectif étant la clarification du cadre conceptuel de la théorie du choix social et de l'économie normative, nous nous attacherons à examiner une classe d'interprétations du concept de préférence qui soit aussi étendue que possible. Néanmoins, nous tenterons de déterminer, parmi ces interprétations, celle qui est à la fois véritablement en adéquation avec le cadre formel lui-même, d'une part, et celle qui est la plus générale, d'autre part. Nous souhaitons éviter ainsi d'introduire des biais interprétatifs.
2) Le second objectif est la conciliation du préférentialisme de la théorie du choix social avec les présupposés conceptuels de l'approche par les capabilités. Il motivera en particulier les choix interprétatifs mis en évidence dans le Chapitre 5, mais nous y reviendrons surtout dans la IIIe partie.

Dans cette IIe partie, nous utiliserons ces critères de choix interprétatifs pour défendre la conception suivante de la préférence : d'une part, la préférence devrait être comprise comme un acte d'évaluation. Cette évaluation est comparative (Chapitre 4). D'autre part, la conception de la préférence la plus à même de surmonter les problèmes relatifs à la position de l'évaluateur et à l'objectivité de la norme du bien-être est la suivante : les préférences devraient être évaluées en adoptant les perspectives positionnelles mises en évidences dans le Chapitre 3 ; par ailleurs, la conception de la préférence la plus à même de surmonter les difficultés que rencontrent les conceptions objectives comme les approches subjectives du bien-être est une conception consistant à filtrer les préférences en lui imposant des conditions de rationalité (Chapitre 5).

Chapitre 4 : Préférences, choix et évaluations

Si la théorie du choix social définit le choix social comme l'agrégation des préférences, la notion de préférence elle-même reste ambiguë. Il convient donc de déterminer, parmi les définitions possibles de la préférence, celle qui est à la fois cohérente avec le cadre conceptuel formalisé de la théorie du choix social et avec l'usage ordinaire du terme « préférence ».

Comme l'écrivent Marc Fleurbaey et Philippe Mongin, la définition de la préférence en théorie du choix social et en économie normative oscille entre les actes effectifs de décision et de choix et les attitudes d'évaluation. La préférence révélée se confond même avec l'action. Nous proposons donc de déterminer le concept de préférence au sein d'un réseau de notions incluant les actions, les décisions, les choix, les désirs, les jugements de valeur, les croyances… L'hypothèse que nous souhaitons vérifier est la suivante : les préférences se situeraient plutôt du côté des évaluations que des actes effectifs de choix. Le problème de la définition de la nature de la préférence (choix ou évaluation) est crucial pour la suite de notre travail, car de sa solution dépend la possibilité d'associer un cadre préférentialiste à l'approche des capabilités. En effet, dans son ouvrage *Commodities and Capabilities,* Sen critique l'approche des préférences révélées en montrant qu'elles ne peuvent pas rendre compte de manière satisfaisante du désavantage.

L'analyse que nous mènerons afin de vérifier notre hypothèse nous amènera à distinguer la préférence du choix et de l'action, et même à prendre nos distances à l'égard de la thèse de l'efficacité causale de la préférence (4.1.) ; ayant ainsi établi que la préférence relève du jugement de valeur, nous en définirons plus précisément les caractéristiques, en distinguant notamment comparaison factuelle et comparaison évaluative, d'une part, et jugement de valeur classificatoire et jugement de valeur comparatif, d'autre part (4.2.).

4.1. Préférence et décision : de la préférence révélée à la préférence fonctionnelle

Les théories de la décision traitent simultanément la préférence et le choix. Choisir, c'est sélectionner une option parmi d'autres. Cette option peut-elle être alors considérée comme l'option *préférée* ? Et, si oui, une préférence détermine-t-elle systématiquement un choix ? Si non, comment peut-on définir la relation entre préférence et choix ? Dans cette section, nous discuterons successivement différentes mises en relation de la préférence et du choix. Cette discussion aboutira à la défense du concept de préférence fonctionnelle, qui connecte la

préférence au choix, mais également à d'autres attitudes, telles que les autres préférences ou les croyances. La préférence est ainsi abordée du point de vue de sa fonction.

Le postulat des préférences révélées en économie a confondu préférence et choix (4.1.1.). Cette confusion a fait l'objet de vives critiques, notamment de la part d'Amartya Sen (4.1.2.). La relation entre préférence et choix a été précisée par Daniel Hausman, et gagne en clarté à être confrontée aux travaux de Donald Davidson sur le rôle explicatif de la causalité (4.1.3.). Les critiques de la réduction de la relation entre préférence et choix à une relation de causalité *seule* permettent de comprendre par la suite l'intérêt du concept de préférence fonctionnelle pour notre analyse (4.1.4.).

4.1.1. Les préférences révélées

Les préférences révélées identifient préférence et choix. Il n'y aurait rien à décrire au-delà de l'observation du comportement de choix. Le choix correspond ici au comportement observable de détermination d'une alternative. La préférence est l'appréciation de cette alternative.

En 1938, l'économiste américain Paul Anthony Samuelson (1915-2009) publie dans la revue *Economica* un article intitulé « A Note on the Pure Theory of Consumer's Behaviour » [*Une note sur la théorie pure du comportement du consommateur*], dans lequel il propose de débarrasser la théorie du consommateur des derniers vestiges du concept d'utilité issu de l'utilitarisme philosophique. D'après l'article de 1938 de Samuelson, la théorie du consommateur pourrait être fondée sur un postulat unique :

> *« L'individu se comporte toujours de manière cohérente [*consistently*] au sens où il ne devrait jamais « préférer » un premier groupe de biens à un second en même temps qu'il préfère le second au premier »*[207]

Ce postulat signifie que le comportement de l'individu correspond à sa préférence : il ne peut d'après lui y avoir de comportement de choix contraire à la préférence. La préférence individuelle est ainsi révélée par le choix effectif : le choix d'un bien ou d'un groupe de bien révèle une préférence pour ce bien ou ce groupe de biens. Les préférences sont construites à partir des choix. Une méthode possible est la suivante[208] :

---

[207] SAMUELSON, Paul A., « A Note on the Pure Theory of Consumer's Behaviour: An Addendum", *Economica*, New Series, Vol. 5, No 19 (Août 1938), pp 353-354.

[208] Pour une brève mais utile introduction aux différentes méthodes de représentation des préférences par le choix, cf. *Stanford Encyclopedia of Philosophy*, article « Preferences », § 3.3., publié le 4 oct. 2006, révisé le 6 sept. 2011, consulté en ligne à l'adresse suivante : http://plato.stanford.edu/entries/preferences/index.html#PreCho

- Définissons *R*, *P* et *I* comme des relations d'ordre que l'agent établit sur les options : *R* signifie « est préféré ou indifférent à » ; *P* signifie « est préféré à » ; *I* signifie « est indifférent à ».*R* est définie sur un ensemble d'options *A*. *C* est une fonction de choix qui détermine un sous-ensemble *B* de *A* dont les éléments n'ont pas été supprimés par le fait de choisir.

  Pour construire la préférence à partir du choix, on dira qu'une option *x* est « préférée ou indifférente » à *y* si et seulement si *x* est choisie dans un ensemble d'options *A* qui contient également *y*.

  x R y si et seulement si, pour *A*, $x \in C(A)$ et $y \in A$.

  x P y si et seulement si *x R y* et $\neg y\ R\ x$.

  x I y si et seulement si *x R y* et *y R x*.

Par conséquent, il n'existe pas pour la théorie du consommateur ainsi définie de préférence « cachée », sous-jacente, qui ne se laisserait pas découvrir par les choix. Comme le fait remarquer Amartya Sen[209], le point de vue de l'économiste suit alors une direction inverse de celui de l'agent qui choisit : nous, « agents », décrivons notre expérience de choix comme celle d'une préférence qui a par la suite déterminé un choix (au moins lorsque les conditions externes y sont favorables), tandis que l'économiste, placé dans la position d'un observateur extérieur, décrit le choix et en déduit l'existence d'une préférence.

Cet article s'inscrit dans le mouvement positiviste soucieux de faire de l'économie une « science ». La description des utilités ne pouvait s'appuyer sur aucun point de référence « objectif », c'est-à-dire, d'après les économistes revendiquant cette conception de la science économique, pouvant être décrit de la même manière par n'importe quel observateur extérieur. En effet, la description des utilités est tributaire de la perception de l'agent ainsi que de l'interprétation de l'observateur : la préférence révélée permettrait au contraire d'éliminer toute référence à un état psychologique dont on ne peut rendre compte que de manière arbitraire ou approximative.

Le postulat des « préférences révélées » a dès lors peu à peu conduit à l'assimilation pure et simple de la préférence et du choix, y compris en économie du bien-être. La confusion touche aussi la théorie du choix social. Ainsi, dans sa monographie sur le théorème d'impossibilité des choix collectifs, Arrow propose de représenter les préférences « et » le choix par des

[209]SEN, Amartya, « Behaviour and the Concept of Preference", *Economica*, 40 (Août 1973), pp 241-259, d'après une conférence inaugurale donnée à la London School of Economics.

relations d'ordre[210] ; or, lorsqu'il parle de cette « représentation des préférences et du choix », il faut comprendre que préférence et choix sont ici la même chose : d'ailleurs, il n'y a pas de représentation du choix distincte de celle de la préférence.

[210] ARROW, Kenneth, *Social Choice and Individual Values,* New York, Wiley, 1951.

4.1.2. Amartya Sen, critique des préférences révélées

Dans le cours inaugural donné à la London School of Economics que nous avons cité, Amartya Sen se montre très critique à l'égard de ceux qui jugent que le concept de préférence révélée permet de faire l'économie de toute référence psychologique. Au contraire, affirme Sen, le postulat des préférences révélées repose sur une hypothèse psychologique, l'hypothèse selon laquelle une personne a tendance à rechercher et à choisir ce qui correspond à ses préférences. Même si cette hypothèse semble raisonnable de prime abord, il ne s'agit ni d'une tautologie, ni de la description d'un fait observable ; elle ne satisfait donc pas davantage les critères de validité d'une approche positiviste que n'importe quelle autre hypothèse psychologique. Les préférences révélées sont donc méthodologiquement défectueuses.

Les tenants des préférences révélées pourraient cependant défendre le caractère « raisonnable » de cette hypothèse ; celle-ci serait « raisonnable » parce qu'elle correspondrait à la plupart de nos observations courantes : ce serait une proposition de bon sens. Mais Sen discute également cet argument en mettant en évidence le fait que nos engagements, les normes sociales ou morales auxquelles nous obéissons, peuvent nous conduire à agir à l'encontre de nos préférences.

> *« Je voudrais dire que la philosophie de l'approche par les préférences révélées sous-estime essentiellement le fait que l'homme est un animal social et que ses choix ne sont pas seulement liés de façon rigide à ses propres préférences. Je ne trouve pas difficile de croire que les oiseaux, les abeilles, les chiens et les chats révèlent leurs préférences dans leurs choix ; c'est avec les êtres humains que la proposition n'est pas particulièrement convaincante. Un acte de choix pour cet animal social est, en un sens fondamental, toujours un acte social. Il peut seulement être vaguement conscient des immenses problèmes d'interdépendance qui caractérisent une société […]. Mais son comportement est plus qu'une simple traduction de ses préférences personnelles. »*[211]

Dans cet article, l'argumentation de Sen conteste le caractère raisonnable de l'hypothèse de révélation des préférences par le choix. Il analysera plus tard[212] la situation suivante : au cours d'un dîner, un panier de fruits circule parmi les convives ; les deux derniers dîneurs doivent se partager une mangue et deux pommes. L'un d'eux préfère les mangues, mais est poli : il a adopté la norme de comportement qui consiste à se servir de manière à laisser au suivant le choix lorsque c'est possible. Il choisit donc une pomme tout en ayant une préférence pour la mangue. Un tel choix ne révèle pas les vraies préférences de l'agent, et cependant il fait partie

[211]SEN, Amartya, *ibidem.*

[212]SEN, Amartya, « Maximization and the Act of Choice", *Econometrica*, vol. 65, No.4 (juillet 1997), pp 745-779.Repris dans *Rationality and Freedom,* The Belknap Press of Harvard Universiy Press, 2002.

des expériences de choix que chacun peut comprendre et est susceptible d'avoir. En revanche, ce choix est révélateur des interactions sociales des questions d'interdépendance évoquées ci-dessus.

On pourrait objecter à Sen qu'il ne faut pas prendre ici le terme « préférence » trop au pied de la lettre, et que celui-ci ne renvoyait pour Samuelson et ses disciples qu'à la seule relation d'ordre mathématiques. Les préférences ne seraient qu'une représentation du choix. Mais les économistes font un double usage du concept de préférence : l'économie positive s'appuie sur la préférence pour décrire et anticiper les comportements des agents ; l'économie normative – qui se limite à l'économie du bien-être dans les décennies qui suivent l'article de Samuelson – en fait la base de ses jugements de valeur et de ses prescriptions en matière de politique sociale. Lorsqu'un économiste du bien-être dit qu'un agent, Sophie, préfère une option à une autre, il faut comprendre que cette option est meilleure pour Sophie que l'autre ; mais ce n'est pas nécessairement ce que Sophie choisit : en effet, l'existence même de l'économie normative et de la théorie du choix social présuppose le fait que la réalisation des préférences ne dépend pas que de l'agent individuel, mais aussi des institutions sociales ou des décisions collectives. L'état social devrait être tel que le choix de Sophie soit conforme à ses préférences, mais il ne l'est pas nécessairement. Le concept économique de préférence oscille donc entre l'évaluation et le choix. Ce que Sen critique ici, c'est un glissement notionnel progressif, de l'action vers le choix, du choix vers la préférence, de la préférence vers l'évaluation de l'agent, de l'évaluation de l'agent vers le jugement de valeur objectif. Soucieux de réintroduire une certaine complexité dans l'analyse économique, il liste les différentes interprétations possibles de la préférence d'une personne pour *x* plutôt que pour *y* :

> *« […] Il existe plusieurs propositions sur les intérêts, actions, etc, d'une personne, liées entre elles mais différentes l'une de l'autre, et qui ont besoin d'être distinguées, même si elles sont souvent confondues dans la littérature :*
>
> *(1) La personne obtient plus de satisfaction dans l'état x que dans l'état y (proposition sur la satisfaction ou le plaisir) ;*
>
> *(2) La personne pense qu'il ou elle est mieux loti(e) avec x qu'avec y (proposition sur le bien-être introspectif) ;*
>
> *(3) La personne est mieux lotie avec x qu'avec y (proposition sur le bien-être individuel, qui peut être introspectif ou pas) ;*
>
> *(4) La personne préfère que x plutôt que y advienne (proposition sur la condition mentale de préférence, ou désir, concernant les états) ;*

(5) *La personne aimerait choisir que x advienne plutôt que y (proposition sur le choix désiré)*

(6) *La personne croit qu'il serait bon [*right*] de choisir de sorte que x advienne plutôt que y (proposition sur le jugement normatif concernant le choix) ;*

(7) *La personne croit qu'il vaudrait mieux que x advienne plutôt que y (proposition sur le jugement normatif concernant les états de choses) ;*

(8) *La personne choisit de sorte que x plutôt que y advienne (proposition sur le choix effectif).*[213]

Sen ajoute que les relations entre ces propositions ne peuvent être déterminées *a priori.* Selon l'usage, explicatif ou normatif, que l'on réserve à la préférence, elles seront mises en relation de manière différente. Autrement dit, la préférence ne se suffit pas à elle-même : elle doit s'adosser à une analyse empirique ou normative plus large. Prenons quelques exemples.

Si nous affirmons que (4) implique (3), cela signifie que nous définissons le bien-être en termes de satisfaction des désirs, quels qu'ils soient. Si nous soutenons que (3) implique (5), nous adossons la préférence à une conception de la rationalité selon laquelle les êtres humains désirent toujours choisir ce qui améliore leur bien-être. Si notre analyse fait dépendre (8) de (2), nous attribuons aux croyances individuelles sur le bien-être une efficacité causale. Si nous supposons que (1) dépend de (2), nous jugeons que le bien-être ressenti est suffisant pour que la personne soit satisfaite. Si notre théorie stipule que (6) implique (4), cela signifie que le désir peut être déterminé par des jugements de valeur sur ce qu'il est correct (*right*) de choisir, quelle que soit la suite. Lorsque (2) implique (7), nous avons affaire à une conception égoïste de la préférence. L'idée qu'il faut retenir ici est la pluralité des théories susceptibles d'être accueillies par la « préférence ».

La critique de Sen a mis en évidence les risques d'un réductionnisme grossier. Et, plutôt que d'amalgamer les différentes significations de la préférence, Sen préconise de s'accommoder et même de profiter de la pluralité sémantique de la préférence.

> *« Cette pluralité est une source de puissance de la vaste catégorie des approches fondées sur les préférences. Dans différents types d'argumentations portant sur l'évaluation de décisions collectives appropriées, divers aspects de la volonté et de l'action d'un individu sont considérés – explicitement ou implicitement – et la richesse de la variété d'interprétations permet à la théorie d'invoquer différentes caractéristiques de l'individu, selon le contexte. »*[214]

Sen lui-même est susceptible d'utiliser le terme « préférence » dans des sens différents et incompatibles entre eux.

---

[213] SEN, Amartya, *Choice, Welfare and Measurement,* Oxford: Basil Blackwell, 1982, Introduction. Traduction libre..

[214] SEN, Amartya, *Rationality and Freedom, op. cit.,*p 242 de la traduction française.

4.1.3. La préférence, cause du choix ?

La critique de Sen ne lève cependant pas les ambiguïtés de la relation entre préférence et choix en économie normative. En effet, son approche ne nie pas la pertinence des préférences révélées. Elle se limite en fait à préciser qu'il existe d'autres manières de penser la relation entre préférence et choix. Mais des conceptions plus précises de cette relation peuvent être défendues. L'une d'elle consiste à penser la préférence comme la cause du choix.

a) La préférence comme cause du choix

Cette conception a été récemment défendue par le philosophe de l'économie Daniel Hausman. Dans son ouvrage consacré à la préférence, *Preference, Value, Choice and welfare,* paru en 2012, Hausman soutient de façon générale que les pratiques des économistes ne sont pas erronées en elles-mêmes. Ce qui ne va pas, c'est la manière dont les économistes eux-mêmes décrivent leurs pratiques. Ainsi, ils se trompent lorsqu'ils identifient préférences et choix.

Hausman critique la thèse selon laquelle les préférences sont identiques au choix, mais il affirme que la représentation de la préférence en elle-même n'est pas problématique, dès lors qu'elle est comprise comme une activité mentale d'évaluation et de classement des options et non comme l'acte de choix lui-même. Hausman réfute la thèse des préférences révélées par le choix en montrant que les croyances interfèrent dans la détermination du choix par la préférence. Il cite le cas de Roméo et Juliette : Juliette a simulé la mort afin de s'enfuir avec Roméo malgré l'interdiction de sa famille, mais une épidémie de peste arrête le messager chargé de mettre Roméo dans le secret. Roméo découvre ce qu'il croit être le cadavre de Juliette et s'empoisonne. Les options réelles offertes à Roméo étaient : fuir avec Juliette ; rester à Vérone ; mourir. Si Roméo n'avait pas été trompé par sa croyance, il aurait choisi selon sa préférence, qui aurait été de fuir avec Juliette.

On peut opposer à cet exemple l'objection selon laquelle l'ensemble d'options proposées au choix correspond à l'ensemble d'options faisables, *étant donné l'état des croyances de l'agent* : l'ensemble des options *faisables* offertes à un Roméo *qui croit Juliette morte* n'inclut alors pas la fuite avec Juliette. Hausman répond à cette objection en rappelant qu'un des objectifs principaux de la science économique est de prévoir les choix des agents, et cela ne peut se faire qu'en faisant l'hypothèse de préférences qui ne sont pas révélées par le choix.

Les joueurs de poker savent que, s'ils croyaient que leur adversaire préfèraient ce qu'ils manifestent dans leur jeu, ils perdraient bien souvent ![215]

> *« Si les économistes se limitaient aux préférences définies par les choix, ils ne pourraient pas faire de théorie des jeux »*[216]

Pour Hausman, la relation entre préférence et choix n'est donc pas définitionnelle, mais causale. Mais il ne suffit pas de dire que les préférences sont simplement des choix hypothétiques : la relation de causalité entre préférence et choix dépend de la croyance. L'analyse d'Hausman permet donc d'exclure l'interprétation n°8 de la préférence parmi celles proposées par Sen, puisque cette interprétation identifiait préférence et choix[217]. La préférence n'est plus de l'ordre de l'action, mais de l'activité mentale.

[215] Ou peut-être bénéficieraient-ils alors de la « chance du débutant », que l'on peut expliquer par le caractère improbable d'un comportement de choix révélant des préférences ?

[216] HAUSMAN, Daniel, *Preference, Value, Choice and Welfare*, Cambridge University Press, 2012, p 30.Traduction libre..

[217] cf. ci-dessus, la citation de l'extrait de l'introduction de *Choice, Welfare and Measurement*

b) La causalité comme rationalisation de l'action

La relation entre préférence et choix est donc causale et non définitionnelle. Une relation causale se caractérise par l'existence d'un lien entre cause et effet, la caractérisation de ce lien par l'assignation d'une direction, et enfin la désignation d'une cause plutôt que d'une autre.. Mais peut-on penser la relation entre une préférence et un choix en termes causaux ? La causalité est-elle un modèle satisfaisant pour expliquer et justifier l'action humaine ? N'appartient-elle pas plutôt au champ d'investigation des processus physiques et à une conception déterministe de l'explication ?

Un détour par la philosophie de l'action permet de mettre en évidence les arguments en faveur d'une relation causale entre préférence et choix. C'est le philosophe de l'action Donald Davidson qui affirme que la relation entre une raison et une action est causale[218]. La « rationalisation », l'explication d'une action par sa raison, est selon Davidson une espèce d'explication causale, et une raison d'agir comprend deux éléments :

a) Une attitude favorable (une « pro-attitude ») à l'égard d'une sorte d'action donnée, qui peut être un désir, une incitation, une envie, un principe moral ou esthétique, une convention…

b) Une croyance que l'action en question est de la sorte d'actions auxquelles l'agent est favorable.

Si l'on revient aux différentes interprétations de la préférence proposées par Sen dans *Choice, Welfare and Measurement*, il apparaît que chacune d'elles peut constituer une raison pour une action au sens de Davidson :

> *« 1) La personne obtient plus de satisfaction dans l'état x que dans l'état y (proposition sur la satisfaction ou le plaisir) ;*

2) La personne pense qu'il ou elle est mieux loti(e) avec x qu'avec y (proposition sur le bien-être introspectif) ;

3) La personne est mieux lotie avec x qu'avec y (proposition sur le bien-être individuel, qui peut être introspectif ou pas) ;

4) La personne préfère que x plutôt que y advienne (proposition sur la condition mentale de préférence, ou désir, concernant les états) ;

5) La personne aimerait choisir que x advienne plutôt que y (proposition sur le choix désiré)

[218] DAVIDSON, Donald, « Actions, Reasons and Causes" *in The Journal of Philosophy*, Vol. LX, No 23, nov. 1963, pp 685-700. Rééd.*In Essays on Actions and Events*, Oxford: Clarendon Press, 1980 ; traduction française par P. Engel, A*ctions et événements*, Paris, P.U.F., 1994.

6) *La personne croit qu'il serait bon [*right*] de choisir de sorte que x advienne plutôt que y (proposition sur le jugement normatif concernant le choix) ;*

7) *La personne croit qu'il vaudrait mieux que x advienne plutôt que y (proposition sur le jugement normatif concernant les états de choses) ; »*[219]

La relation causale entre préférence et choix est donc une relation entre une évaluation, la préférence, et une disposition à agir, suivie d'une action spécifique sélectionnée à partir d'un ensemble d'options. L'action ainsi expliquée est décrite de telle sorte que la connexion avec sa raison soit signifiante. Une raison justifie une action, dit Davidson, dans la mesure où elle a un rôle explicatif ; elle permet de reconstruire le raisonnement qui attribue à l'action la caractéristique d'être désirable.

Mais la justification est-elle une explication causale ? Davidson affirme que la rationalisation, c'est-à-dire l'explication d'une action par sa raison, est une espèce (*species*) d'explication causale. La justification est la propriété de la rationalisation qui la différencie d'autres espèces d'explication. Connaître la raison d'une action, c'est être pourvu d'une interprétation, d'une description nouvelle de cette action qui la rend intelligible et permet de la replacer « dans un contexte social, économique, linguistique ou évaluatif plus large ». Connaître la raison du suicide de Roméo, c'est savoir qu'il ne peut vivre sans Juliette – son attitude – et qu'il croit celle-ci morte – sa croyance - . Le suicide de Roméo paraîtrait absurde aux yeux d'un observateur ignorant des circonstances du drame, mais sa raison permet d'en donner une description nouvelle qui fait sens pour les spectateurs de la tragédie.

L'explication causale ne renvoie donc pas ici à la description d'un processus physique. Il n'y a pas de production du choix par la préférence. Son rôle est de connecter une raison et une action, ou, dans notre cas, une préférence et un choix[220]. Cette connexion est de l'ordre du langage et s'effectue par la description recommencée de l'action justifiée par sa raison.

[219] SEN, Amartya, *Choice, Welfare and Measurement,* Oxford: Basil Blackwell, 1982, Introduction. Traduction libre.

[220] Le choix n'étant pas un événement, mais une disposition, la connexion semble encore plus aisée.

c) Deux objections à la thèse de la relation causale

La thèse selon laquelle la préférence est la cause du choix fait face à deux objections. La première concerne le problème de la caractérisation d'une relation entre des attitudes humaines en termes de causalité. Cette objection peut être surmontée. La seconde, mise en évidence par Davidson lui-même, met en évidence les problèmes créés par l'isolation de la relation causale entre préférence et choix. Cette objection met véritablement en difficulté la conception causale ou dispositionnelle de la relation entre préférence et choix.

Mais la mise en relation de la cause et de l'effet ne présuppose-t-elle pas l'existence d'une loi ? Peut-on définir des lois articulant certaines préférences à certains choix ? Et, si oui, quelles sont-elles ? En réponse, Davidson fait remarquer que, si chaque explication causale requérait une loi capable de prédire l'occurrence de l'effet, bien peu d'explications causales seraient possibles[221]. Il cite les prévisions météorologiques à cet effet : si celles-ci nous paraissent approximatives, c'est parce que les descriptions des événements qui nous intéressent (la pluie, le beau temps, le froid…) sont difficiles à mettre en relation avec les concepts utilisés pour formuler des lois capables de prévoir des changements. De même, le raisonnement pratique ne peut être mis en relation que de façon approximative avec un modèle prédisant les actions futures.

Mais Davidson lui-même a mis en évidence un autre problème auquel est confrontée la thèse de l'efficacité causale de la préférence.

Lorsque nous affirmons que la préférence est une disposition à choisir, ou encore est cause du choix, nous nous exposons à l'objection suivante : réduire la relation entre préférence et choix à une relation de causalité occulte l'influence d'autres attitudes, non seulement des croyances, mais également des autres préférences. La description strictement causale de la relation entre préférence et choix court le risque d'être trop abstraite et de ne pas rendre compte avec suffisance de tous les facteurs qui expliquent une action. Davidson propose le cas suivant afin de clarifier cette question : il s'agit de deux variantes du mythe d'Œdipe. Dans la première, Oedipe a de bonnes raisons de tuer son père, et ensuite il tue son père. Mais il ne savait pas que l'homme qu'il a tué était son père. Le meurtre de Laïos n'est donc pas causé par les raisons qu'a Œdipe de tuer son père. Dans la seconde variante, Œdipe se précipite sur la route pour tuer son père (celui qu'il croit être son père), et tue l'homme qu'il rencontre sur la route (qui s'avère être son père) et qui l'empêchait de passer. La cause du meurtre de Laïos est

[221] *Ibidem.*

effectivement le désir d'Œdipe de tuer son père, et cependant nous ne pouvons pas dire qu'il a tué Laïos parce qu'il voulait tuer son père (il l'a fait parce qu'il voulait passer). Davidson en conclut que les croyances, désirs et causes considérés isolément du processus de raisonnement et du système de croyance de l'agent ne peuvent offrir une explication suffisante de l'action.[222]

On ne peut citer les travaux de Davidson à l'appui de notre réflexion sans rappeler sa position à l'égard des théories de la décision. Lorsque les hypothèses comportementales de la théorie de la décision sont testées expérimentalement, Davidson, ayant lui-même contribué à ce type d'expériences, affirme que les chercheurs peuvent toujours retrouver ces hypothèses de rationalité dans cette expérience. Pourquoi ? Lorsque le seul fondement des assertions sur les désirs et les croyances des agents est l'observation de leur comportement, plusieurs interprétations sont toujours possibles. Néanmoins, l'interprétation choisie dans ce contexte s'appuie sur un fond de croyances et de désirs que l'expérimentateur peut lui-même comprendre. Le problème de l'interprétation du comportement est analogue à celui de l'interprétation du discours. Le point de départ de l'interprétation d'un discours n'est pas l'analyse des termes de celui-ci, pris un à un ; le point de départ est le système de croyances du locuteur et ce qu'il tient pour vrai. Nous avons besoin de postuler que ce système de croyances est cohérent. Même si nous identifions des erreurs dans le discours de la personne, cela n'est possible que parce que nous sommes déjà en possession d'un accord minimal sur certaines croyances ou propositions (sur l'existence des objets désignés, sur la valeur de la vérité, etc). Si mon interlocuteur me dit : « la Terre est plate », je peux comprendre et interpréter ce qu'il dit même si je juge que c'est faux, et ce parce que je suppose que nous sommes d'accord sur ce que signifie le mot « Terre », sur le fait qu'une proposition décrivant la Terre doit avoir une valeur de vérité et peut faire l'objet d'une investigation scientifique…[223]. Lorsque nous interprétons une action, nous voyons les autres « comme nous-mêmes » : notre interprétation s'appuie sur un fond de croyances et de désirs que nous sommes également susceptibles d'avoir et d'éprouver. Nous décidons de voir les autres comme rationnels, ou en partie rationnels, ou susceptibles d'être rationnels, et comme sujets à une évaluation morale. Les généralisations de la psychologie dépendent donc de ce

[222] Cette conclusion s'écarte de la version primitive de l'article de 1963, dans lequel Davidson pensait pouvoir offrir une formulation des causes *suffisantes* de l'action. Cf. DAVIDSON, Donald, *Essays on Actions and Events,* Oxford: Clarendon Press, 2001, Introduction.

[223] Néanmoins, il n'est pas impossible que mon interlocuteur pense en fait que la Terre est une assiette, ou encore qu'il s'agit d'un objet fictif, comme une licorne, auquel il est libre d'attribuer telle ou telle propriété. De telles hypothèses sont intéressantes pour les philosophes, amusantes pour les artistes, mais, les prendre au sérieux dans notre pratique quotidienne, c'est peut-être mettre en danger notre santé mentale.

« biais »[224]. La « redescription » de l'action sous la forme d'une explication causale, qui donne la raison principale de l'action, dépend d'une suite d'idées, de raisonnements, n'appartenant pas à cette relation causale.

[224]DAVIDSON, Donald, "Psychology as Philosophy", *in* BROWN S. C. (éd), *Philosophy of Psychology*, The Macmillan Press and Barnes, Noble, Inc, 1974.

4.1.4. La préférence fonctionnelle

Comment penser la relation entre préférence et choix ? Nous avons vu que l'on ne peut sans implications problématiques identifier préférence et choix. De même, penser la relation entre les deux comme une relation causale isolée mène à des conclusions paradoxales. Il nous faut donc décrire la relation entre préférence et choix de manière à prendre en compte les différents facteurs susceptibles d'influencer cette relation. Le concept de « préférence fonctionnelle » développé par Philip Pettit pourra nous y aider[225].

Pettit reconnaît sa dette à l'égard de Daniel Hausman pour la compréhension de la préférence. Souhaitant articuler la préférence à ses travaux sur la délibération entrepris en collaboration avec Michael Smith, il entreprend de préciser les différents types de relation qu'entretient la préférence avec d'autres expériences et d'autres états mentaux. Soucieux de ne pas exposer son analyse de la préférence aux objections qui visent la réduction de celle-ci à une disposition à choisir toujours efficace, il distingue sa méthode de conceptualisation de la préférence de deux autres :

- La conception comportementale [*behavioral*], dont nous avons présenté la critique faite par Sen ci-dessus, affirme que la préférence est, et n'est que ce qui est révélé par le comportement de choix ; on y reconnaît les thèses du courant behaviouriste[226]. Pettit souligne le caractère étrange de la théorie : il n'existerait pas de préférence en dehors du comportement de choix, même lorsque des obstacles de diverses natures empêchent l'agent de choisir ce qu'elle préfère. Si cette approche a permis des développements mathématiques très élégants, ajoute Pettit, elle ne rend en revanche aucunement compte de la notion commune de préférence. Sen a mis en évidence le caractère réducteur et insuffisant de cette méthode.
- La conception dispositionnelle de la préférence comprend celle-ci comme une disposition à choisir. Lorsqu'un agent préfère *x* à *y*, il est disposé à choisir *x* plutôt que *y*. Cette conception conserve le lien causal entre préférence et choix qui constitue l'hypothèse psychologique implicite de la théorie des préférences révélées, tout en reconnaissant que les préférences puissent rencontrer des obstacles empêchant leur

[225]PETTIT, Philip, « Preference, Deliberation and Satisfaction", *in* OLSARETTI, Serena (ed.), *Preferences and Well-Being*, Cambridge University Press, 2006.

[226] Cette conception est associée à la préférence révélée, mais Pettit souligne que tous les partisans de cette dernière ne sont pas nécessairement behaviouristes (ils ont pu, comme notre objecteur, adopter la préférence révélée pour traiter certaines questions bien définies sans pour autant défendre le behaviourisme dans toutes les sciences de l'Homme et de la société.

réalisation. Mais Pettit la juge insuffisante, dans la mesure où elle se focalise uniquement sur ce lien, faisant abstraction des relations entre la préférence et d'autres attitudes[227], telles que la croyance. Il s'agit ici de la conception défendue par Davidson.

- La conception fonctionnelle de la préférence, que Pettit identifie comme sienne, permet de définir une préférence pour *x* plutôt que pour *y* même si celle-ci ne dispose pas l'agent à choisir *x* plutôt que *y*. La préférence cesse d'être une simple disposition à choisir lorsqu'elle est liée à d'autres attitudes : autres préférences, croyances... Par exemple, même si je préfère vivre longtemps en bonne santé, mes choix ne procèdent pas de cette préférence si en plus de cela je préfère fumer des cigarettes. De même, si je préfère lire *Guerre et Paix* plutôt qu'un autre livre, et si je crois que *Guerre et Paix* se trouve à la bibliothèque, je m'y rendrai. Mais c'est conjointement à la croyance que ce livre se trouve en ce lieu que je me rends à la bibliothèque, et non pas parce que je préfère me rendre à la bibliothèque plutôt que de rester chez moi.

> *« D'après une telle analyse, dire que quelqu'un préfère une option à une autre revient à dire qu'il est dans un état tel que, en l'absence de conditions perturbatrices, cet état le disposera à choisir la première option plutôt que la seconde, et se liera de telle et telle manière avec d'autres préférences et d'autres états mentaux. Les préférences, en gros, ne sont pas seulement n'importe quelle disposition à choisir ; ce sont des dispositions qui sont collatéralement sensibles à divers autres états. »*[228]

A la différence de la préférence pensée comme seule cause du choix, la préférence fonctionnelle est « collatéralement sensible » à d'autres états. Cela signifie que c'est conjointement à d'autres états mentaux, à d'autres attitudes, que la préférence peut nous disposer à choisir..

La préférence fonctionnelle, en étant collatéralement sensible à d'autres attitudes de l'agent, et en n'étant pas directement causalement efficace, peut figurer comme la composante d'une délibération dont le rôle ne se limite pas à apporter des informations, mais aussi à clarifier et à résoudre des conflits de valeurs, qu'ils s'agissent de conflits intrapersonnels ou interpersonnels.

Revenons à l'exemple d'Œdipe proposé par Davidson afin de comprendre l'intérêt et la portée du concept de préférence fonctionnelle. Lorsqu'Œdipe a l'intention de tuer celui qu'il croit

[227] Pettit parle également d' « états mentaux » (*states of the mind*), sans surdéterminer son approche. Il nous semble que cette analyse est compatible avec des conceptions assez différentes de l'esprit, allant du cerveau des neurosciences à la conscience de la phénoménologie.

[228] *Ibidem*. Traduction libre.

être son père, et tue par accident son véritable père, nous ne pouvons pas dire qu'il a tué ce dernier parce qu'il voulait tuer son père (il l'a fait parce qu'il voulait passer). En revanche, si sa préférence pour le meurtre de son père est décrite comme collatéralement sensible à certaines croyances, telles que la croyance que son père n'est pas l'homme qu'il a croisé au carrefour, nous pouvons éviter le paradoxe mis en évidence par Davidson. De même, si cette la description de cette préférence est mise en relation avec la différence entre la colère profonde qui le motive à vouloir la mort de son père et l'accès de fureur passager qui l'entraîne à tuer Laïos, nous éviterons les conséquences problématiques d'une description de la relation entre préférence et choix en termes causaux.

4.2. Préférences et évaluations : des jugements de valeur comparatifs

Ayant proposé un mode de description de la relation entre préférence et choix susceptible de vaincre certaines des résistances des avocats des capabilités, il nous faut à présent statuer sur la nature de la préférence proprement dite. Dans cette section, nous tenterons de montrer que la préférence est un jugement de valeur comparatif. Cette conception de la préférence peut être associée à l'approche des capabilités, comme le verrons plus loj, dans la sous-section 7.2.2.. Notre argumentation procédera en quatre étapes. Nous présenterons les raisons qui nous mènent à identifier la préférence à un jugement – ainsi que la conception du jugement que cela sous-tend – (4.2.1.), avant d'étudier les déterminations successives de ce jugement, à savoir le fait qu'il s'agit d'une comparaison de descriptions (4.2.2.), que cette comparaison est évaluative et non factuelle (4.2.3.), et qu'une évaluation comparative doit être distinguée d'une évaluation classificatoire (4.2.4.).

4.2.1. La préférence est un jugement

Dans l'article « Social Choice Theory : a Reexamination »[229], Sen défend la possibilité d'interpréter les préférences dans la théorie du choix social soit comme des jugements, soit comme des intérêts. Une étude plus approfondie des préférences de la théorie du choix social montre qu'elles sont une espèce de jugement. Nous dégagerons la conception du jugement que nous utilisons ici : il s'agit de la conception frégéenne du jugement comme assertion.

[229] SEN, Amartya, « Social Choice Theory : A Reexamination", *Econometrica*, 45, Janvier 1977, pp 53-89. Rééd.*In Choice, Welfare and Measurement*, Oxford: Basic Blackwell Publishers, 1982

Dans le formalisme de la théorie du choix social, la préférence est une relation entre deux éléments ; ces deux éléments appartiennent à un ensemble d'états sociaux concevables. La relation de préférence ordonne ces éléments : soit un élément est classé au-dessus de l'autre (il est dit alors « être préféré » à l'autre), soit il est jugé équivalent (on parle alors d' « indifférence »). La préférence correspond ainsi à une opération cognitive consister à apprécier les propriétés respectives de chaque élément et à les comparer en fonction de critères variés[230]. Du point de vue de la description formelle, les préférences sont des relations exprimées sur un ensemble d'états sociaux.

Cette opération cognitive peut être qualifiée de jugement, dans la mesure où le jugement est l'affirmation ou la négation d'un contenu. Plus précisément, la conception du jugement que nous retiendrons ici est la conception frégéenne de l'assertion. En effet, l'assertion est indifférente à la forme logique du jugement : la proposition assertée peut être positive, négative, conditionnelle, possible, nécessaire, l'assertion fonctionne toujours. Une telle souplesse nous autorise ainsi à transcrire les préférences en termes de jugements.

La conception du jugement comme assertion a été développée par Gottlob Frege[231]. Dans son idéographie, langage formalisé qu'il a inventé et qui n'est plus utilisé aujourd'hui, Frege introduit la barre du jugement (|), qui représente le fait que le contenu de la proposition est asserté pour vrai[232]. Comme le montre l'analyse d'Ali Benmakhlouf, la force d'assertion ne peut résider ni dans l'ajout d' « il est vrai que » à une proposition donnée, ni dans la copule, ni dans le prédicat. Ajouter une barre de jugement à la proposition n'est donc pas superflu[233].

L'acte d'assertion, poursuit Ali Benmakhlouf, correspond à celui de la « clôture » du jugement, « c'est un signe qui ne laisse plus ouverte la proposition à des débats futurs ou à des révisions »[234]. Cela ne signifie pas que l'acte d'assertion équivaut à un simple acte d'assentiment, d'ordre psychologique. L'assertion n'est pas arbitraire ; elle ne traduit pas un « point de vue » ou un « ressenti », mais plutôt la conclusion des investigations.

Les étapes de cette investigation sont distinguées sur le plan conceptuel par Frege dans ce passage, que nous citons :

> *« On distinguera donc :*

[230] La question de la définition de ce critère fait l'objet de la section 5.2..

[231] Sur l'intérêt de la conception frégéenne du jugement pour la théorie de l'agrégation logique, cf. MONGIN, Philippe et DIETRICH, Franz, « Un bilan interprétatif de la théorie de l'agrégation logique », *Revue d'économie politique*, 120 (6), pp 929-972.

[232] FREGE, Gottlob, *Begriffsschrift*, 1879. Trad. Fr. BESSON, Corine, *Idéographie*, Paris : Vrin, 1999.

[233] BENMAKHLOUF, Ali, *Frege. Le nécessaire et le superflu*, Paris : Vrin, 2002, pp 194-196.

[234] *Ibidem*, p 196.

1. *La saisie de la pensée – l'acte de penser*
2. *La reconnaissance de la vérité d'une pensée – le jugement.*
3. *La manifestation de ce jugement – l'affirmation »*[235]

Selon Frege, la saisie de la pensée correspond à la première étape d'une démarche scientifique et peut-être exprimée sous forme de proposition interrogative. Dans le cas des préférences – que Frege ne traite pas – la saisie de la pensée correspondrait à la comparaison entre deux états sociaux ou deux options. En reprenant notre exemple, elle pourrait être exprimée ainsi :

- « le chocolat est-il préféré à la pomme ? ».

La deuxième étape consiste à reconnaître la vérité de la proposition pensée. Ainsi, l'agent pourrait reconnaître que le chocolat est préféré à la pomme. La préférence correspond donc à la reconnaissance de la vérité d'une comparaison en termes de préférences. La troisième étape, la manifestation du jugement – ici, de la préférence – est son expression.

La préférence peut alors être définie comme l'assertion d'un classement[236]. « Face à » toutes les propositions énonçant des relations d'ordre entre tous les états sociaux concevables, l'agent « approuve » certaines d'entre elles : ce sont ses préférences. La préférence est donc un jugement en tant qu'acte d'assertion de certaines relations d'ordre. L'acte d'assertion peut être décrit comme l'appartenance à un ensemble :

- Soit X, ensemble de tous les états sociaux concevables ;
- Soit Ξ, l'ensemble de toutes les préférences logiquement possibles sur X,
- Soit A_i le sous-ensemble des préférences que l'agent i reconnaît comme siennes (ou asserte) ; A_i est un sous-ensemble de Ξ.
- Sachant que $X = \{x, y, z\}$, si l'agent i préfère faiblement x à y et y à z, alors $A_i = \{xR_iy, yR_iz\}$.

L'appartenance des relations de préférence xR_iy, et yR_iz à l'ensemble A_i signifie que l'individu i asserte ces relations de préférences.

Cette conception distingue le jugement de la proposition. Toute proposition n'est pas nécessairement transcrite dans un jugement. Si la préférence est la reconnaissance de la vérité d'une comparaison en termes de préférences, alors il est faut distinguer les préférences

[235] FREGE, Gottlob, *Recherches logiques,* trad. Claude Imbert, *in* FREGE, *Ecrits logiques et philosophiques*, Paris, Seuil, 1971.
[236] L'axiomatique définit ensuite les propriétés de ce classement : complétude, transitivité, réflexivité, etc.

assertées des préférences non assertées. Cela nous amène à distinguer les préférences exprimables des préférences exprimées. Les préférences exprimables sont des préférences logiquement possibles, mais qui n'ont pas encore été assertées. Les préférences exprimées sont des préférences qui ont été assertées. Afin de mettre en évidence le sens de cette distinction, examinons les deux propositions suivantes, dans lesquelles la négation est située soit au niveau de la préférence (a), soit au niveau de l'assertion (b) :

(a) Sophie juge qu'elle ne préfère pas les pommes au chocolat ;

(b) Sophie ne juge pas qu'elle préfère les pommes au chocolat.

La proposition (a) est l'assertion de la négation d'une préférence. Cette « non-préférence » est exprimée. La proposition (a) fait partie des préférences que Sophie reconnaît comme siennes. Elle peut être impliquée par l'une des deux propositions suivantes[237] :

(c) Sophie juge qu'elle préfère le chocolat aux pommes ;

ou

(d) Sophie juge qu'elle est indifférente entre le chocolat et les pommes

En revanche, la proposition (b) dénote la négation de l'assertion d'une préférence. (b) signifie que la comparaison en termes de préférences n'a pas, ou pas encore été faite. Il n'y a pas encore lieu de statuer sur la vérité des préférences ou de l'indifférence de Sophie à l'égard du chocolat ou des pommes. En d'autres termes, le débat est encore ouvert.

Cette conception permet également de distinguer clairement le contenu descriptif de la préférence avec la préférence proprement dite. Préférer, c'est adopter une certaine attitude à l'égard de deux contenus mis en relation par une comparaison, et cette comparaison ne détermine pas en tant que telle la préférence.

[237] Nous supposons ici que la relation de préférence est complète, c'est-à-dire que Sophie statue sur toutes les paires d'options.

4.2.2. Un jugement comparatif

La préférence est un jugement issu d'une comparaison d'états de choses. Ces états peuvent être des états sociaux (c'est le cas de la théorie du choix social), des paniers de biens (en économie) ou plus généralement des mondes possibles.

La préférence dépend donc d'une description empirique de ces états sociaux. Cette description devrait être aussi complète que possible, incluant, comme le rappelle Arrow, les « statues d'hommes célèbres »[238]. Mais ce petit exemple montre justement que toute description est une sélection : les statues d'hommes célèbres ne sont pas toujours un élément d'information pertinent pour le choix social. Comme le dit Sen, la description suppose un choix[239]. Lorsque nous décrivons un état social, nous sélectionnons à l'avance les caractéristiques que nous jugeons pertinentes pour cette description. Si nous commençons ici l'analyse de la préférence en énonçant qu'elle repose sur des descriptions, cela ne correspond pas pour autant à une chronologie de la préférence : il n'y aurait pas d'abord des descriptions, suivies ensuite d'évaluations. En fait, le type d'évaluation que nous souhaitons mener détermine déjà en grande partie la méthode de description adoptée. Il ne faudrait pas non plus en inférer que la description ne sert que de contenu, voire de prétexte à l'évaluation : lorsqu'on veut évaluer la pauvreté, on s'intéressera aux revenus plutôt qu'aux statues d'hommes célèbres, sans pour autant avoir l'intention de désigner à l'avance telle ou telle classe de personnes comme « pauvre ».

Dans la mesure où la préférence repose sur une description des états sociaux proposés, elle dépend intrinsèquement de la connaissance que nous avons de ces états sociaux, ou plus exactement des croyances que nous nourrissons à leurs égards. Si, comme nous le montrons dans le paragraphe suivant, les préférences relèvent du jugement de valeur, elles n'en sont pas moins étroitement liées à nos jugements concernant les faits présents et à venir. Ainsi, de nouvelles informations sur un état social peuvent nous amener à modifier nos préférences, et certaines préférences ne sont pas compatibles avec certaines croyances. Comme l'écrivent Grüne-Yanoff et Hansson :

> *« Il s'en suit que, si une personne croit (correctement) que Le Corbusier et Charles-Edouard Jeanneret-Gris sont une seule et même personne, alors il serait incohérent de sa part de préférer les*

[238]ARROW, Kenneth, *Social Choice and Individual Values*, New Haven, Yale university Press, 1951 (2ème édition révisée, 1961).Traduction française par l'Association de Traduction Economique de l'Université de Montpellier, Paris, Calmann-Levy, 1974

[239]SEN, Amartya, "Description as a choice", *Oxford Economic Papers,* New Series, Vol. 32, No 3, Nov. 1980, pp 353-369.

maisons construites par Le Corbusier aux maisons construites par Charles-Edouard Jeanneret-Gris. »[240]

La préférence suppose la description non pas d'un état social, mais de deux. Ainsi, le jugement de préférence dépend de la mise en rapport de ces deux descriptions. En ce sens, la préférence est moins immédiate que le désir pour un état, ou l'approbation d'un état. Pour exprimer une préférence, nous devons avoir réalisé une opération cognitive de mise en relation de deux descriptions d'états sociaux possibles.

Afin qu'une telle opération ait un sens, les deux états sociaux ne doivent pas être identiques. Ils ne doivent même pas être des états identiques présentés sous des dénominations différentes, comme dans l'exemple où Le Corbusier et Charles-Edouard Jeanneret-Gris sont considérés comme des personnes différentes. Les états sociaux sur lesquels porte la comparaison sont mutuellement exclusifs. Si deux états sociaux diffèrent, ne serait-ce que d'une caractéristique, la réalisation de l'un exclut la réalisation de l'autre.

Mais on peut se demander si deux états sociaux ainsi comparés ne doivent pas être *trop* différents. Dans les *Topiques*, Aristote décrit la préférence comme un des « lieux » courants du discours, et l'examen du préférable portera « sur des choses voisines et au sujet desquelles nous sommes dans l'incertitude pour savoir à laquelle il faut accorder la préférence, du fait que nous n'apercevons aucune supériorité de l'une sur l'autre. »[241]. Mais comment rendre compte des « petites » ou des « grandes » différences dans la représentation formelle des préférences ? Il apparaît ici que la notion logique de préférence diffère quelque peu de la notion intuitive. La préférence, au sens logique, suppose une comparaison entre deux états sociaux *distincts*, et le degré de différence de ces états sociaux n'importe pas ; en revanche, l'usage du terme préférence – et c'est à cela que se réfère Aristote dans les *Topiques* – restreint les comparaisons pertinentes à celles effectuées entre des états sociaux aux caractéristiques proches. Cette différence entre la structure logique de la préférence et son usage effectif peut créer des difficultés d'application : le cadre logique risque en effet d'être trop général et d'imposer des conditions qui ne sont pas nécessaires au traitement du problème auquel nous avons affaire.

De cela suit que la description des états sociaux à propos desquels nous exprimons une préférence passe au crible, d'une part, du projet d'évaluation qui a motivé la description, et

[240] GRÜNE-YANOFF, Till, et HANSSON, Sven Ove, « From belief Revision to Preference Change », in GRÜNE-YANOFF, Till, et HANSSON, Sven Ove, *Preference Change. Approaches from philosophy, economics and psychology*, Springer, 2009. Traduction libre..

[241] ARISTOTE, *Topiques*, 116b.

d'autre part, de la mise en rapport avec un autre état social. Cette mise en rapport met en évidence les ressemblances et les différences entre ces états sociaux. La préférence suppose donc une opération d'abstraction : nous abstrayons les caractéristiques pertinentes pour la comparaison de la description des états sociaux.

4.2.3. Jugements de fait comparatifs et jugements de valeur comparatifs

Mais la préférence ne se contente pas de dégager les ressemblances et les différences des états comparés. C'est l'évaluation qui motive la comparaison. Préférer *x* à *y*, c'est affirmer que *x* est préférable à *y* , que *x* possède des caractéristiques qui le rendent supérieur à *y* en termes de valeur. Une autre formulation identifie le fait de préférer *x* à *y* au fait de juger que *x* est meilleur que *y*.

Une première difficulté consiste en ce que les comparaisons en termes de préférences peuvent être justifiées par des valeurs substantiellement différentes. Ainsi, une comparaison en termes de justice ne correspond pas nécessairement à une comparaison du point de vue du meilleur. Reprenons l'histoire biblique du jugement de Salomon : deux femmes affirment être la mère d'un petit enfant. Afin de trancher, Salomon décide de couper l'enfant en deux et d'en attribuer une moitié à chacune des femmes ; la vraie mère alors supplie le roi de donner l'enfant à l'autre plutôt que de le tuer, s'identifiant comme mère, c'est-à-dire comme seule réellement soucieuse du bien de l'enfant. Au moment où Salomon prononce sa sentence, il affirme qu'il est plus juste de couper l'enfant en deux plutôt que de le donner à une des deux femmes sans être certain qu'il s'agit vraiment de la mère. Mais la valeur invoquée par cette préférence, la justice, diffère substantiellement de la valeur du jugement que la vraie mère exprime en suppliant le roi d'épargner l'enfant : pour l'enfant – et pour celle qui aime l'enfant – il est meilleur qu'il soit épargné, y compris au prix d'une décision injuste. Une préférence peut ainsi être justifiée par des valeurs substantiellement différentes (Salomon finit d'ailleurs par rendre l'enfant à la vraie mère, qu'il a reconnue grâce aux supplications de cette dernière).

La difficulté ne s'arrête pas ici : dans le langage courant, la préférence est utilisée fréquemment dans des contextes indépendants de la morale ou des jugements de valeur : lorsque Sophie préfère la robe bleue à la robe rouge, sa préférence est dictée par des considérations esthétiques[242]. Peut-on assimiler le jugement de Sophie :

[242] Nous supposons ici que Sophie sélectionne la robe qu'elle pense être la plus jolie, sans se soucier d'une série d'autres facteurs qui interviennent fréquemment dans le choix des vêtements, comme le souci des normes sociales, le prix…

(a) La robe bleue est plus belle que la robe rouge.

au jugement

(b) La robe bleue est préférable à la robe rouge ?

Certainement : entre la robe bleue et la robe rouge, Sophie sélectionne la première. Mais (a) équivaut-elle à :

(c) La robe bleue est meilleure que la robe rouge ?

Le terme « meilleur » fait entrer des considérations morales dans l'évaluation de Sophie, ce qui, dans la plupart des contextes, est inapproprié. La comparaison en termes de préférences peut donc être déterminée par des valeurs indépendantes de la morale[243].

Si le critère de « préférabilité » n'est pas circonscrit, même des comparaisons factuelles pourraient être confondues avec la préférence telle que la représente formellement la théorie du choix social. Intuitivement, la proposition suivante ne nous semble pas être une préférence :

« Paris est plus au Nord de la France que Nice »

Et cependant, elle pourrait être représentée par une relation d'ordre, à l'instar des préférences de la théorie du choix social. Le critère de classement serait la localisation géographique en Europe. Nous avons ici un problème : qu'est-ce qui fait qu'un critère de classement est un critère de préférence ? Quel critère nous permet de trancher et d'identifier, parmi les jugements comparatifs, ceux qui sont à proprement parler des préférences ? Le critère du « meilleur » est vague.

Si nous reprenons l'exemple de la comparaison entre Paris et Nice, il nous semble curieux d'affirmer que le fait que Paris soit au Nord de Nice soit constitutif de son caractère « préférable ». Ici, quelques éléments de clarification sont nécessaires. L'exemple suivant nous permet de comprendre ce qui distingue les critères de comparaison factuelle des critères de préférence. Il arrive que nous disions que, pour une personne donnée que nous prénommerons Arlette, Paris est préférable à Nice parce que située plus au Nord : Arlette préfère Paris à Nice parce que Paris est plus au Nord. Mais cela ne fait pas de la localisation de Paris le caractère constitutif de la préférence pour Paris. Supposons par exemple qu'Arlette

[243] On pourrait cependant avancer que la distinction entre valeurs éthiques et valeurs esthétiques n'a pas toujours été aussi évidente, et date du manifeste de l'Art pour l'Art. En ce cas, le problème des comparaisons en terme de beauté et de goût est soit réductible à celui des comparaisons en termes de « meilleur », soit similaire à celui des comparaisons faisant appel à des valeurs substantielles autres que le bien ou le bon. Néanmoins, il nous semble que l'idée selon laquelle les jugements de goût relèvent d'une catégorie différente de celle des jugements moraux nous semble assez largement répandue pour que le problème que nous traitons puisse être bel et bien considéré comme pertinent.

n'aime pas la chaleur l'été. Paris étant plus au Nord, elle peut y éviter la canicule l'été. Ce n'est donc pas le fait que Paris soit plus au Nord qui rend Paris préférable à Nice. C'est le fait que cette localisation de Paris rende Paris plus fraîche et donc plus apte à réaliser le but d'Arlette, qui est d'éviter la chaleur et de veiller à son confort. Si la localisation de Paris au Nord de Nice rend Paris préférable à Nice aux yeux de cette personne, c'est parce que cette localisation rend Paris plus apte à réaliser une fin bonne pour Arlette, ou jugée comme telle. Autrement dit, ce qui rend un objet préférable à un autre, c'est son aptitude à réaliser une fin.

Afin de qualifier plus précisément cette caractéristique de la préférence, nous mobiliserons ici l'idée de structure téléologique des comparaisons que le philosophe et économiste John Broome a développée. Les préférences, en tant que jugements de valeur comparatifs, sont *téléologiques*. Broome propose de définir le bien tel qu'il apparaît dans les jugements comme :

> *« x est meilleur que y »,*

non par son contenu, mais par sa structure.

> *« Une théorie téléologique implique qu'entre les actes, il y a une relation « est meilleur que » (betterness relation) : _ est au moins aussi bon que _ (où les espaces sont à compléter par des actes), et que cette relation détermine la rectitude (rightness) de l'acte. Une relation « est meilleur que » a nécessairement une structure particulière. »*[244]

La structure de la relation « est meilleur que » est celle d'un ordre, ajoute Broome : c'est aussi celle de la « préférence » dans la théorie du choix social.

Du point de vue des propriétés formelles, il s'agit d'un quasi-ordre, dans la mesure où cette relation devrait être réflexive et transitive, mais pas nécessairement complète, puisqu'il existe des biens incommensurables. Rappelons qu'une relation d'ordre réflexive est une relation telle qu'un des *relata* ne peut être strictement préféré à lui-même, mais peut lui être indifférent, tandis qu'une relation d'ordre transitive est une relation d'ordre telle que, pour tout A, B et C, si A est classé au-dessus de B et B au-dessus de C, alors A sera classé au-dessus de C. Broome considère que ce sont des vérités logiques, pouvant être établies *a priori.*

> *« Une éthique téléologique, donc, dit qu'il y a un ordre d'actes qui détermine la rectitude de cet acte. Et c'est en fait suffisant pour définir une éthique téléologique. Toute théorie éthique avec cette implication est téléologique. Je définis donc la téléologie par sa structure. La téléologie standard a une structure*

[244]BROOME, John, *Weighing Goods*, Oxford : Blackwell, 1991, p 11. Traduction libre.

maximisante : elle dit qu'il y a un ordre d'actes tel que, si un acte parmi ceux qui sont disponibles est classé devant les autres dans l'ordre, alors cet acte devrait être fait. »[245]

Broome défend donc la thèse selon laquelle, au sein de chaque théorie éthique téléologique, c'est l'ordre qui définit une conception du bien. L'ordre, le fait de hiérarchiser les actes – ou les états de choses, en l'occurrence – constitue la structure même de la relation « est meilleur que ». Cette conception de la téléologie est structurelle et non substantielle.

Juger une option préférable à une autre du point de vue d'un critère – justice, bien, bien-être, beauté -, c'est juger que cette option est davantage susceptible de réaliser un monde correspondant à ce critère. Juger qu'il est plus juste de couper en deux l'enfant, c'est juger que l'action de couper l'enfant afin d'en attribuer une moitié à chaque femme contribuerait, davantage que les actions de donner l'enfant à l'une, l'autre ou aucune des deux femmes, à la réalisation d'un monde satisfaisant le critère de justice. Lorsque Sophie juge que la robe bleue est plus belle que la robe rouge, elle juge que l'action de mettre cette robe contribuerait à la réalisation d'un monde « plus beau », toutes choses égales par ailleurs, que celle de mettre la robe rouge. Mais encore faut-il que Salomon ait le désir d'un monde plus juste, ou que Sophie ait le désir d'un monde plus beau. En ce sens, lorsque plusieurs valeurs sont en jeu, l'agent peut adopter une attitude détachée à l'égard de certaines préférences : c'est le cas de Salomon lorsqu'il revient sur sa décision face aux supplications de la vraie mère.

Nous reviendrons sur la question du choix et de la hiérarchisation des normes et des préférences elles-mêmes. Nous avons ici simplement mis en évidence le critère téléologique afin de distinguer les préférences des comparaisons purement factuelles, qui, quant à elles, ne dépendent pas de l'hypothèse de la réalisation d'une fin.

[245] *Ibidem*,p 12-13.

4.2.4. Jugements de valeur comparatifs et jugements de valeur non-comparatifs

Les préférences sont des jugements de valeur, mais tous les jugements de valeur ne sont pas des préférences. La nature des jugements moraux est un champ d'investigation qui excède notre recherche. En revanche, nous pouvons apporter quelques éléments de clarification à l'aide de la distinction de structure entre jugements de valeur classificatoires et jugements de valeur comparatifs.

Grüne-Yanoff et Hansson distinguent de la manière suivante ces deux types de jugements de valeur[246]. Les jugements du premier type font référence à des concepts monadiques, ou classificatoires, tels que « bon » ou « mauvais ». Un jugement tel que : « le café est mauvais pour la santé », est un jugement classificatoire : le café est classé parmi les objets « mauvais pour la santé ». Les jugements du second type font référence à des concepts dyadiques, ou comparatifs. Ainsi, le jugement : « le café est plus mauvais pour la santé que le thé » compare ces deux boissons du point de vue de leur aptitude à favoriser une bonne santé. Parmi les concepts monadiques, Grüne-Yanoff et Hansson incluent aussi les concepts tels que « le meilleur » (*best*) et « le pire ». Cette inclusion est un peu inexacte : déterminer ce qui est le meilleur suppose un processus antérieur de comparaison. D'ailleurs, le meilleur ne se confond pas avec le bon : le meilleur des mauvais coureurs n'est pas un bon coureur, même s'il a été déclaré meilleur que tous les autres membres de la classe des mauvais coureurs. Grüne-Yanoff et Hansson écrivent alors que :

> *« Il y a des relations structurelles proches entre les valeurs monadiques et dyadiques. Il semblerait paradoxal d'affirmer à la fois que A est meilleur que B et que B est le meilleur. Il serait à peu près aussi étrange d'affirmer à la fois que A est meilleur que B et que A est mauvais tandis que B est bon. Il est généralement admis dans l'étude formelle des préférences que le prédicat de valeur monadique « le meilleur » (*best*) peut être défini en termes de prédicat de valeur dyadique. Un objet A est le meilleur d'un groupe d'objets s'il est meilleur que tous les autres objets dans ce groupe (ou alternativement si aucun autre objet de ce groupe n'est meilleur que A). A l'autre extrémité de l'échelle de valeurs, « pire » est défini de manière analogue. »*

L'argumentation confond ici des concepts tels que « le meilleur » ou « le pire » avec les concepts tels que « bon » ou « mauvais ». S'il est paradoxal de dire que A est meilleur que B, et B est le meilleur, en revanche, il n'est pas paradoxal de dire que A est meilleur que B et que

[246] *Op. cit.*

B est bon. Un pianiste peut être meilleur qu'un autre sans que ce dernier ne soit mauvais. Il faut donc distinguer le cas des concepts comme «le meilleur » et « le pire » qui peuvent être en effet définis en termes de prédicats de valeur dyadique, peut-être parce qu'en fait ils en sont *déjà*, du cas des concepts comme « bon » ou « mauvais » qui ne dépendent que de façon très lâche des prédicats de valeur dyadique. En effet, le jugement de valeur comparatif :

(a) « le coureur A est meilleur que le coureur B »

Ne contredit pas les jugements de valeur classificatoires suivants :

(b) « le coureur A et le coureur B sont bons. »

(c) « le coureur A est bon et le coureur B est mauvais. »

(d) « le coureur A et le coureur B sont mauvais ».

Il ne contredit que ce dernier jugement classificatoire :

(e) « le coureur A est mauvais et le coureur B est bon. »

Il serait donc erroné d'affirmer qu'un jugement de valeur comparatif est totalement indépendant d'un jugement de valeur classificatoire. Les jugements de valeur comparatifs doivent varier dans le même sens que les jugements de valeur classificatoires, mais ils ne rendent pas nécessairement compte de la dichotomie entre ce qui est absolument « bon » (ou juste, ou beau…) et ce qui ne l'est absolument pas. Un jugement de valeur comparatif peut donc exprimer plusieurs jugements de valeur classificatoires différents ; à ce titre, les préférences d'une personne ne nous délivrent qu'une information incomplète sur ses valeurs substantielles, notamment lorsque l'environnement de choix de cette personne – les options faisables, ou qu'elle croit faisables – est restreint. Lorsqu'une personne dit préférer la torture à la mort, il serait bizarre d'en déduire qu'elle juge que la torture est *bonne* ; en revanche, on peut en déduire qu'elle juge que la torture est la meilleure des options qui lui sont offertes, dans le cas où elle est contrainte à choisir entre les deux (et rien d'autre). Les jugements de valeur comparatifs peuvent ne pas rendre compte des jugements de valeur classificatoires de cette personne lorsqu'ils dépendent du menu, c'est-à-dire des choix offerts à l'agent[247]. Mais une information incomplète n'est pas une absence d'information.

Afin de saisir plus précisément encore la distinction entre jugement de valeur comparatif et jugement de valeur classificatoire, ainsi que les relations entre les deux, nous pouvons également définir un jugement classificatoire en termes de jugement comparatif. Une solution

[247]Cf. SEN, Amartya, *Rationality and Freedom*, The Belknap Press of Harvard University Press, 2002; traduction française par Marie-Pascale d'Iribane-Jaawane :*Rationalité et liberté en économie*, Paris, Editions Odile Jacob, 2005, chap. 4.

consiste à établir qu'un objet « bon » est meilleur que sa négation[248]. Par exemple, un bon coureur est meilleur qu'un coureur qui n'est pas bon, c'est-à-dire que la négation du bon coureur.

La définition d'un jugement de valeur classificatoire en termes de jugement de valeur comparatif pose également le problème de la pertinence de la comparaison pour des jugements de valeur de type déontologique[249]. Lorsqu'on juge qu'il est mal de tuer des êtres humains, il est difficile d'établir ensuite des gradations. Le jugement :

(f) Il est mal de tuer des êtres humains

peut sans difficulté être redéfini ainsi en termes comparatifs :

(g) Tuer des êtres humains est plus mauvais que de ne pas en tuer.

En revanche, peut-on accepter de le redéfinir ainsi ?

(h) Tuer un être humain est meilleur que d'en tuer deux.

La proposition (h) est étrange. Faut-il affirmer que, sous contrainte, nous devrions être disposés à choisir de tuer une personne plutôt que deux ? C'est une conclusion gênante. Nous dirions que nous tuons un être humain plutôt que deux « parce que nous n'avons pas le choix », et non parce que c'est « mieux ». Si, d'un point de vue logique, les jugements de valeur classificatoires peuvent être transcrits sous la forme de jugements de valeur comparatifs, en revanche, certaines valeurs monadiques ne se laisse pas utiliser dans des comparaisons, du moins dans l'usage ordinaire que nous en faisons.

En conclusion, même si, en redéfinissant les préférences comme des jugements de valeur comparatifs, nous les avons rendues aptes à intégrer bien plus de considérations évaluatives que le seul bien-être ou la seule satisfaction des désirs, il ne faudrait pas en déduire que les préférences peuvent représenter n'importe quel jugement de valeur, et que n'importe quel type de jugement de valeur peut être exprimé en termes de préférence.

[248]BROGAN, Albert, « The fondamental value universal », *Journal of Philosophy, Psychology and Scientific Method*, 16, pp 96-104.

[249] Cf. également la sous-section 2.1.2.b de notre travail.

Conclusion du Chapitre 4 : nature de la préférence et approche des capabilités

Cette analyse exploratoire de la préférence nous a permis de clarifier ce que Mongin et Dietrich appellent le « point de vue évaluatif de la préférabilité »[250], et de définir un concept de préférence qui puisse être utilisé dans l'approche des capabilités.

La conception de la préférence que nous avons développée dans le chapitre 4 s'accorde en effet avec certaines des caractéristiques de l'approche des capabilités :

1) Nous avons défendu le concept de préférence fonctionnelle, qui stipule que la relation entre la préférence et le choix peut être comprise de la manière suivante : la préférence peut remplir la fonction de disposition à choisir, mais conjointement à d'autres attitudes. Le choix n'est donc pas exclusivement déterminé par la préférence. Cette approche de la relation entre préférence met en évidence la pluralité des facteurs et des motivations qui influencent le choix, et l'importance du contexte pour les comprendre. Or, cet intérêt pour une telle pluralité est caractéristique de l'approche des capabilités.
2) La conception de la préférence que nous défendons définit celle-ci comme une opération cognitive, et plus précisément un jugement. La préférence ne peut donc pas être réduite à un comportement observé. Cela permet de penser une préférence insatisfaite. Or, le fait que les privations puissent empêcher la satisfaction d'une préférence est un des problèmes que traite l'approche des capabilités.
3) Nous avons également insisté sur le fait que la préférence est l'aboutissement d'un processus de comparaison d'états sociaux reflétant l'appréciation de ces derniers. Il est apparu que la manière dont ces états sont décrits peut influencer de façon déterminante leur évaluation. L'importance de la description des états sociaux pour une comparaison informée fait écho à l'une des motivations majeures de l'approche des capabilités, à savoir le souci de fournir une description du désavantage qui soit la plus riche et la plus pertinente possible.
4) Une théorie reposant sur les préférences, comme la théorie du choix social, est une théorie conséquentialiste. Ce conséquentialisme se retrouve dans l'approche des capabilités. A l'origine des jugements de valeur, nous trouvons une comparaison de

[250]MONGIN, Philippe et DIETRICH, Franz, « Un bilan interprétatif de la théorie de l'agrégation logique », *Revue d'économie politique*, 120 (6), pp 929-972.

conséquences, qui peut être vue comme l'un des points d'ancrage de la « tradition comparative » mise en évidence par Sen dans *l'Idée de Justice.*

Chapitre 5 : Que décrit la préférence ? Position de l'évaluateur et critères d'évaluation

Dans le chapitre précédent, nous avons dégagé une conception de la préférence cohérente avec le cadre conceptuel de la théorie du choix social et en ligne avec certaines des préoccupations majeures de l'approche des capabilités.

Toutefois, il sera difficile d'aller plus loin en ce sens si nous ne remettons pas en cause l'interprétation subjectiviste de la préférence. Cette interprétation subjectiviste semble difficilement conciliable avec l'une des ambitions principales de l'approche des capabilités, qui est de proposer un critère d'évaluation sociale qui ne soit pas biaisé par des fausses croyances, des préférences adaptatives ou des illusions objectives[251]

Les préférences disent ce qui est « meilleur » pour l'agent. Elles décrivent donc le bien de celle ou de celui à qui nous les attribuons. Mais cette affirmation n'est pas dénuée d'ambiguïtés : le « bien » est-il défini indépendamment des préférences, ou se réduit-il à leur satisfaction ? Ce bien est-il objectif ou subjectif ? Correspond-il au bien-être de l'individu ? Dans la littérature, les termes *welfare*, *well-being* et même *good* sont utilisés à tour de rôle pour désigner ce que l'on traduit en français par « bien-être ». Quelle est donc la nature de ce « bien-être » ?

Une difficulté, distincte de celle de la détermination du critère ou de la norme de « préférabilité », mais néanmoins associée à celle-ci, est que la position de celui qui décrit la préférence n'est pas conçue de la même manière selon que l'on se situe dans le paradigme propre à la théorie du choix social, issu des études des procédures de vote, et dans celui de la construction des fonctions de bien-être social, plus proche dans son esprit de l'évaluation sociale utilitariste.

Déterminer en quel sens la préférence décrit le bien des agents est une question qui, en fait, en comprend trois distinctes l'une de l'autre : la première est la question de la position de l'évaluateur ; la deuxième est celle de la nature du critère d'évaluation et de son rapport à la préférence ; la troisième est celle du sujet de l'évaluation : s'agit-il des intérêts de l'agent dont il est question, ou bien de celui de la société dans son ensemble ? Le tableau suivant recense les différentes possibilités et les associe entre elles.

[251] Cf. section 3.1.2..

	L'évaluateur et l'agent de la préférence sont la même personne	L'évaluateur est distinct de l'agent auquel la préférence est attribuée
Le bien est la satisfaction de la préférence	*L'agent exprime lui-même une préférence, dont la satisfaction réalisera sa conception du bien.*	*L'évaluateur exprime à la place de l'agent sa préférence, dont la satisfaction réalisera sa conception du bien.*
Le bien est défini indépendamment de la préférence	*L'agent exprime lui-même une préférence ; sous certaines conditions (si l'agent est rationnel, bien informé, soucieux de ses intérêts), la satisfaction de cette préférence correspondra à son bien.*	*L'évaluateur exprime à la place de l'agent ce qui est préférable pour lui, en se fondant sur une conception du bien définie indépendamment de la préférence.*

Tableau 5

Ce tableau synthétique offre une vision d'ensemble de ces différents problèmes. Parmi ces choix interprétatifs, il nous faut déterminer et défendre ceux qui reflètent une compréhension fidèle de la théorie du choix social tout en étant en adéquation avec les préoccupations et les ambitions de l'approche des capabilités.

Nous défendrons donc l'interprétation suivante du processus d'évaluation décrit par la préférence : ce processus d'évaluation est, dans la lignée de l'argumentation développée dans le Chapitre 3, caractérisé comme positionnellement objectif ; les préférences prises en compte dans l'évaluation sont des préférences soumises à des procédés de filtration permettant d'exclure celles qui posent problème pour l'approche des capabilités. Cette interprétation est en fait un double compromis. Premièrement, il s'agit d'un compromis entre l'évaluation de la situation de l'agent par l'agent, d'une part, et l'évaluation par un observateur impartial, d'autre part. Deuxièmement, il s'agit d'un compromis entre conception subjective et conception objective du bien.

Nous étudierons dans cette perspective la position de l'évaluateur par rapport à celle de celle, celui ou ceux auxquels la préférence est attribuée (5.1.), puis la question de l'objectivité du critère de « préférabilité » et son rapport au contenu de la préférence elle-même (5.2.).

5.1. La position de l'évaluateur et la rationalité de l'évaluation[252]

La préférence consiste en une évaluation par comparaison. Encore faut-il déterminer la position la plus appropriée à cet acte d'évaluation. Nous avons discuté dans la première partie l'hypothèse selon laquelle l'individu serait l'origine de fait des préférences, pour en montrer les limites. Dans ce chapitre, nous examinons non pas l'origine de fait des préférences, mais le point de vue ou la perspective à partir de laquelle celles-ci pourraient être énoncées de façon à ce que les désavantages comme les avantages des membres de la société soient énoncés de la manière la plus précise et la plus pertinente.

En économie normative, il existe deux paradigmes dominants de l'évaluation. L'un correspond aux jugements de bien-être social conceptualisés par Abram Bergson et stipule que ces jugements devraient être énoncés comme émanant d'une position d'observation, extérieure et impartiale : il s'agit du paradigme de l' « observateur impartial ». L'autre paradigme est celui de la décision collective et du vote, initié par Kenneth Arrow. Ce paradigme suppose que les préférences sont énoncées de la position même des individus : il s'agit du paradigme de la « voix » [*voice*], terme emprunté aux travaux sur le vote. Chacun de ces paradigmes présente des désavantages du point de vue de l'approche par les capabilités. Le premier tend à nier l'importance de l'activité d'agent, qui consiste à former ses propres préférences et à les défendre dans l'espace public. Le second tend à occulter les influences extérieures susceptibles de biaiser la perception qu'a l'agent de son propre avantage et de ses propres attentes. Nous examinerons successivement le paradigme de la voix (5.1.1.) et celui de l'observateur impartial (5.1.2.)

Nous proposerons ensuite de jeter les bases d'un paradigme alternatif qui s'appuie notamment sur le concept senien d' « objectivité positionnelle ». Nous y suggérons que la position de l'évaluateur peut être comprise comme une position à la fois concrète, située, et dotée des prérequis cognitifs nécessaires à une évaluation adéquate de l'avantage et des attentes (5.1.3.).

[252] Cette section a beaucoup bénéficié des échanges que j'ai eus avec Muriel Gilardone et Antoinette Baujard.

5.1.1. « Voix » et choix social

Dans cette section, nous présentons le paradigme de la « voix », selon lequel l'agent effectue lui-même l'évaluation. Il nous est apparu que, bien que ce paradigme rende justice à la qualité d'agent et à la valeur d'une démarche personnelle de formation et d'expression des préférences, il tend à ignorer le fait que la capacité de faire valoir ses préférences dépend des circonstances dans lesquelles se trouve l'agent (voire de son existence, dans le cas des choix collectifs affectant les générations futures).

On peut identifier deux arguments en faveur de l'évaluation des préférences à partir de la position de la personne concernée elle-même, que nous discuterons tour à tour. Le premier est que cette personne connaît mieux que personne ce qu'elle préfère – personne ne peut savoir comme moi où et à quel point ma chaussure me fait mal – (a) ; le second est que le fait d'exprimer soi-même ses préférences a une valeur propre : cela confère à la personne la dignité et l'autonomie du citoyen (b).

a) Connaît-on ses propres préférences mieux que personne ?

Un premier argument en faveur de l'évaluation de ce qui est meilleur pour un individu ou un groupe à un moment donné est que celui-ci a une meilleure connaissance de ce qui lui convient. Cet argument a beaucoup de force dans le cas où l'on estime que la préférence est déterminée par la perspective d'une expérience agréable. Une conception hédoniste de la préférence accueillera ainsi plus aisément l'idée que l'évaluateur et l'évalué ne font qu'un : il est en effet très difficile de vivre par procuration les expériences d'autres personnes. Ainsi, le philosophe R. Hare souligne que le principal obstacle à l'énonciation des préférences d'une personne par une autre personne est le fait que celle-ci ne peut savoir ce que c'est que d'*avoir* cette préférence.

> *« Examinons à présent notre connaissance de ce que c'est que d'être* quelqu'un d'autre *qui est en train de souffrir (par exemple parce qu'il s'est rompu le cou). Peut-il être correct de dire de moi que je sais ce que cela est pour lui (non pas juste de savoir que son cou est rompu), à moins que je n'aie moi-même une aversion égale au fait que cela m'arrive, si je devais être dans sa position avec ses préférences ? »*[253]

Hare distingue ici l'information selon laquelle un autre s'est rompu le cou, information à laquelle la victime aussi bien que d'autres personnes peuvent avoir accès, et l'expérience

[253]HARE, R. M., *Moral Thinking: its levels, methods and point*, Oxford: Oxford University Press, 1982, 5.3. Traduction libre.

(*what it is like to…*) individuelle d'avoir le cou rompu. Cette expérience dépend du degré d'aversion que chacun éprouve à l'égard de ce type de douleur. Or, Hare doute que cette expérience puisse être communiquée aussi aisément qu'une information factuelle :

> *« Supposons que j'aie dit : « Oui, je sais exactement comment tu te sens, mais cela ne me fait rien du tout si quelqu'un me fait maintenant la même chose » : ne devrais-je pas montrer à présent que je ne savais pas vraiment, ou même ne croyais pas vraiment, que ce serait comme* cela *? Mon manque de connaissance, voire de sincérité, ne serait-il pas mis en évidence si quelqu'un disais : « très bien, si cela ne te fait rien, pourquoi ne pas essayer ? » »*[254]

Nous ne pouvons donc jamais être absolument certains de savoir ce que c'est que de faire l'expérience d'une douleur ou d'un plaisir tant que nous ne faisons pas actuellement cette expérience : une préférence énoncée à la place de quelqu'un d'autre sera toujours basée sur une expérience hypothétique, et il n'est pas certain que cette expérience hypothétique corresponde à l'expérience actuelle de celui dont on énonce ainsi la préférence. Et même, en imaginant que nous ayons la capacité de nous représenter exactement l'expérience d'une autre personne dans une situation donnée, il n'est pas certain que nous aurons les *préférences* de cette personne à l'égard de ce type d'expérience. Hare souligne ainsi la distinction entre les deux propositions suivantes :

> *« (1) Je préfère maintenant avec la force F que, si j'étais dans cette situation,* x *doive arriver plutôt que non.*
>
> *(2) Si j'étais dans cette situation, je préfèrerais avec la force F que* x *doive arriver plutôt que non. »*

La préférence d'une personne dans une situation présente pour une situation future n'est pas la même que la préférence d'une personne dans une situation future pour cette situation. Les attirances ou les aversions de cette personne peuvent changer, et avec elles ses préférences. Or, si l'on ne peut pas savoir maintenant quelles seront nos préférences effectives plus tard, l'on ne peut *a fortiori* pas savoir dans notre position actuelle quelles sont les préférences d'une autre personne. Il doit y avoir au moins autant de différences entre les préférences actuelles et les préférences futures d'une même personne qu'entre les préférences de deux personnes différentes.

Cette difficulté soulevée par Hare constitue un argument contre l'évaluation de ce qui est meilleur pour un agent par un autre agent, ou, si on préfère, contre l'attribution d'une préférence à un agent par un autre agent. Mais on peut opposer à cet argument trois objections. La première est que cette difficulté n'est pas insurmontable : Hare lui-même

[254] *Ibidem.*

propose une méthode pour la contourner; la deuxième est que le raisonnement de Hare repose sur l'hypothèse selon laquelle c'est une expérience subjective singulière qui détermine la préférence : or, cette hypothèse n'est pas nécessaire au préférentialisme ; la troisième découle de la deuxième : comment prendre en compte les préférences des *absents*, de ceux qui ne peuvent pas, pas encore ou plus éprouver une expérience dans la situation présente, et qui néanmoins ont des revendications à faire valoir dans le processus de décision collective ?

On peut s'identifier aux prescriptions plutôt qu'à l'expérience de l'autre

La première objection consiste dans la proposition suivante de Hare : lorsqu'une personne s'identifie à une autre personne afin d'énoncer à sa place la préférence de cette personne, elle n'a pas besoin de s'identifier à cette personne en tant que celle-ci vit une expérience qui fait qu'elle préfère une option à une autre. Elle n'a pas besoin de vivre elle-même l'expérience de la personne à la place de laquelle elle se met pour lui attribuer des préférences. Il suffit, nous dit Hare, qu'elle s'identifie aux prescriptions qui découlent des préférences de cette personne.

> *« […] 'Je' n'est pas un mot totalement descriptif ; il est en partie prescriptif. En m'identifiant avec une personne quelconque, que ce soit actuellement ou hypothétiquement, je m'identifie avec ses prescriptions. En des termes plus simples, penser à la personne qui est sur le point d'aller chez le dentiste comme étant moi-même, c'est avoir maintenant la préférence qu'elle ne devrait pas souffrir comme je crois qu'elle va souffrir. Dans la mesure où je pense qu'il s'agira de moi-même, j'ai maintenant par anticipation la même aversion que je pense qu'elle aura. »*

Il n'est donc pas nécessaire de partager les préférences d'une autre personne, et surtout l'expérience qui fait que cette personne a ces préférences, pour reconnaître que la situation de la personne implique certaines prescriptions, requiert la mise en œuvre de certaines actions susceptibles de réaliser la préférence dont il est question.

Par conséquent, même si l'on ne connaît pas mieux ses préférences que celle de l'autre, il est tout à fait concevable de formuler des prescriptions pour l'autre. La perspective de la « voix » n'est pas nécessaire ici.

Préférences et expérience

L'objection suivante remet en cause la thèse même selon laquelle la préférence est déterminée par une expérience subjective privée connaissable de l'agent seul. Il suffit pour cela de démontrer qu'une préférence qui n'est pas déterminée par ce type d'expérience *privée* peut être prise en compte dans la procédure d'évaluation et de choix social pour qu'il soit possible dès lors de confier le soin d'évaluer cette préférence à une autre personne ou une autre entité

que l'agent lui-même. Or, les préférences étant définies comme des jugements de valeur comparatifs et non comme une satisfaction expérimentée, il nous est possible d'être parfaitement agnostique à l'égard de la définition de ce qui motive lesdites préférences.

Les préférences des absents

La troisième objection stipule que, non seulement l'expérience subjective n'est pas nécessairement constitutive de la préférence, mais aussi que, dans certains cas, elle *ne peut pas* la constituer, parce que l'on attribue une préférence à un ou plusieurs individus qui sont *absents,* qui ne peuvent pas vivre d'expériences de plaisir ou de peine. Ce cas concerne les personnes plongées dans le coma, ou même celles qui sont paralysées et ne sont pas, en partie, capables d'éprouver de la douleur ou du plaisir. Or, si le choix social et l'évaluation sociale doivent prendre en compte les préférences de tous les membres de la collectivité, alors on ne peut exclure ces personnes en raison de leur incapacité temporaire à éprouver des douleurs, des plaisirs, des aversions et des attirances, et donc à avoir des préférences déterminées par ces dernières. Le choix collectif doit également prendre en compte les préférences de personnes qui ne sont pas présentes. Ainsi, les politiques publiques en matière de dette auront des effets sur les générations futures, des personnes qui n'existent pas encore et dont cependant on peut juger qu'il convient de prendre en compte les préférences[255]. Nous devons dès lors définir la position de l'évaluateur de telle sorte qu'elle n'ait pas à coïncider systématiquement avec celle de l'agent dont on énonce les préférences.

Le problème se répète, quoiqu'en des termes légèrement différents, pour les personnes qui n'existent plus. Lorsque nous obéissons aux volontés d'une personne décédée, nous agissons comme si ses préférences passées étaient encore actuellement valables.

Le problème apparaît également dans la mesure où les préférences présentes ou actuelles d'une même personne peuvent différer de ses préférences futures. On peut donc se demander si l'agrégation des préférences ne pourrait pas intégrer les préférences futures probables d'une personne.

En conclusion, l'argument en faveur de l'énonciation des préférences par l'agent lui-même, selon lequel seul l'agent peut faire l'expérience de ce que c'est que d'avoir une telle

[255] Sur le thème lié de la justice entre générations, cf. RAWLS, John, *A Theory of Justice*, The Belknap Press of Harvard University Press, 1971; traduction française par Catherine Audard sur un texte révisé par l'auteur en 1975: *Théorie de la Justice*, Paris, Editions du Seuil, 1987, 1997. § 44.PARFIT, Derek, *Reasons and Persons*, Oxford :Clarendon Press, 1984, Part 4 ; GOSSERIES, Axel, *Penser le justice entre les générations : de l'affaire Perruche à la réforme des retraites*, Paris : Aubier, 2004.

préférence, n'est ni requis par le cadre conceptuel de la théorie du choix social, ni approprié à certains de ses contextes d'application.

b) Qualité d'agent et expression des préférences

Le second argument justifiant la coïncidence de la position de l'évaluateur et de celle de l'agent consiste à affirmer que cela permet d'éviter le paternalisme. L'idée n'est pas que le paternalisme conduit à une mauvaise prise en compte des intérêts des personnes, mais que la procédure elle-même ne satisfait pas certaines exigences normatives relatives à la liberté des personnes. En d'autres termes, il peut être requis *a priori* que les individus expriment eux-mêmes leurs propres préférences. L'idée est que le fait de participer soi-même à la décision collective et de défendre son point de vue dans l'espace public a une valeur intrinsèque ; à l'opposé, l'argument précédent ne donnait à la coïncidence de la position de l'évaluateur et de celle de l'évalué qu'une valeur instrumentale. Les préférences exprimées par les personnes concernées auraient donc plus de valeur, ou elles seules auraient de la valeur, du seul fait qu'elles sont exprimées par ces personnes.

Hilary Putnam a également mis en évidence l'importance du fait d'évaluer par soi-même les différentes propositions. Son article intitulé « On the Rationality of Preferences » discute une objection faite à sa critique de la légitimité de l'axiome de complétude des préférences[256]. L'axiome de complétude, rappelons-le, stipule que l'agent doit toujours pouvoir ordonner par préférence les options alternatives qui lui sont proposées : soit elle préfère A à B, soit elle préfère B à A, soit elle est indifférente. Mais, quoiqu'il arrive, les options, sous la forme d'états sociaux dans la théorie du choix social, sont toujours comparables et évaluables. Formellement, si nous avons deux états sociaux *x* et *y, R* étant la relation d'ordre stipulant qu'une option « est préférée ou indifférente » à une autre, l'axiome de complétude stipule que :

$$\forall x \forall y \, (xRy \vee yRx)$$

A cet axiome, Putnam oppose le contre-exemple de Thérésa ; celle-ci est tiraillée entre le désir de mener une vie pieuse et celui de jouir des plaisirs terrestres. Chacun de ces genres de vie a de la valeur en tant que tel, et Thérésa se révèle impuissante à les hiérarchiser dans un ordre de préférence. Mais un théoricien de la décision aurait peut-être une solution pour elle, avance Putnam : supposons que lui, théoricien de la décision, décide de lui attribuer un des deux genres de vie – par exemple, le genre de vie ascétique. Si Thérésa se plaint, cela montre

[256]PUTNAM, Hilary, *The Collapse of the Fact/Value dichotomy and other essays,* Harvard University Press, 2002, chapitre 5.

qu'elle préfère en fait une existence où les plaisirs de la chair auront leur part. Si elle accepte, alors le théoricien de la décision en inférera qu'elle préférait en fait cette vie de piété. En d'autres termes, d'après le théoricien de la décision, Thérésa ne se trouvait pas dans une situation de dilemme, elle n'avait simplement pas encore formé ses préférences.

L'argument que Putnam oppose à cette solution nous intéresse dans la mesure où il nous apprend que la position de l'évaluateur a son importance. En effet, Putnam avance qu'une évaluation digne de ce nom ne saurait être construite de cette manière. Ce n'est pas un théoricien de la décision ou toute autre personne qui doit achever la formation des préférences de Thérésa ; celle-ci doit procéder à l'évaluation elle-même, jusqu'à son achèvement.

> *« Cela fait partie des choix comme celui de Thérésa, que les options incluent, en tant que part essentielle de ce qui est évalué, le fait d'être sélectionnées par l'agent elle-même, par son propre libre arbitre, et non pas d'être décidées pour elle par un théoricien de la décision (ou par un bureaucrate, un psychologue, ou à pile ou face…). »*[257]

Selon Putnam, la valeur des choix de Thérésa dépend du fait qu'elle effectue elle-même ce choix.

Mais un « Challenger » anonyme objecté à Putnam que le théoricien de la décision – ou tout autre évaluateur différent de Thérésa – peut effectivement permettre à celle-ci de révéler ses préférences. D'après lui, la préférence de Thérésa provient soit d'une source « interne », soit de facteurs externes. Dans le premier cas, cette source détermine déjà ce que cette préférence sera, Theresa n'étant simplement pas au fait de celle-ci[258]. Dans le second cas, des facteurs environnementaux connus de l'évaluateur (le théoricien de la décision) déterminent la décision de Theresa, sans que celle-ci en ait à nouveau conscience. En d'autres termes, le « Challenger » juge que l'autonomie du choix elle-même n'est pas rationnelle. Ce qui est rationnel pour le Challenger, c'est le contenu même de l'évaluation, c'est-à-dire l'ordre des préférences. Peu importe le processus par lequel il est *révélé*. Au contraire, Putnam affirme que cela ne suffit pas : les préférences sont fondées sur des raisons ; lorsque Theresa tente de comparer les deux genres de vie qui s'offrent à elle, elle compare les raisons qu'elle aurait de choisir l'un ou l'autre. Ces raisons lui sont propres. Ce ne sont pas de simples rationalisations après coup d'une préférence qui serait « déjà là »[259].

[257] *Ibidem.* Traduction libre.

[258] Il serait intéressant de mettre cette solution en perspective avec la conception frégéenne du jugement comme assertion.

[259] On peut par ailleurs se demander « depuis quand » les préférences en question seraient déjà là. La conception du « Challenger » néglige tout à fait le fait que les préférences ne peuvent qu'être formées au cours de l'existence de l'individu ; il semble en effet étrange de penser que les préférences puissent même préexister à l'individu.

On doit souligner que le cadre conceptuel de la théorie du choix social permet de rendre compte de cette différence : comme Putnam le montre, on peut substituer des ordres de préférence incomplets aux ordres complets pour représenter la perplexité de Theresa. Et, suivant une proposition de Sen, nous pourrions intégrer à la description même des états sociaux le fait que ceux-ci sont choisis par l'agent, pour des raisons qui lui sont propres. Enfin, il est possible d'affiner la position de l'évaluateur, afin de lui permettre de se représenter non seulement le résultat du choix de Theresa, mais également le processus délibératif qui a généré ce résultat.

Le souci de l'autonomie est louable, mais il ne peut être respecté dans certains cas.

L'objection selon laquelle les préférences de personnes *absentes* devraient aussi être prises en compte dans le choix collectif affecte également cet argument en faveur du paradigme de la « voix ». Mais on peut y ajouter que l'idéal selon lequel exprimer soi-même ses préférences aurait une valeur intrinsèque peut avoir pour effet d'exclure tous ceux qui ne sont pas capables de faire valoir ces préférences. Nous avons évoqué le cas de ceux qui sont exclus de la communication humaine parce qu'ils sont dans le coma ; mais les handicaps qui privent certains de la possibilité d'exprimer leurs préférences sont de natures et de degrés divers. La capacité de faire valoir par soi-même ses préférences dans la procédure d'évaluation et de choix collectifs ne doit donc pas être présupposée *a priori.* Elle peut en revanche faire partie des objectifs poursuivis par les actions mises en place suite à la décision collective.

Néanmoins, la discussion sur l'autonomie a mis en évidence l'importance des *raisons* de préférer et de choisir, importance au moins égale à celle du contenu de l'évaluation proprement dite. La rationalité de la préférence dépend de la rationalité du processus de délibération qui détermine l'évaluation.

5.1.2. L'observateur omniscient et impartial

Cette section discute la portée et les limites du paradigme de l'observateur omniscient et impartial pour l'évaluation de l'avantage et des attentes des membres de la société. Ce type de position a été élaboré avec des contraintes diverses susceptibles de garantir l'impartialité des évaluations.

Nous analyserons ici la théorie de l'observateur idéal et impartial et la position originelle rawlsienne. Ces positions sont également analysées par R. Brandt (BRANDT, Richard, *A Theory of the Good and the Right,* Oxford : Oxford University Press, 1979, ch. XII), dans la

mesure où elles sont autant de réponses à la question de savoir quelles actions seraient permises ou prescrites par un code moral qu'une personne pleinement rationnelle défendrait, dans une société où elle est supposée vivre tout le temps de son existence. Notre question consiste à déterminer si l'on peut concevoir le point de vue d'une personne capable d'évaluer au mieux ce qui serait le plus rationnel pour chacun des agents de préférer. Il est important de rappeler que ce qui est le plus rationnel de préférer pour les agents ne correspond pas nécessairement *a priori* à leurs intérêt privé : comme le souligne Arrow, l'évaluation sociale est concernée aussi bien par les goûts que par les valeurs des personnes, et ces dernières jouent généralement un rôle important dans les décisions publiques.

Cette analyse met en évidence les limites d'un recours à un tel dispositif.

a) L'observateur idéal

Le point de vue de l'observateur idéal comprend des déterminations cognitives et motivationnelles[260]. Les déterminations cognitives consistent généralement dans le fait de disposer de toutes les informations nécessaires à l'évaluation : idéalement, l'observateur idéal est omniscient. Les déterminations motivationnelles supposent le désintéressement de cet observateur : ainsi, tout agent disposant de toutes les informations nécessaires à l'évaluation de la situation d'autrui, et capable de faire abstraction de ses intérêts propres, peut se trouver dans la position de l'observateur idéal.

Les déterminations cognitives permettent à l'observateur idéal de définir au mieux les préférences que les individus devraient rationnellement avoir. Si ces préférences devaient être déterminées par une conception objective du bien, qu'elles révèleraient, alors l'observateur idéal aurait bien évidemment une position privilégiée pour les identifier. Faire énoncer les préférences par cet observateur au lieu de laisser ce soin aux agents réels permettrait d'éviter les erreurs à ce sujet. L'observateur aurait à la fois connaissance de la fin que devraient poursuivre les individus (le bien humain ou la réalisation de soi, par exemple), et de la manière dont chaque configuration sociale, économique et politique peut contribuer à l'accomplissement de cette fin. Et même si les préférences des agents devaient être déterminées par leur propre conception du bien-être, par la recherche exclusive de leur satisfaction subjective, alors l'observateur omniscient jouirait toujours du privilège de connaître ces conceptions ainsi que l'aptitude de chacun des états sociaux à les réaliser. Enfin,

[260]Cf. FIRTH, Roderick, « Ethical absolutism and the ideal observer », *Philosophy and Phenomenological Research* 12, 1952.

quelle que soit la valeur qui détermine les préférences des individus, le fait de définir la position d'évaluation indépendamment de l'existence des individus permet d'attribuer des préférences à ceux qui ne peuvent d'eux-mêmes les faire valoir dans le processus du choix collectif, soit parce qu'ils n'ont pas accès aux moyens d'expression de ces préférences, soit parce qu'ils ne sont tout simplement pas présents.

Les déterminations motivationnelles mettent en évidence la fonction d'arbitrage de l'observateur idéal, impartial et désintéressé par définition. Ainsi, Harsanyi conceptualise cette fonction en distinguant les « préférences personnelles » des « préférences morales ». Tandis que les deux sont endossées par les individus dans le cadre arrovien du choix social, Harsanyi distingue les positions à partir desquelles ces deux types de préférences sont énoncés.

> *« Les préférences morales diffèrent des simples préférences personnelles du fait qu'elles reposent sur un critère impersonnel et impartial. Supposons à présent qu'un individu* j *choisisse entre deux situations dans une complète* ignorance*, ou sans ne tenir aucunement compte de ce que serait sa position personnelle dans chacune des situations. Plus spécifiquement, supposons qu'il fasse son choix en postulant qu'il aurait exactement la* même probabilité *d'occuper n'importe laquelle des positions sociales disponibles et, de fait, d'endosser l'identité personnelle de n'importe quel individu particulier dans la société. Il est clair que, avec cet arrangement, son choix satisferait nécessairement les exigences d'impersonnalité et d'impartialité au plus haut degré possible et, par conséquent, serait qualifié de déclaration de préférence morale, c'est-à-dire de jugement de valeur moral. »*[261]

En distinguant la position de l'évaluateur de celle de l'agent, il devient possible d'abstraire certaines caractéristiques de l'agent, afin de rendre l'évaluation de la préférence valable pour plusieurs agents en même temps.

Néanmoins, ces déterminations motivationnelles peuvent sembler insuffisantes pour garantir que l'observateur omniscient et impartial veuille effectivement la satisfaction des préférences des membres de la société. Etre désintéressé signifie que l'on n'a pas d'intérêt ; cela ne signifie pas que l'on soit intéressé par le bien d'autrui. Il peut donc être souhaitable d'enrichir la description des motivations à l'œuvre dans le processus d'évaluation.

L'agrégation des préférences que l'évaluateur social omniscient et impartial attribue aux membres de la société permet de produire un jugement de bien-être social. Il suffit dès lors de postuler qu'un tel jugement revêt un caractère d'obligation morale : dire qu'un état social

[261]HARSANYI, John, *Essays in Ethics, Social Behaviour and Scientific Explanation*, Dordreicht, Reidel, 1976, chap. V.

donné est préférable à tous les autres états sociaux faisables, étant données les préférences des membres de la société, équivaudrait à affirmer qu'il faut réaliser cet état social. L'évaluation sociale basée sur l'agrégation des préférences d'un profil donné à partir d'une position d'évaluation omnisciente et impartiale se transforme alors en une théorie normative : l'état social qui doit venir à l'existence est celui qui correspond à ce jugement de bien-être collectif.

Remarquons qu'un individu dont les préférences personnelles coïncideraient par accident avec celles de l'évaluateur social se retrouverait de fait dans la position du dictateur arrovien : il s'agit du paradoxe du « dictateur démocrate », qui apparaît uniquement lorsque les conditions du théorème d'impossibilité d'Arrow sont appliquées à une approche à profil simple ; ce paradoxe disparaît aussitôt que l'on passe d'une approche à profil simple à une axiomatique valable pour tous les profils de préférences possibles.

> *« Dans le modèle d'Arrow, un dictateur est une personne dont les préférences prévalent quelles que soient les préférences des autres personnes. Il n'est alors pas difficile de reconnaître que l'existence que l'existence d'un dictateur est indésirable. Dans le modèle à profil simple, un dictateur est une personne dont les préférences individuelles coïncident avec les préférences sociales dans le seul et unique profil considéré. Il n'y a rien de nécessairement problématique avec cela ; le processus de décision peut être parfaitement démocratique, et une seule personne s'est simplement retrouvée du côté gagnant pour toutes les questions. »*[262]

Cet élément d'analyse mis en évidence par Aanund Hylland montre que la signification de la position de l'évaluateur n'est pas la même selon que l'on adopte une approche à profils simples ou à profils multiples. Dans les approches à profil simple, où l'on dérive un jugement de bien-être social d'un profil de préférences *donné*, l'évaluateur n'est – artificiellement – doté du désir de réaliser l'état social jugé bon en raison de ces préférences que sur la base de ce profil de préférences. Ce désir de réaliser l'état social jugé bon n'est en fait défini que négativement : il s'agit du désir de ne pas privilégier la préférence d'un groupe ou d'un individu au détriment de celle de la collectivité dans son ensemble. Cela permet de comprendre que l'évaluateur social n'est à proprement parler caractérisé que par sa position, et non par une motivation substantielle que l'on pourrait attribuer à un agent individuel ou collectif. Il suffit que cette position équivaille à une condition nécessaire et suffisante d'énonciation d'un jugement de bien-être collectif qui serait bien fonction d'un profil de préférences des membres de la société, pour que n'importe quelle personne ou groupe

[262]HYLLAND, Aanund, « Purpose and significance of social choice theory », *in* ELSTER, Jon and HYLLAND, Aanund (eds), *Foundations of social choice theory*, Cambridge University Press, 1986, p 51 n. 16.

susceptibles de préférer – y compris par accident - l'état social qui correspondrait à la maximisation du bien-être collectif puisse l'occuper.

b) La position originelle

Une autre possibilité consiste à énoncer la préférence des individus à partir d'une position déterminée de telle sorte que ces préférences soient filtrées, afin d'être à la fois plus rationnelles et plus appropriées à la coopération sociale. La position originelle que Rawls a définie dans sa *Théorie de la Justice* peut être comprise dans cette perspective, bien qu'elle soit conçue par Rawls comme un élément-clef de l'argumentation en faveur .des principes de justice qu'il défend.

Théorie de la justice et théorie du choix rationnel

La position originelle justifie la sélection des biens premiers qui permettent d'évaluer l'avantage dans la théorie de la justice rawlsienne. Elle garantit en effet l'impartialité et la rationalité de l'évaluation de la contribution de ces biens premiers à l'avantage individuel. Pour ce faire, Rawls utilise de fait les préférences dans la description du choix des principes de justice à partir de la position originelle. A ce titre, un rapprochement intéressant avec la théorie du choix social mérite d'être fait, comme le suggère cet extrait :

> *« Une condition supplémentaire pour une conception du juste est qu'elle doit imposer une relation d'ordre à des revendications en conflit. Cette exigence naît directement du rôle des principes qui doivent arbitrer des demandes concurrentes. Mais il est difficile de décider ce qui doit compter comme relation d'ordre. Il est clairement souhaitable qu'une conception de la justice soit complète, c'est-à-dire capable de hiérarchiser toutes les revendications qui peuvent surgir (ou qui risquent de le faire en pratique). Et la relation d'ordre devrait être en général transitive ; si, par exemple, une première organisation de la structure sociale est classée comme plus juste qu'une seconde, et la seconde comme plus juste qu'une troisième, alors la première devrait être plus juste que la troisième. Ces conditions formelles sont tout à fait naturelles, quoique pas toujours faciles à remplir[263].»*[264]

Les principes de justice choisis dans la position originelle devraient, d'après cet extrait, déterminer une relation d'ordre, c'est-à-dire une évaluation qui a la même structure que la préférence sociale de la théorie du choix social. Rawls attribue d'ailleurs à cette relation d'ordre les propriétés qu'Arrow confère à la préférence sociale : complétude et transitivité. En revanche, ce sont des « conceptions du juste » et non des «états sociaux » qui sont ainsi

[263] « Pour une étude des relations d'ordre et de préférence, voir A.K. Sen, *Collective Choice and Social Welfare, op. cit.* , chap. I et I*, et K.J. Arrow, *Social Choice and Individual Values* (New York, John Wiley and Sons, Inc., 1963), chap. II (trad. Française : *Choix collectifs et Préférences individuelles*, Paris, Calmann-Lévy, 1975). » (note de J. Rawls, complétée par C. Audard)

[264]RAWLS, John, *op. cit.*, § 23.

évalués. Autrement dit, le classement de « préférence » sociale décrit par Rawls est déterminé par la norme du juste, et la manière dont il dépend des préférences des agents appelés à produire cette évaluation sera déterminée par cette norme du juste. De fait, c'est en énonçant les préférences des agents à partir de la position originelle, et non pas de la position réelle qu'ils occupent dans la société, que Rawls peut intégrer cette relation d'ordre à une réflexion sur la justice. Le classement des conceptions du juste en fonction des principes de justice diffère du choix social en ce qu'il est arbitré à partir d'une position d'évaluation caractérisée de manière appropriée, à savoir la position originelle.

Il peut sembler curieux à première vue d'intégrer la théorie rawlsienne de la justice à une réflexion sur l'agrégation des préférences, cadre conceptuel plus proche de l'utilitarisme contre lequel la philosophie politique de Rawls s'est construite. Quelques précisions s'imposent donc.

De son propre aveu, la pensée de Rawls a évolué pendant les années quatre-vingt : il s'agit du fameux « tournant kantien »[265], lors duquel Rawls a abandonné l'idée selon laquelle la théorie de la justice est une partie de la théorie du choix rationnel[266]. Après le « tournant kantien », la position originelle relie une conception morale de la personne, analogue à la conception kantienne, aux principes de justice : elle définit la manière dont des personnes morales choisiraient les principes qui régissent une société juste. Cette conception de la personne permet de mettre en évidence ce qui, chez ceux qui vivent dans les sociétés démocratiques contemporaines, permettrait aux principes de justice d'être choisis. Elle est donc à la fois postulée par la théorie de la justice et nécessaire à la constitution de cette dernière. Mais, au moment où Rawls publie sa *Théorie de la Justice*, la position originelle n'est occupée que par des « personnes rationnelles », sans que la question ne soit traitée avec plus de précision. C'est la position originelle qui fait que les personnes choisissent les principes de justice, et non le fait que cette position soit occupée par des personnes morales :

> *« Du point de vue de la position originelle, il est rationnel pour les partenaires de supposer qu'ils veulent une part [de biens premiers sociaux] plus large, puisque, de toute façon, ils ne sont pas obligés de l'accepter s'ils ne le souhaitent pas. Ainsi, même si les partenaires sont privés d'information sur leurs buts particuliers, ils en ont suffisamment pour hiérarchiser les solutions. Ils savent que, d'une manière générale, ils doivent essayer de protéger leurs libertés, d'élargir leurs possibilités et les*

[265]RAWLS, John, « Kantian Constructivism in Moral Theory », *The Journal of Philosophy*, vol. 77, n° 9, 1980, p 515-572 ; traduction Catherine Audard, « Le constructivisme kantien dans la théorie morale », *in* RAWLS, John, *Justice et démocratie*, Seuil, « La couleur des idées », 1993.

[266] RAWLS, John, *A Theory of Justice*, The Belknap Press of Harvard University Press, 1971; traduction française par Catherine Audard du texte révisé par l'auteur en 1975: *Théorie de la Justice*, Paris, Editions du Seuil, 1987, 1997. § 3.

moyens de favoriser leurs objectifs quels qu'ils soient ; guidés par la théorie du bien et les faits généraux de la psychologie morale, leurs réflexions ne sont pas des devinettes. Elles peuvent conduire à une décision rationnelle, au sens ordinaire. »[267]

Pour Rawls,

« l'idéal serait de montrer que [la reconnaissance des principes de justice] est le seul choix compatible avec la description complète de la position originelle. »[268]

Mais il reconnaît qu'il ne peut strictement déduire les principes de justice de la position originelle, car il doit avoir recours à certaines hypothèses psychologiques substantielles. Donc, dans la *Théorie de la Justice*, les partenaires de la position originelle sont guidés à la fois par une information sélective, leur rationalité et le fait qu'ils savent qu'ils poursuivront la réalisation d'un plan de vie ; les « faits généraux de la psychologie morale » comprennent l'amour des enfants pour leurs parents, l'affection pour les membres d'un système de coopération auquel on appartient, le sens de la justice, et le sentiment de culpabilité qui accompagne les actions allant à l'encontre de ces trois tendances. Ces « faits généraux » relèvent davantage d'une psychologie empirique que d'une conception de la personne morale. La genèse de la réflexion qui a abouti à la *Théorie de la Justice* met par ailleurs en évidence le rôle important des connaissances que Rawls a acquises en économie et en théorie du choix rationnel dans la construction du modèle de la position originelle. L'influence des économistes Vickrey et Harsanyi notamment a été déterminante[269].

[267] *Ibidem,* § 25.

[268] *Ibidem,* § 20.

[269] La communication faite par Rima Hawi à ce sujet a été très éclairante sur ce point : HAWI, Rima, « Rawls et l'économie : les leçons des archives de Harvard », *Communication dans le cadre du Colloque « Philosophie Economique »,* Sciences Po Lille, 21-22 juin 2012.

Ayant levé les éventuels doutes sur l'association du concept de position originelle au cadre conceptuel de la théorie du choix social, il nous faut à présent répondre à deux questions : peut-on utiliser la position originelle pour évaluer les préférences des membres de la société qui devront par la suite constituer la matière de l'évaluation sociale et du choix collectif ? Et, si cela est possible, est-ce un procédé souhaitable ? Quels avantages en retirerait-on ?

La position originelle rawlsienne est en premier lieu caractérisée par le voile d'ignorance. Celui-ci garantit le caractère équitable (*fair*) de la procédure de choix des principes de justice. Ce voile dissimule la manière dont chaque structure de base affectera leur situation à chacun des agents chargés de choisir les principes de justices : en substance, il garantit leur désintéressement. Néanmoins, les partenaires ne peuvent pas non plus ne *rien savoir.* Ils ignorent certes leur position sociale, leurs atouts naturels, leur conception du bien, les « traits particuliers de [leur] psychologie comme l' « aversion pour le risque », la « propension à l'optimisme ou au pessimisme ». Ils ignorent également les caractéristiques particulières de leur société, qu'elles soient économiques, politiques, culturelles ; enfin, ils ne savent pas à quelle génération ils appartiendront[270] [271]. En revanche, ils disposent des connaissances générales requises pour choisir en conscience de cause des principes de justice appropriés : « les affaires politiques et les principes de la théorie économique », « la base de l'organisation sociale et les lois de la psychologie humaine »[272]. La position originelle efface les différences interpersonnelles et rend les différents partenaires sensibles à la même argumentation. Donc, tout comme dans le cas de l'observateur idéal, la position est définie indépendamment des individus particuliers concernés par les décisions qui y sont prises. Le seul fait d'occuper la position originelle détermine des êtres humains relativement similaires – puisque caractérisés par des faits psychologiques généraux – à évaluer les différentes solutions de la même manière.

Nous avons montré que le fait que l'évaluation sociale soit effectuée à partir d'une position différente de celle des agents concernés ne nous interdit pas de parler de choix social. Mais on pourrait se demander si le maximin, et *a fortiori* le leximin n'appartiennent pas à l'extension du concept de choix social, tel que l'a construit la théorie du choix social. Si la pratique des

[270] *Op. cit.*, § 24.

[271] Dans une première version, toutes les générations sont représentées. Par la suite, confronté aux problèmes de la non-existence (l'existence des générations future dépend des actions de la génération présente) et de la non-identité (l'identité des générations futures dépend des actions de la génération présente), problèmes soulevés par Derek Parfit, Rawls a proposé que les partenaires de la position originelle ne sachent pas à quelle génération ils appartiennent, mais sachent appartenir à la même génération.

[272] *Ibidem.*

théoriciens du choix social définit leur compréhension du choix social, alors le maximin et le leximin font partie des axiomes utilisés[273].

Le choix effectué à partir de la position originelle est un « choix social », s'il s'agit d'une espèce d'agrégation des préférences. Il est effectivement, en partie, question de préférences, à savoir les préférences des agents placés dans cette position particulière. Ces agents décident de principes de liberté et d'égalité équitable des chances, mais également d'un principe de distribution des biens qui accorde une priorité absolue à l'avantage des plus mal lotis, le principe de différence. Ce critère de distribution des avantages est le maximin. Peut-on dire que le maximin résulte de l'agrégation des préférences dans la position originelle ?

Si le problème posé dans la position originelle ne concernait que l'indexation de biens premiers de valeur et leur allocation entre les différents individus en fonction de leur position sociale, alors on pourrait concevoir les choix effectués dans ladite position comme des variantes du choix social. Des développements formels ont d'ailleurs été faits dans ce sens.

Mais le problème de Rawls est plus fondamental. Rawls souhaite, en un sens, répondre à la question : qu'est-ce qu'une société juste ?, afin de déterminer comment y parvenir. Précisément, son ambition est de déterminer les principes de justice qui devraient régir une structure de base « juste ». La structure de base ne se limite pas à la description de l'allocation des avantages. Elle inclut en fait toutes les institutions, évaluées selon leurs effets sur les existences concrètes des personnes. En intégrant la position originelle aux dispositifs d'évaluation de l'avantage individuel en économie normative, nous n'avons donc offert qu'une vision restreinte du projet rawlsien. Ce projet est plus ambitieux politiquement que celui de la théorie du choix social dans la mesure où il vise à changer la structure institutionnelle même. En revanche, l'ambition descriptive de la théorie du choix social, et particulièrement de ses derniers développements, est plus importante puisqu'elle ne se résout pas à appréhender l'avantage au travers de l'indice des biens premiers.

Il nous faut à présent élaborer un paradigme d'évaluation susceptible de satisfaire cette ambition descriptive.

[273] Dans son manuel d'introduction à la théorie du choix social, Wulf Gaertner parle à plusieurs reprises du maximin et du leximin : GAERTNER, Wulf, *A Primer in Social Choice Theory*, Oxford : Oxford University Press, 2006, pp 15, 62, 81, 110, 112, 115-117, 120, 122, 151, 153-155, 157.

5.1.3. Les évaluations positionnelles : de multiples possibilités

Le paradigme de la « voix » préserve la perspective évaluative de l'agent, mais tend à négliger les conditions concrètes d'une telle perspective et à considérer comme acquise la capacité d'énoncer des préférences représentatives de son avantage propre. Le paradigme de l'observateur impartial et omniscient fait droit à une évaluation informée, mais tend à ignorer les conditions pratiques de mise en œuvre d'une évaluation de cet ordre. Dans cette section, nous proposons un mode d'évaluation alternatif, qui repose en grande partie sur le concept senien d'objectivité positionnelle.

Amartya Sen a défini le concept d'objectivité positionnelle afin de caractériser à la fois les observations et les jugements individuels ou collectifs méritant d'être contextualisés. Sen caractérise chaque individu par sa position[274] [275]. En effet, nous avons montré (Ière Partie, 3.1.4. de ce travail) qu'il était plus cohérent avec la pensée de Sen d'attribuer les évaluations qui constituent la matière du choix collectif à des positions qu'à des individus. A présent, nous franchissons une étape supplémentaire en argumentant en faveur de l'hypothèse selon laquelle nous pouvons attribuer des *préférences* à des positions, et même que cela est plus souhaitable que de les attribuer à des individus, d'une part, et que cette attribution peut être effectuée à partir d'une autre position, d'autre part. Les positions en question sont des positions réelles et non pas hypothétiques, comme c'était le cas avec l'observateur idéal ou la position originelle. La nature du raisonnement évaluatif s'en trouve dès lors modifiée.

Rappelons qu'une position est définie par Sen comme un ensemble de paramètres qui influencent ou déterminent[276] conjointement les observations, croyances et décisions des individus qui l'occupent. Une position n'est pas qu'une location. Ainsi, le fait d'avoir une mauvaise vue est un paramètre positionnel, sans pour autant correspondre à une location déterminée.[277].

---

[274] SEN, Amartya, « Positional Objectivity », *Philosophy & Public Affairs*, Vol. 22, No. 2, Spring, 1993, pp. 126-145 ; traduction française par Marie-Pascale d'Iribane-Jaawane : *Rationalité et liberté en économie*, Paris, Editions Odile Jacob, 2005, chap. 15.

[275] Nous devons beaucoup à ce que nous avons entendu de GILARDONE, Muriel et BAUJARD, Antoinette, « A possible social choice : democracy and positional judgments », communication dans le cadre du colloque *Philosophie Economique*, Sciences Po Lille, 21-22 juin 2012.

[276] La question du déterminisme n'est pas traitée directement par Sen.

[277] Les paramètres positionnels chevauchent ainsi allègrement certaines dichotomies, telles que la dichotomie individu-société ou la dichotomie nature-culture.

Nous avons vu que le fait de passer d'une analyse des caractéristiques individuelles à une analyse des paramètres positionnels nous permet de concevoir le fait que différents individus puissent occuper une même position et partager les mêmes observations, croyances et décisions, d'une part, et qu'un même individu puisse occuper simultanément ou successivement différentes positions, d'autre part.
Sen parle d'*observations,* de *croyances*, de *jugements* et de *décisions* positionnellement objectives, mais jamais de *préférences*. Toutefois, cela ne doit pas nous inquiéter outre mesure, car ce sont précisément ces observations, croyances, jugements, qui déterminent les préférences des agents, que les préférences ne viennent pas de « nulle part », mais trouvent leurs sources dans des motifs susceptibles d'être décrits en termes de paramètres positionnels.

Nous expliciterons l'attribution de préférences à des positions (a). Nous proposerons ensuite notre propre reconstruction de l'évaluation positionnelle à partir de la contribution de Sen. Tout d'abord, l'évaluateur peut être l'agent lui-même ou non ; le critère déterminant n'est pas son identité, mais son aptitude à formuler au mieux l'avantage et les attentes de l'agent concerné, aptitude qui dépend de la connaissance des paramètres positionnels pertinents pour l'évaluation (b). Nous proposons ensuite un « test positionnel » permettant de vérifier si les évaluations effectuées par une personne autre que l'agent correspondent effectivement aux finalités de l'agent lui-même (c). S'en suit un bref bilan (d).

a) L'attribution des préférences aux positions

Les paragraphes qui suivent montrent comment il est possible d'attribuer une préférence à une position, au sens où Sen entend ce terme : la position est l'ensemble des paramètres susceptibles d'influencer le jugement de l'individu. Les positions dont il est question ici sont caractérisées en fonction des objectifs de l'évaluation sociale. Autrement dit, sera pris en compte tout ce qui est pertinent pour déterminer dans quelle mesure l'individu est désavantagé[278].
Les préférences sont généralement attribuées aux individus, encore que le formalisme laisse la possibilité de les attribuer à n'importe quel type d'agent, individuel ou collectif. Ici, nous souhaitons montrer que les préférences peuvent également être attribuées aux positions.

[278] Cette sous-section complète ainsi le raisonnement élaboré dans la section 3.1.4.. Cependant, la problématique n'est pas tout à fait la même. Dans la section 3.1.4., nous souhaitions montrer que, de fait, ce sont des caractéristiques positionnelles qui sont à l'origine des jugements de valeur recensés dans le choix social. Notre perspective était celle du diagnostic, et les préférences étaient des descriptions potentiellement biaisées du désavantage. Dans cette sous-section, nous envisageons de mettre en place une conception de l'évaluation telle qu'elle devrait être. Cette perspective se veut plus constructive, et les préférences y sont comprises comme les revendications que devrait formuler un individu, étant donnée la position qu'il occupe. Néanmoins, il ressort assez clairement de ces deux analyses que diagnostic et construction sont corrélés.

Rappelons dans un premier temps la représentation formelle courante des préférences dans la théorie du choix social. Afin de simplifier l'analyse, nous nous limiterons au cas des préférences strictes, mais notre argumentation fonctionne également dans le cas des préférences faibles et des relations d'indifférence.

- Soit $X = \{x, y, z\}$ l'ensemble d'états sociaux que le jugement de préférence doit classer. P est une relation d'ordre binaire de préférence stricte énoncée sur X. Cette relation est anti-symétrique ; elle est réflexive (x ne peut être préféré à lui-même) ; elle est transitive (si x est préféré à y et y à z, alors x est préféré à z ; il n'est pas nécessaire de la définir comme complète. P_i est la relation d'ordre attribuée à l'individu I sur l'ensemble$X$. Donc, la proposition:

  $$x \, P_i \, y$$

  signifie que i préfère x à y, ou juge x meilleur que y. Ou tout au moins attribuons-nous cette préférence à l'individu i en vertu de ce que nous pouvons observer d'elle.

Supposons que i représente Sophie, une aveugle. x est l'état social dans lequel Sophie possède un chien guide d'aveugle, et y celui où elle doit se déplacer sans. x et y ne diffèrent qu'à cet égard. Si la préférence de Sophie est prise en compte, alors il faut faire en sorte qu'elle dispose du chien guide d'aveugle.

Si nous attribuons la préférence à Sophie sans rappeler qu'elle est aveugle, cette préférence peut passer pour arbitraire : est-ce que Sophie aime simplement les chiens ? Se sent-elle plus en sécurité avec ? Veut-elle faire croire qu'elle est aveugle, par plaisanterie ? Nous ne savons pas. Lorsque nous attribuons une préférence à un individu substantiel, nous ne disposons que d'un regard extérieur sur cette préférence et sur les motifs de celle-ci, ce qui peut faire apparaître la préférence comme purement arbitraire. Cela ne nous permet pas de distinguer une préférence susceptible d'avoir une certaine légitimité dans le débat public d'une pour laquelle ce n'est pas le cas.

Au contraire, si la préférence est reliée à un paramètre positionnel, elle peut être plus aisément justifiée. Ainsi, la proposition $x \, P_i \, y$ pourrait être comprise de ces deux manières:

(j) "Etant aveugle, Sophie devrait préférer un état social dans lequel elle dispose d'un chien d'aveugle à un où ce n'est pas le cas."

(k) "Comme Sophie est aveugle, un état social dans lequel elle dispose d'un chien d'aveugle est meilleur pour elle qu'un état social où ce n'est pas le cas. »

La proposition (j) fait état de la propre déclaration de préférence de Sophie, tandis que (k) pourrait également être énoncée par un évaluateur indépendant. (k) se réfère explicitement au critère du *meilleur*, qui n'apparaît pas clairement dans (j). Mais la caractéristique commune de (j) et de (k) est que la préférence pour un état social avec un chien d'aveugle est liée à un des paramètres de la position qu'occupe Sophie, à savoir le fait qu'elle est aveugle. Mentionner le paramètre justifie la préférence : sans ce paramètre, la préférence pour le chien d'aveugle ne serait pas. Ce que (j) et (k) signifient, c'est que, *parce que* Sophie est aveugle, elle préfère avoir un chien plutôt que de ne pas en avoir.

Cela signifie que l'on peut attribuer cette préférence à tous ceux dont la position est caractérisée par ce même paramètre. En d'autres termes, la raison pour préférer une un objet à un autre correspond à un fait descriptible en termes indépendants des désirs et des volitions de l'individu concerné. Ce n'est pas parce que Sophie *croit* ou *sait* qu'elle est aveugle qu'elle préfère avoir un chien d'aveugle ; c'est en raison du *fait* qu'elle est aveugle qu'elle a, ou devrait avoir cette préférence.

b) La position de l'évaluateur

Nous savons à présent qu'il n'est pas absurde d'attribuer des préférences à des positions plutôt qu'à des individus, et que cela peut même s'avérer indispensable pour l'évaluation de l'avantage. Il nous faut donc ensuite déterminer qui est le plus à même d'identifier les paramètres positionnels à prendre en compte pour attribuer une préférence ou une revendication à une personne ou un groupe de personnes.

Est-ce l'agent individuel concerné ? Le fait que les préférences dépendent de positions et non d'une source proprement individuelle n'interdit pas l'énonciation de ces préférences par l'individu concerné lui-même. En fait, *l'aptitude à exprimer par soi-même des préférences positionnelles peut être considérée comme un paramètre positionnel en soi* : cette aptitude peut dépendre de la capacité physique d'énoncer une préférence, de la présence sur le lieu d'énonciation des préférences, de la reconnaissance institutionnelle d'un droit à l'expression et à la parole, ou encore de ce que Rawls nomme les bases sociales du respect de soi-même. Mais ce n'est parfois pas possible, soit parce que l'évaluation de l'individu concerné se trouve être biaisée[279], soit parce qu'il n'est de fait pas présent ou capable de faire valoir ses attentes[280]. Il faut donc songer à définir une position d'évaluation des préférences de l'agent concerné différente de la position qu'occupe ce même agent. Ainsi, si Sophie ne sait pas

[279] Cf. section 3.2.1..
[280] Cf. section 5.1.1.

qu'elle peut obtenir un chien d'aveugle, ou si elle ne perçoit pas cette revendication comme légitime, une personne occupant une autre position peut affirmer en son nom qu'il est préférable pour elle de profiter de ce type d'aide.

c) Le test positionnel

Des difficultés surgissent cependant lorsque la position de l'évaluateur ne coïncide pas avec celle de l'évalué. En effet, comment l'évaluateur peut-il légitimement attribuer une préférence à une position, y compris en dépit des propres déclarations de l'individu, et comment peut-il dire en lieu et place de ce dernier que tel état social est préférable à tel autre ? Une première difficulté est que l'évaluateur peut se tromper, non pas sur les caractéristiques de la position de l'agent, mais sur ce que ce dernier valorise. Une seconde difficulté est que l'évaluateur adopte de ce fait une posture paternaliste. Cette posture peut être justifiée lorsqu'il est urgent de remédier à une privation grave. Mais, à long terme, elle risque de miner l'égalité des membres de la société en attribuant à certains d'eux le privilège de mieux savoir que les autres ceux qu'est le bien.

Afin d'éviter en partie ces difficultés, nous proposons ici un « test » permettant d'examiner de manière critique les préférences d'un ou plusieurs individus situés dans une position à partir d'une autre position. Cette proposition peut être mise en relation avec l'idée d' « examen critique » (*critical scrutiny*) de Sen, ainsi qu'avec celle d' « impartialité ouverte ». Sen fait référence à Adam Smith pour définir le concept d'impartialité ouverte :

> *« Les illusions positionnelles peuvent imposer de sérieuses limites à la poursuite de la justice ; ces limites doivent être surmontées par l'élargissement de la base informationnelle d'évaluation, qui est une des raisons pour lesquelles Adam Smith demanda que les perspectives venant d'ailleurs, y compris de très loin, dussent être systématiquement invoquées (voir Chapitre 6). »*
>
> *« Dans nos réflexions spéculatives nous pouvons, bien entendu, envisager d'aller au-delà des points d'ancrage qui semblent fixer le monde dans lequel nous vivons et les activités corporelles qui gouvernent notre discernement et notre réflexion.[…] Tout cela n'indique toutefois pas que la positionnalité ne puisse être partiellement ou entièrement surmontée de manière à nous installer dans un point de vue moins confiné. »*[281]

L'impartialité ouverte n'est pas le point de vue « confiné » de celui qui n'a pas bénéficié du diagnostic de l'illusion positionnelle. Ce n'est pas davantage un point de vue « de nulle part ». Ces extraits situent l'évaluation positionnelle entre l'évaluation purement déterminée par les paramètres de la position occupée par l'agent, et une évaluation effectuée d'un point de vue

[281] SEN, Amartya K., *The Idea of Justice*, London: Allen Lane, 2009, p 169 et 170 (extraits). Traduction libre.

radicalement indépendant de ladite position. L'évaluation positionnelle est une « prise de distance » sans pour autant être une abstraction. Reste à spécifier la, ou les « manières » dont cette évaluation pourrait être menée. C'est ce que nous tentons de faire ici, à l'aide de l'expérience de pensée des tests de positionnalité.

L'évaluateur doit, de sa propre position, avoir une connaissance au moins suffisante des paramètres positionnels qui justifient la préférence qu'il attribue à l'agent dont il évalue la situation. Ce peut être l'agent concerné lui-même, une personne qui veut le bien de cet agent, ou une institution. Il faut donc distinguer deux types d'évaluations positionnelles :

(l) Les agents occupant la position p^2 jugent *x* préférable à *y* pour eux.

(m) A partir la position p^1, on juge *x* préférable à *y* pour les agents occupant la position p^2.

La préférence énoncée à partir de la position p^2 peut être déterminée par de fausses croyances ; si celles-ci sont identifiables à partir de la position p^1, alors une évaluation d'une position à partir d'une autre est requise. En d'autres termes, la proposition (m) pourrait être une « meilleure » énonciation de la préférence individuelle que la proposition (l). Mais comment s'en assurer ?

Lorsqu'une préférence est satisfaite, les paramètres de la position de l'agent concerné sont modifiés. Dans l'exemple proposé plus haut, Sophie obtient un chien d'aveugle et se trouve apte à se déplacer plus aisément, sans l'aide d'autres personnes. Ce changement de position effectué, on pourrait lui demander si elle souhaite rendre le chien et revenir à son ancienne position, maintenant qu'elle sait que cela est possible et qu'elle a fait l'expérience de ce changement de qualité de vie. Si Sophie répond par la négative, et dit vouloir garder le chien et donc rester dans l'état social *x,* et ce même si auparavant elle n'avait pas personnellement exprimé de préférence pour *x*, alors cela signifie que sa position s'est réellement améliorée et que la préférence qu'on lui attribuait auparavant, même sans son assentiment explicite, était effectivement sienne. C'est ce que nous qualifierons de « test positionnel » : lorsqu'un agent atteint la position préférée par une évaluation positionnelle, et ne souhaite pas revenir en arrière, alors il était légitime de lui attribuer cette évaluation, ou préférence.

La perspective critique sur les préférences positionnelles émane d'une autre position[282]. Cette position est réelle et non pas hypothétique, mais elle peut être qualifiée pour l'évaluation pour deux raisons :

- d'une part, elle peut être caractérisée comme une « position d'expert », c'est-à-dire la position de personnes susceptibles d'avoir une connaissance approfondie des paramètres des positions occupées par d'autres agents. Les positions d'experts sont celles de sociologues, d'économistes, de travailleurs sociaux, ou de personnes ayant elles-mêmes occupé précédemment les positions qu'elles peuvent à présent décrire.
- D'autre part, le fait même d'occuper une autre position, quelle qu'elle soit, crée une distance critique qui favorise une évaluation plus objective de nos préférences et des conditions de satisfaction de celles-ci. Le procédé de distanciation est bien connu en littérature : les *Lettres Persanes* de Montesquieu font ainsi jouer un rôle critique au point de vue d'un Persan visitant la France au XVIIIe siècle.

Il faut ajouter à cela que l'évaluateur occupant la position *critique* doit être disposé à vouloir les fins que l'agent lui-même sera prêt à reconnaître une fois le changement de position effectué. Il doit donc être parfaitement altruiste[283], et en aucun cas se substituer à l'évaluation de l'agent sur toute la durée : une fois l'agent parvenu à la position positionnellement préférée, si cette position passe le test positionnel, alors l'agent fait sienne la volonté de parvenir à cette position. En ce sens, les fins d'agent restent les fins d'agent, mais des fins soumises à l'examen critique (*critical scrutiny*) d'une autre position.

Ce test permet également d'éviter l'objection de Putnam faite à la contribution d'un évaluateur extérieur à la formation de la préférence de l'agent. L'agent, en effet, se réapproprie la préférence et les raisons de préférer la position dans laquelle elle se trouve à présent, plutôt que d'y consentir passivement. Si nous reprenons l'exemple de Theresa, un évaluateur occupant une autre position mais procédant correctement à l'évaluation positionnelle ne négligera pas les raisons différentes qu'aurait Theresa de choisir tel ou tel genre de vie, car ces raisons font précisément partie des paramètres positionnels qu'il lui faut

[282] Nous construisons cette analyse sur les travaux de Sen sur l'objectivité positionnelle et l'impartialité, mais lui-même s'appuie sur une interprétation possible des écrits d'Adam Smith sur le spectateur impartial. Cf. SMITH, Adam, *The Theory of Moral Sentiments*, 1759, rééd. Oxford : Clarendon Press, 1976. Trad. Fr. BIZIOU, M., GAUTIER, C., PRADEAU, J.-F., *Théorie des sentiments moraux*, Paris : P.U.F., 2011.

[283] Pour une théorie complète du bien-être comme « sollicitude rationnelle » (*rational care*), cf. DARWALL, Stephen, *Welfare and Rational Care*, Princeton University Press, 2004. La thèse de Darwall est que le bien [*good*] d'une personne correspond à ce qu'une autre personne (qui peut être la personne elle-même dans les faits) voudrait pour elle *du fait qu'elle se soucie d'elle* [cares about her]. Autrement dit, le point de vue de « la deuxième personne » est un point de vue privilégié pour évaluer le bien-être de la première personne, ce qui va à l'encontre des théories du bien-être standard, qui reposent soit sur le point de vue de la première personne, soit sur un point de vue impartial et détaché.

connaître – y compris le fait que de choisir par soi-même contribuera à la qualité de vie de Theresa. L'évaluation positionnelle ne se cantonne pas à l'observation des comportements visibles de consentement, ce que reprochait Putnam à la solution offerte par le théoricien de la décision de « forcer » la révélation des préférences de Theresa en lui imposant l'un ou l'autre des genres de vie ; l'évaluation positionnelle, au contraire, fait état du processus de formation des préférences, ainsi que des éventuels échecs de ce processus.

d) La position de l'évaluateur : bref bilan

La discussion de la position de l'évaluateur fait apparaître que les préférences peuvent être décrites du point de vue de l'agent lui-même, mais également de deux autres points de vue : le point de vue hypothétique d'un observateur impartial et disposant des connaissances requises pour définir l'évaluation appropriée, et le point de vue réel d'une personne disposée à déterminer la préférence qui correspond le mieux aux paramètres de la position qu'occupe l'agent (le point de vue de la « deuxième personne » qu'a développé par ailleurs Stephen Darwall en philosophie morale[284]). Cette dernière solution suppose des interactions en amont entre les agents, la mise en œuvre de mobiles altruistes et de mécanismes d'empathie.

[284] DARWALL, Stephen, *Welfare and Rational Care*, Princeton University Press, 2004.

Implications pour le formalisme de la théorie du choix social

Concernant la théorie du choix social, nous avons vu que la distinction entre évaluation par l'agent et évaluation par l'observateur indépendant correspondait à la distinction entre une approche en profils multiples et une approche en profil simple de la fonction de bien-être social. La proposition consistant à comparer les évaluations positionnelles actuelles des agents avec les évaluations positionnelles dans certaines conditions cognitives, remplies soit par les agents dans le futur, soit par une autre personne qui prend fait et cause pour les intérêts de l'évalué, pourrait donner lieu à des opérations d'agrégation des préférences d'une même personne dans différentes positions : ce serait une agrégation interpositionnelle, en somme. Les travaux initiés par Sen sur les métaclassements de préférences vont dans ce sens. A ceci près que l'agrégation des préférences ne serait pas seulement intrapersonnelle, mais également intertemporelle, dans la mesure où nos préférences évoluent avec les paramètres des positions que nous occupons[285]. Par ailleurs, on pourrait imaginer des conditions de restriction du domaine des préférences admissibles. Enfin, lorsqu'il s'agit d'agréger les préférences positionnelles d'individus différents afin de résoudre un problème de choix social donné, certaines préférences seront communes à des individus partageant une même position, et plus le tissu social sera homogène, plus la restriction du domaine des préférences admissibles sera envisageable. Dans la mesure où le problème du choix social sera interprété comme celui de l'agrégation des préférences propres à des positions sociales ou économiques données, il sera peut-être possible de définir certaines conditions relatives aux relations entre ces différentes positions, telles que les conditions d'absence d'envie et leurs variantes. Nous ne poursuivrons pas ces quelques suggestions qui requièrent des propositions techniques étoffées avant d'être validées.

[285]Sur les métaclassements, cf. IGERSHEIM, Herrade, *Liberté et Choix social - Contribution à l'analyse de la liberté en économie normative.* Thèse de doctorat, Université Louis Pasteur, Strasbourg, 2004 et IGERSHEIM, Herrade, « Du paradoxe libéral-parétien à un concept de métaclassement des préférences », *Recherches économiques de Louvain* 2/2007 (Vol. 73), p. 173-192.

Implication pratiques

Concernant la pratique politique, les évaluations positionnelles requièrent trois types de pratiques et d'activités, qui s'apparentent à la délibération en ses différents sens, et permettent par ailleurs de spécifier celle-ci tout en rétablissant le lien avec le choix social traditionnellement davantage associé à l'agrégation et au vote.

Un premier type de pratique requis pour mener à bien une évaluation positionnelle progressive est la réflexion individuelle. Chaque personne est amenée à s'imaginer dans une position différente de la sienne afin de comprendre ce qui, dans sa position actuelle, influence ses jugements. Chacun compare ainsi les paramètres de sa position avec ceux d'une position qu'il pourrait occuper, et en déduit les préférences qu'il a et celles qu'il peut avoir ensuite. Comment peut-on imaginer ou concevoir une position différente de la sienne ? La connaissance commune d'autres positions, c'est-à-dire de positions occupées par des personnes que nous connaissons, favorise ce type de processus de réflexion. Une connaissance plus théorique, voire scientifique, étayée par diverses sources d'informations, peut compléter cette connaissance de première main. Enfin, la lecture des romans permet de s'exercer à s'imaginer dans la position d'autrui. Il s'agit ici non pas de concevoir logiquement une autre position à l'aide de paramètres connus de façon factuelle, mais de faire preuve de ce que Martha Nussbaum appelle l' « imagination narrative »:

> *« La logique ou la connaissance factuelle seules ne suffisent pas à mettre les citoyens en rapport avec le monde complexe qui les entoure. Une troisième capacité du citoyen, étroitement liée aux deux premières, est ce qu'on peut appeler l'imagination narrative. J'entends par là la capacité à imaginer l'effet que cela fait d'être à la place d'un autre, à interpréter intelligemment l'histoire de cette personne, à comprendre les émotions, les souhaits et les désirs qu'elle peut avoir. Le développement de cette sympathie se trouve au cœur des meilleurs projets modernes d'éducation démocratique, en Occident et ailleurs. »*[286]

Il apparaît ici, avec le terme « sympathie », que la comparaison des évaluations effectuées dans différentes positions sert à la fois une meilleure connaissance de notre avantage propre et un altruisme approprié. Cela nous permet d'introduire le second type de pratique induit par l'évaluation positionnelle : il s'agit des interactions entre les personnes. Ces interactions peuvent être verbales ou non. Parmi les interactions verbales, la conversation permet l'échange d'informations sur la situation de chacun. Quant à la discussion, elle offre l'occasion d'échanger des arguments et de ce fait des raisons de préférer tel ou tel état social.

[286]NUSSBAUM, Martha, *Not for profit: Why Democracy needs the humanities*, Princeton University Press, 2010. Traduction française Solange Chavel, *Les émotions démocratiques*, Paris: Flammarion, 2011, Chapitre 6.

Quant aux interactions non verbales, à la pitié ou à la joie que l'on peut ressentir en voyant autrui souffrir ou être heureux, elles nous permettent d'étoffer ici encore les comparaisons de positions que nous effectuons lors des évaluations positionnelles.

Enfin, la troisième pratique induite par les évaluations positionnelles est l'expertise. Afin d'établir avec précision le lien entre les contenus des évaluations et les paramètres positionnels, il convient d'identifier ces derniers clairement. Selon la nature des paramètres (biologiques, médicaux, sociaux, économiques, juridiques, politiques…), différents types d'expertises pourront être mobilisés.

5.2. Le critère d'évaluation

La préférence est un jugement de valeur qui statue qu'une chose est préférable à une autre. Mais quel est le critère d'évaluation permettant cette assertion ? Quel type de « bien » décrit la préférence ? Que doit-il être pour que la préférence puisse convenir à l'approche des capabilités ?

On peut distinguer trois conceptions du bien ou du bien-être, que nous analyserons dans les sections suivantes. Cette analyse s'achève par une étude du bien-être lorsqu'il est déterminé par des préférences rationnelles. Cette conception du bien-être nous apparaît comme la plus adéquate pour intégrer une interprétation des préférences de la théorie du choix social à l'approche des capabilités. Mais, avant cela, il nous faut présenter la conception hédoniste, la conception du bien comme satisfaction des préférences et la conception objective du bien. Selon la conception hédoniste, les préférences décrivent un accroissement du plaisir ou une diminution de la douleur (5.2.1.) ; la conception du bien-être comme satisfaction des préférences décrit la réalisation d'un état social conforme à la préférence de l'agent (5.2.2.) ; la conception objective du bien stipule que la préférence décrit un état social dont la valeur serait appréciée par des personnes idéalement capables de connaître le bien (5.2.3.).

5.2.1. L'hédonisme

Une première interprétation possible de la préférence consiste à associer celle-ci à une définition hédoniste (du grec ancien ἡδονή, qui signifie « plaisir ») du bien-être. Ce serait le plaisir que nous préférerons. Cette interprétation associe le préférentialisme de la théorie du choix social à l'utilitarisme hédoniste. Une telle association n'a rien de surprenant, dans la mesure où, comme nous l'avons signalé à plusieurs reprises, la théorie du choix social et

l'économie normative s'enracinent dans la tradition utilitariste en philosophie morale et sociale.

Avant de procéder à l'analyse de cette interprétation hédoniste, le lien définitionnel entre préférence et utilité mérite d'être rappelé.. Une préférence peut être, selon l'interprétation hédoniste, exprimée en termes de comparaison d'utilités. Formellement, cela revient à une proposition telle que (α) :

- Soit *x, y* des options ;

  Soit *P* une relation de préférence stricte, réflexive, transitive et complète ;

  Soit *u(x)* l'utilité de *x* (c'est-à-dire, la mesure de la satisfaction expérimentée grâce à l'obtention de *x* ;

  (α) *x P y* est équivalente à *u(x) > u(y)*

Selon l'interprétation hédoniste, l'expérience du plaisir accroît l'utilité, tandis que celle de la peine la fait décroître. L'hédonisme associe donc l'utilité à un « contenu descriptif indépendant »[287]. Ce contenu correspond à une expérience qualifiée en termes de plaisir ou de peine. Lorsqu'un ordre de préférence représente une fonction d'utilité, il représente la magnitude de satisfaction que l'agent retire de l'existence de chacun des états sociaux comparés. C'est en ce sens que la valeur de l'utilité est « fonction » de l'existence d'un état social donné. L'utilité est la valeur assignée au degré de satisfaction résultant de l'existence d'un état social donné. Interpréter la préférence comme une comparaison d'utilités, c'est présupposer que l'ordre de préférence est déterminé par des états de satisfaction évaluables et comparables. Une fois l'analyse de la préférence comme plaisir effectuée (a), nous distinguerons deux variantes de l'hédonisme (b). Nous présenterons ensuite les principales objections contre cette interprétation de la préférence (c).

[287]SEN, Amartya ,*Commodities and Capabilities,* Amsterdam: North-Holland, 1985, p 36. Traduction libre.

a) Que décrit la préférence dans la conception « hédoniste » ?

Dans la conception hédoniste, des expériences dites de plaisir et de peine existent indépendamment de la description qu'en donne la préférence. L'utilitarisme hédoniste peut donc être dit réaliste, dans la mesure où il asserte la réalité d'une expérience, indépendamment de la description que le philosophe ou l'économiste en donne. Cette expérience est l'effet, en termes de plaisir et de peine, de l'existence d'un état social sur un individu. Toute la difficulté consiste dès lors à construire une description de cet état mental qui puisse représenter cet état mental de façon adéquate. Les débats autour des comparaisons interpersonnelles, présentées ci-avant, peuvent être compris comme l'une des manifestations de cette difficulté.

La paternité de l'utilitarisme hédoniste peut être attribuée à Jeremy Bentham. Bentham pose en effet, comme une vérité axiomatique, que la nature de l'être humain le porte à rechercher des expériences de plaisir et à fuir les expériences pénibles. Le principe d'utilité, principe permettant à la fois de décrire les comportements humains et de déterminer des normes morales et sociales appropriées, fait état de cette vérité. Dans une note de bas de page de son *Introduction aux Principes de la Morale et de la Législation,* Bentham décrit ainsi la relation empirique entre l'état de satisfaction des individus et les caractéristiques des états ou des objets que l'on évalue :

> *« Nos idées sont dérivées, toutes sans exception, des sens ; celles qui sont plaisantes ou pénibles le sont donc comme les autres : par conséquent, elles sont dérivées de l'opération des objets sensibles sur nos sens. Le bonheur d'un homme peut ainsi être dit dépendre plus ou moins de la* relation *qu'il entretient avec un objet sensible, lorsque tel objet est de sorte qu'il ait une chance, plus ou moins grande, de produire en lui, ou de lui éviter, de la peine ou du plaisir. Maintenant cela, le cas échéant, doit être fait de l'une ou l'autre de ces deux façons ; 1. de façon* active*, à proprement parler ; c'est-à-dire par un mouvement : ou, 2. De façon* passive *ou en repos, en étant mû par, ou en subissant une action : et, dans chaque cas, soit, 1. De façon* immédiate*, en agissant sur, ou en subissant l'action des organes des sens, sans l'intervention d'un objet externe quelconque : ou, 2. D'une façon plus ou moins* éloignée*, en agissant sur, ou en subissant l'action d'un objet extérieur quelconque, qui (avec l'intervention d'un nombre plus ou moins élevé de tels objets, et au bout d'un intervalle de temps plus ou moins important) en viendra à la longue à agir sur, ou à subir l'action de ces organes. Et cela est également vrai, que les objets externes en question soient des choses ou des personnes. C'est aussi également vrai des peines et des plaisirs de l'esprit, comme de ceux du corps : toute la différence est que, dans la production de ceux-ci, le plaisir ou la peine peut résulter immédiatement de la perception*

qu'ils accompagnent : dans la production de ceux de l'esprit, il ne peut résulter de l'action d'un objet des sens, autrement que par association ; *autrement dit, au moyen d'une connexion que la perception a effectué avec certaines perceptions antérieures, déjà logées dans la mémoire. »*[288]

Cette analyse détaillée de la relation entre les objets externes et l'état de satisfaction que l'on en retire met en évidence le postulat empiriste de départ. Le plaisir et la peine, comme « idées », sont le résultat d'une modification de l'expérience sensible. Que la relation entre la sensation et l'état de plaisir ou de peine soit le fait d'une action effectuée ou subie, qu'elle implique ou non des objets ou des personnes dont l'existence est distincte de celle de nos organes sensoriels, cette expérience sensible demeure le fondement de tout plaisir et de toute peine. Même les plaisirs de l'esprit résultent indirectement de l'association de perceptions passées retenues grâce à la mémoire.

Bentham élabore ici une description des mécanismes de production de l'utilité. L'interprétation hédoniste des préférences consiste à affirmer que ce sont ces relations entre perceptions sensibles et idées de plaisir ou de peine qui déterminent les préférences et sont décrites par celles-ci.

[288]BENTHAM, Jeremy, *An Introduction to the principles of Morals and Legislation*, Londres, 1780, 1823. Réédition Oxford, Clarendon Press, 1907.Ch. XVI, § 2, n.1 p 209.Traduction libre.

b) Deux conceptions de l'hédonisme

La nature et la qualité de ce « plaisir » et de cette « peine » peuvent être comprises de manière plus ou moins étroite. Cette distinction conceptuelle a été mise en évidence par Derek Parfit, dans le livre *Reasons and Persons.*

L'hédonisme étroit

L'hédonisme étroit repose sur une définition du plaisir que le philosophe utilitariste victorien Henry Sigdwick énonce ici.

> *« Comment définir le plaisir ? Il semble évident de le définir comme une sorte de sentiment qui nous plaît, que nous aimons ou (si l'on suppose la comparaison) que nous préférons. »*[289]

Cette définition du plaisir semble redondante : le plaisir est un sentiment qui nous plaît. Une telle définition ne peut nous apprendre ce qu'est le plaisir que dans la mesure où le « plaisir » et la « peine » sont des propriétés distinctes de nos expériences. Certaines de ces propriétés sont plaisantes par elles-mêmes, d'autres sont déplaisantes également par elles-mêmes. Ainsi, le plaisir serait la propriété commune d'expériences aussi différentes que celles « d'étancher une soif intense ou un désir sexuel, d'écouter de la musique, de résoudre un problème intellectuel, de lire une tragédie, et de savoir que son enfant est heureux. »[290].

Cette première version de l'hédonisme fait de l'expérience psychologique et physiologique de « plaisir » le critère des évaluations individuelles ou collectives. Une préférence décrit donc une évaluation comparative des états sociaux en fonction de la quantité de plaisir et de peine qu'ils procurent à l'agent. Ces quantités dépendent de la durée et de l'intensité du plaisir ou de la peine[291].

L'hédonisme est généralement décrit comme une conception subjective du bien-être[292]. Et en effet, les expériences de plaisir et de peine sont vécues exclusivement par le sujet, et en ce sens elles sont effectivement « subjectives ». Toutefois, ces expériences existent indépendamment de l'énoncé de la préférence, puisqu'un être vivant peut éprouver du plaisir sans manifester ou exprimer une préférence. L'expérience du plaisir consiste en effet en des

---

[289]SIGDWICK, Henry, *Methods of Ethics*, London: McMillan & Co, 1874, 1907[7], Book II, Chapter 2, § 2. Traduction libre.

[290]PARFIT, Derek, *Reasons and Persons*, Oxford :Clarendon Press, 1984, Appendix I. Traduction libre.. Parfit donne cette liste afin de critique le caractère réducteur de cet hédonisme étroit (*Narrow Hedonism*).

[291] BENTHAM, Jeremy, *An Introduction to the principles of Morals and Legislation*, Londres, 1780, 1823. Réédition Oxford: Clarendon Press, 1907, chapter IV, p 29-30.

[292] Cf. , entre autres, PARFIT, Derek, *Ibidem*, et SCANLON, Thomas, « Preference and Urgency », *Journal of Philosophy* 72 (19), 1975, pp 655-669, ANGNER, Erik, "Subjective Measures of Well-Being: Philosophical perspectives," in KINCAID, Harold and DON ROSS (Eds.) *The Oxford Handbook of Philosophy of Economics*, Oxford: Oxford University Press, 2009, pp. 560-579.

réactions physiologiques, des évènements biochimiques, une activité neurologique, qui peuvent être décrits indépendamment de la formulation du jugement de préférence ou même de la simple évaluation de cette expérience. Autrement dit, l'hédonisme étroit suppose la réalité de cette expérience, et ce indépendamment de l'expression des préférences. De ce fait, les préférences ne font que décrire une réalité de plaisir ou de peine qui existe indépendamment d'elles.

De telles préférences ne sont pas arbitraires dans la mesure où elles se réfèrent à un contenu descriptif précis. En revanche, la description de la valeur de ce contenu en termes de plaisir et de peine peut être vue comme arbitraire.

L'hédoniste étroit peut donc prétendre à une description objective des expériences de plaisir et de peine. Selon lui, il est logiquement possible d'accéder à une description de ces dernières sans passer nécessairement par la recension des préférences des individus. Il serait possible d'associer aux termes « plaisir » et « peine » la description de phénomènes biologiques ou chimiques, tels que des réactions corporelles ou des modifications neurologiques. Les préférences ne seraient que l'une des multiples voies d'accès à l'évaluation des états sociaux selon les critères de plaisir et de peine[293].

En revanche, on peut se demander si l''hédonisme étroit rend compte de manière satisfaisante de la multiplicité d'expériences susceptibles de constituer le bien-être. Par exemple, Richard Hare souligne la difficulté en soulignant qu'il faut distinguer la douleur de la souffrance.

> *« Dans des cas peu courants, il est possible d'éprouver de la douleur sans souffrir, et sans avoir de motif d'arrêter ou d'éviter la douleur, même* ceteris paribus. *Bien que le mot 'douleur' puisse être, et souvent l'est, utilisé dans l'évaluation de la même manière que 'souffrance', de sorte que l'expérience de douleur constitue* eo ipso *un motif pour arrêter ou éviter la douleur, il est aussi parfois utilisé dans la description comme le nom d'une sensation reconnaissable ou d'une classe de celles-ci ; et il y a des exemples attestés de personnes qui disent avoir de la douleur mais que cela ne leur fait rien. Tandis que ce serait une contradiction que de dire que l'on souffre tout en affirmant que cela ne nous fait rien,*

[293] Comment l'hédonisme étroit explique-t-il les préférences – avérées – pour des expériences apparemment déplaisantes (telle qu'une préférence pour le jeûne ou la mortification) ? Trois explications peuvent être avancées:

- Ces expériences ne sont en fait pas si déplaisantes qu'elles en ont l'air
- Ces expériences ne sont recherchées que comme des moyens d'obtenir ensuite des expériences vraiment plaisantes (ainsi, on jeûne afin d'obtenir le plaisir de prier avec plus d'intensité, de jouir de la reconnaissance d'autrui, de voir son pays libéré, de profiter des plaisirs offerts dans le Paradis plus tard…)
- Les agents qui recherchent ce type d'expériences se trompent, ou sont trompés ; ils croient que le bonheur se trouve dans la négation de nos tendances naturelles, alors que, conformément au principe d'utilité, nous devrions au contraire rechercher rationnellement la satisfaction de ces tendances.

et que nous n'avons pas de motif d'arrêter ou d'éviter cette souffrance, même ceteris paribus. *S'il n'y avait pas un tel motif, il n'y aurait pas de souffrance. »*[294]

La douleur est un évènement qui influence nos motivations et nos préférences, mais sans les déterminer nécessairement. Une douleur est localisable. En revanche, la souffrance telle que Hare la conçoit est constitutive de nos motivations. La souffrance est par définition ce que l'on fuit ou ce que l'on évite, et ce quelle que soit l'expérience sensible à laquelle elle correspond. Il est fréquent que la douleur soit associée à la souffrance, mais ce n'est pas une nécessité.

Une conception moins restrictive de l'hédonisme serait plus susceptible d'éviter de réduire la variété des expériences constitutives du bien-être aux seuls sentiments de plaisir et de peine. C'est le réalisme même de l'hédonisme, que nous avons souligné, qui limite la pertinence de l'hédonisme étroit : en faisant le pari de rendre compte d'un contenu descriptif complexe à l'aide de la seule dichotomie plaisir/peine, l'hédonisme étroit met en danger sa pertinence descriptive même.

[294]HARE, R. M., *Moral Thinking: its levels, methods and point*, Oxford: Oxford University Press, 1982, 5.2..

L'hédonisme du désir

La distinction entre douleur et souffrance est un cas qui montre que les expériences décrites par les préférences ne sont pas nécessairement des expériences de plaisir ou de douleur. Les préférences devraient plutôt décrire les expériences que nous recherchons, que le terme « plaisir » soit approprié pour les décrire ou non. Ainsi, l'hédonisme du désir promeut des préférences qui décrivent les expériences désirables. Les états sociaux qui suscitent ces expériences seront hissés au sommet du classement de préférences. Corollairement, les derniers du classement de préférence devraient être les états sociaux qui génèrent des expériences pour lesquelles nous éprouvons de l'aversion.

C'est Henry Sigdwick encore qui expose clairement la relation entre les états mentaux dits « plaisants » et le désir que nous en avons.

> *« Si, toutefois, nous définissons ainsi le plaisir, lorsque nous comparons les plaisirs et considérons lequel est le plus grand, nous pourrions devoir dire que les plaisirs sont plus grands ou moins grands exactement en proportion de l'influence qu'ils exercent en stimulant la volonté à agir. »*[295]

Ici, le plaisir n'est pas vu comme la propriété d'une expérience indépendamment de la préférence comme du désir que nous éprouvons pour cette expérience. Le plaisir est la caractéristique des expériences que nous désirons, pour lesquelles nous éprouvons de l'attrait. Les expériences que nous préférons sont préférées en vertu du désir éprouvé pour elles. La préférence décrit donc ce désir. Cela signifie que désir et préférence ne coïncident pas exactement.

A ce propos, notons qu'une distinction similaire, entre deux formes d'hédonisme, est mise en place par Derek Parfit. Parfit baptise cette seconde version de l'hédonisme l' « hédonisme des préférences » (*Preference-Hedonism*). Lui identifie désir et préférence. Nous ne suivrons pas ces choix terminologiques qui pourraient jeter de la confusion dans notre analyse. Nous avons vu précédemment (section 4.2.1. de notre travail) que la préférence, telle qu'elle est représentée dans la théorie du choix social, est une espèce de jugement. En revanche, le désir est une tendance vécue qui préexiste à la préférence. Nous parlerons donc d'hédonisme du désir plutôt que d'hédonisme des préférences.

Les choix terminologiques peuvent paraître troublants ici. En effet, l'' « hédonisme » du désir est mal nommé. Il qualifie de « plaisants » des états mentaux qui ne le sont pas

[295] SIGDWICK, *Ibidem.*

nécessairement au sens courant du terme[296]. Pour l'écrivain, l'expérience de l'écriture est souvent pénible ; pourtant les écrivains la désirent et la recherchent. Le philosophe James Griffin, qui a contribué à cette discussion dans son ouvrage *Well-Being* (1986), illustre cette question par le cas suivant : à la fin de sa vie, Freud, souffrant terriblement, refusa cependant tout analgésique à l'exception de l'aspirine, parce qu'il préférait rester lucide[297]. Son état mental sous analgésique aurait été plus « plaisant » au sens courant du terme ; mais il désirait plutôt l'expérience de la pensée lucide.

Le bien-être décrit par la préférence est ici subjectif en deux sens : il l'est au sens où son contenu correspond à une expérience vécue par un sujet individuel ; et il l'est au sens où la valeur de cette expérience dépend du désir de ce même individu.

Récapitulons : selon l'hédonisme étroit, le plaisir est la qualité commune d'une catégorie d'états mentaux pouvant être expérimentés de la même manière par tous les êtres capables d'éprouver des sensations. Le désir est la tendance naturelle à rechercher ces états de plaisir. La préférence décrit cette réalité naturelle. En revanche, selon l'hédonisme du désir, le plaisir est l'attribut des états mentaux recherchés ou désirés par les personnes. Le désir est une tendance à rechercher certaines expériences, mais la nature de ces expériences n'est pas déterminée et leur valeur dépend de leur attrait pour l'agent. La préférence décrit la relation entre l'état social et l'état mental, d'une part, et l'état mental et le désir d'autre part. Un état social est préféré par l'agent à un autre parce qu'il produit un état mental qui se trouve être désiré par ce même agent.

L'hédonisme étroit et l'hédonisme des désirs coïncident lorsque les expériences « plaisantes » au sens courant du terme, c'est-à-dire procurant des sensations agréables, sont également celles qui sont désirées par l'agent. Lorsque l'agent désire d'autres types d'expériences, les deux hédonismes ne coïncident plus.

[296]Il faut remarquer d'ailleurs que tous les auteurs n'utilisent pas le terme « plaisir » pour qualifier ces états mentaux. Cf., par exemple, KYMLICKA, *Contemporary political philosophy: an Introduction*, Oxford University Press, 1990, 1999 ; traduction française Marc Saint-Upéry, *Les théories de la justice: une introduction*, Paris: La Découverte, 1999. Plutôt que de « plaisir », il y est question d' « états mentaux appréciables » ou d' « expériences dignes d'être vécues ».

[297]GRIFFIN, J.P., « Are there incommensurable values ? », *Philosophy and Public Affairs 7*, No. 1, Automne 1977.

c) *Les critiques de la préférence basée sur les états mentaux*

La préférence décrit-elle correctement le bien-être lorsque ce contenu consiste en un état mental plaisant ou désiré ? Et toute préférence décrit-elle ces états *in fine* ? On peut adresser à cette conception hédoniste de la préférence deux critiques. Une première critique vise le caractère arbitraire de la représentation d'états mentaux par des préférences. Une seconde critique met en évidence le caractère réducteur de l'interprétation des préférences comme représentation d'états mentaux.

Une mesure arbitraire

La préférence peut décrire une comparaison d'états sociaux en fonction des états mentaux qu'ils procurent à l'agent, que ces états soient intrinsèquement plaisants, ou plaisants car désirés. Cependant, elle n'en offre pas la description la plus adéquate. Considérons l'exemple suivant : un observateur demande à deux personnes différentes, Awa et Balla, de goûter deux gâteaux, l'un au chocolat et l'autre au miel, et de les noter sur 10 selon le plaisir qu'ils en ont retiré. Les résultats sont les suivants :

	Utilité d'Awa	**Utilité de Balla**
Gâteau au chocolat	10/10	1/10
Gâteau au miel	0/10	0/10

Tableau 6

La fonction d'utilité, c'est-à-dire la mesure du degré de plaisir qu'Awa ressent en mangeant la gâteau au chocolat, est la plus élevée possible. Celle correspondant à la consommation du gâteau au miel est la plus faible possible. Concernant Balla, le plaisir trouvé à la dégustation du gâteau au miel est nul, mais celui procuré par le gâteau au chocolat n'est guère plus élevé. Or, si la paire constituée par ces deux types de gâteaux constitue la seule alternative, l'énoncé des préférences ordinales d'Awa et de Balla sera identique : Awa, comme Balla, préfère le gâteau au chocolat au gâteau au miel. Et, cependant, les notes qu'ils donnent au gâteau au chocolat montrent qu'ils n'ont pas eu une expérience de dégustation similaire[298].

[298] On pourrait ajouter que la note ne dépend pas que du plaisir retiré lui-même, mais aussi, par exemple, de l'enthousiasme de la personne ou de sa disposition à avoir une expérience agréable Balla est peut-être d'un naturel pessimiste, avec une faible propension à la joie ?.

Les préférences ordinales qui constituent la base d'information du cadre arrovien sont donc de bien mauvaises candidates à la description des expériences de plaisir ou de satisfaction. De fait, il est presque curieux de parler d' « utilitarisme ordinal » tant les ordres de préférences semblent peu appropriés à la représentation des états mentaux. La transformation des utilités en préférences se fait donc au coût d'une description beaucoup moins proche de la réalité des expériences vécues par les personnes. Si la préférence peut décrire des états mentaux ou des expériences, il faut avouer qu'elle n'y réussit guère.

Une interprétation réductrice

La seconde critique de l'hédonisme vise son caractère réducteur. Le caractère préférable d'un état social ne doit-il dépendre que de sa propension à plaire aux personnes, ou même à satisfaire leurs désirs, y compris les plus triviaux ? Les états sociaux ne devraient-ils être évalués qu'en vertu du plaisir ou de la satisfaction des désirs qu'ils procurent ?

Si les expériences désirées ou plaisantes sont ce qui a le plus de valeur dans l'existence d'un être humain, alors elles constituent le critère de préférence. Mais est-ce le cas ?

Si tel est le cas, seules les expériences de plaisir, de satisfaction d'un désir, ou encore de soulagement de la douleur et d'arrêt de la frustration, ont de la valeur. Mais alors l'existence de l'état social à l'origine de cette expérience est indifférente à l'évaluation. Cette conclusion heurte nos intuitions. En quel sens ? Voyons plutôt la question qu'adresse Julie de l'Espinasse au médecin Bordeu dans *Le Rêve de d'Alembert* de Diderot :

> *« Si le rêve m'offre le spectacle d'un ami que j'ai perdu, et me l'offre aussi vrai que si cet ami existait ; s'il me parle et que je l'entende; si je le touche et qu'il fasse l'impression de la solidité sur mes mains ; si, à mon réveil, j'ai l'âme pleine de tendresse et de douleur, et mes yeux inondés de larmes ; si mes bras sont encore portés vers l'endroit où il m'est apparu, qui me répondra que je ne l'ai pas vu, entendu, touché réellement ? »*[299]

Julie de l'Espinasse décrit le rêve d'un ami disparu ; bien que cet ami ne soit plus, le rêve a suffi à susciter en elle des émotions et des expériences aussi vives que s'il avait été présent. Dans le vocabulaire de l'économie normative (moins élégant que celui de Diderot, disons-le), ce cas serait décrit ainsi : il est indifférent pour l'évaluation d'un état social que cet état existe ou non, dès lors que les états mentaux recherchés sont effectivement éprouvés. Mais Julie a des difficultés à se satisfaire de cela. La présence de l'ami n'a-t-elle pas aussi une valeur ? Peut-on se satisfaire d'illusions ? Si nous répondons par la négative, la préférence ne devrait

[299]DIDEROT, *Le Rêve de d'Alembert, Entretien entre d'Alembert et Diderot* et *Suite de l'entretien*, 1830. Edition critique d'après le manuscrit autographe de Diderot, thèse complémentaire pour le doctorat ès lettres présentée à la Faculté des lettres de l'Université de Paris, par Paul VERNIERE, Paris, Didier, 1951

pas être seulement déterminée par les expériences plaisantes ou désirées, mais également par l'existence des objets qui à l'origine de ces expériences.

A cette critique, les défenseurs de l'hédonisme pourraient répondre que l'expérience du rêve de l'ami n'est en fait pas exactement la même que celle de la présence de l'ami. Elle est de qualité ou d'intensité plus faible. Cette valeur amoindrie devrait alors se traduire dans l'énonciation des préférences : Julie préfèrera l'état social avec son ami que l'état social dans lequel elle ne fait qu'en rêver.

Nozick a élaboré une expérience de pensée plus aboutie afin de réfuter la thèse selon laquelle ce sont les expériences plaisantes ou désirées seules qui ont de la valeur. Il s'agit de la machine à expériences[300]. Supposons, propose Nozick, l'existence d'une machine capable de nous faire éprouver n'importe quel type d'expériences : des neuropsychologues peuvent nous stimuler ou nous injecter des substances de telle sorte que nous soyons capables d'éprouver *exactement* les mêmes expériences que celles que nous aurions réellement lorsque nous effectuons des activités aussi variées que lire ou écrire un roman, déguster du saumon, etc. De surcroît, ces expériences sont bel et bien choisies, puisque, tous les deux ans, nous sortons de la machine afin de sélectionner dans un menu les expériences que nous souhaitons vivre ensuite (et nous oublions juste après cette excursion hors de la machine, excursion dont le souvenir pourrait nuire à la qualité des expériences recherchées). A ceci près que nous ne réalisons pas ces activités : nous sommes plongés dans une sorte de liquide nutritif, branchés à des électrodes qui nous envoient des stimuli. Nozick affirme que la thèse selon laquelle une vie de cette sorte a de la valeur est contredit nos intuitions. Cette contradiction suffit, selon lui, pour faire de l'hédonisme une doctrine fausse. Car, ce qui a de la valeur pour nous, affirme Nozick, ce n'est pas seulement ce que nous expérimentons ou ce que nous éprouvons, mais également ce que nous *faisons* et ce que nous *sommes.* Un poète ne veut pas seulement avoir les états mentaux d'une personne qui écrit des poèmes et les publie avec succès : il veut *écrire les poèmes* et *être un poète.*

Cette critique est valide dans la mesure où nous acceptons un raisonnement dont la pierre de touche est l'intuition. D'autre part, le raisonnement en question postule que, selon cette intuition justificatrice, la valeur dépend nécessairement de la réalité des objets de nos expériences de plaisir et de peine. Mais l'utilitarisme hédoniste récuse ce point de départ même, et stipule que ce qui peut avoir de la valeur, c'est l'expérience de ce qui nous arrive.

[300]NOZICK, Robert, *Anarchy, State and Utopia*, New York, Basic Books, 1977; traduction française par Evelyne d'Auzac de Lamartine, *Anarchie, Etat et Utopie*, Paris, Presse Universitaires de France, 1988, 2003, 2008.

L'ambition de Nozick est de réfuter l'utilitarisme en montrant son incohérence : mais l'utilitarisme n'a pas à être cohérent avec des hypothèses de départ qu'il ne revendique pas. Un utilitariste devrait répondre à Nozick que, si l'expérience est ce qui a de la valeur, alors, de fait, une vie avec la machine à expérience est une bonne vie. En revanche, notre propre ambition est plus modeste : nous souhaitons montrer que l'hédonisme n'est pas la seule interprétation possible des préférences, ou encore que les préférences ne décrivent pas *nécessairement* ou *exclusivement* des états mentaux de plaisir ou de satisfaction des désirs. Elles le peuvent, mais elles peuvent également décrire d'autres formes de bien-être. Et, pour notre propos, un argument intuitif tel que celui de Nozick est tout à fait suffisant : il suffit qu'il existe au moins une personne susceptible de refuser la machine de Nozick pour démontrer que les préférences ne décrivent pas exclusivement des expériences plaisantes ou désirées.

5.2.2. Le bien-être comme satisfaction des préférences

La conception du bien-être comme satisfaction des préférences constitue une alternative à l'hédonisme[301]. Il s'agit d'une autre conception subjective du bien décrite par les préférences des agents. Cette conception fait office de rasoir d'Occam en faisant abstraction des états mentaux ou des expériences des agents, pour ne retenir que le jugement de préférence. Le bien-être de l'agent est réalisé lorsque l'état social situé au sommet de son classement de préférences est réalisé. Selon cette conception, la préférence décrit l'état social dont l'existence serait un bien selon l'agent concerné, ou encore pour l'agent concerné. Cette dernière proposition mérite quelques explications.

Tout d'abord, l'état social préféré est un bien pour l'agent auquel cette préférence est attribuée. Mais cela n'implique pas que l'agent doive faire l'expérience, plaisante ou désirable, de cet état social. C'est ici même que se situe la différence avec la conception hédoniste de la préférence. Cela ne manque pas de susciter quelques interrogations : si le critère de l'expérience disparaît, quel rapport l'agent entretient-il avec l'état social préféré ? En connaît-il les caractéristiques ? Doit-il les connaître ? Cette connaissance est-elle une raison de désirer cet état social ?

[301] Derek Parfit parle de satisfaction des désirs (*Desire-Fulfilment Theory*) ; mais, comme nous l'avons vu, la distinction entre le désir comme tendance et la préférence comme jugement sur laquelle repose notre argumentation requiert un changement de terminologie. D'autre part, le terme « désir » connote encore trop un contenu empirique subjectif. Or, c'est l'éviction de ce contenu qui caractérise justement la conception du bien-être comme satisfaction des préférences par rapport à la conception du bien-être comme expérience plaisante ou désirable décrite par la préférence.

En deuxième lieu, l'état social préféré peut constituer un bien *pour* l'agent, ou *selon* lui. Il y a ici une ambiguïté. En effet, dans le cas où l'état social est un bien *pour* l'agent, on considère que l'état social sert les intérêts de l'agent, y compris des intérêts autres que la recherche d'une expérience agréable ou désirée. Cela signifie que la préférence peut être énoncée par tout évaluateur connaissant ces intérêts, qu'il s'agisse de l'agent lui-même, d'un observateur, ou encore de l'agent estimant ses préférences futures. En revanche, dans le cas où l'état social est un bien *selon* l'agent, il est fait appel à son opinion ou à son jugement, y compris lorsque les caractéristiques de l'état social n'ont aucun effet sur l'agent. Des préférences pour des états du monde ne concernant pas du tout l'agent peuvent dès lors être incluses dans la base d'information, ce qui crée des difficultés théoriques et pratiques.

Enfin, si le critère de détermination des préférences est la réalisation d'un état social, la meilleure manière de satisfaire nos préférences n'est-elle pas d'adapter nos ambitions et de faire nôtres les préférences les plus aisées à réaliser ? Lorsque le bien-être correspond à la satisfaction des préférences, les plus heureux ne sont-ils pas ceux qui ont le plus de préférences satisfaites, et de ce fait, ceux qui ont les préférences les plus aisées à satisfaire ? En ce sens, les préférences peuvent devenir elles-mêmes objets de préférence, avec des critères de facilité de réalisation, ce qui sème le doute sur l'aptitude des préférences à décrire effectivement le bien-être des agents.

a) Satisfaction des préférences et connaissance des états sociaux

Si le bien-être correspond à la satisfaction des préférences par un état social donné, cela signifie qu'une préférence décrit, par anticipation, l'état social qui la satisfera (et ceux qui la satisferont moins, voire pas du tout).Et, lorsque cette préférence est décisive, cela signifie que l'agent a la garantie que l'état social résultant du choix collectif correspondra à la description donnée par la préférence[302]. Or, il semble paradoxal de connaître les déterminations de notre satisfaction avant que cette dernière n'ait eu lieu. Ce paradoxe est identifié par Wittgenstein :

> *« Il semble que le souhait sache déjà ce qui le satisfera ou pourrait le satisfaire, et la phrase – la pensée – ce qui la rend vraie, même si cela n'est aucunement là ! D'où vient donc cette détermination de ce qui n'est pas encore là – cette exigence despotique ? (« La dureté du « Il faut » logique ».) »* [303]

Wittgenstein souhaite souligner le caractère trop exigeant d'une thèse sur le caractère intentionnel de la pensée : ce serait trop demander à la pensée que de lui assigner la tâche de déterminer ce qui n'est pas encore là. Si notre enquête ne concerne pas directement la nature de la pensée et la manière dont elle se rapporte au sens, mais plus spécifiquement les conditions dans lesquelles une préférence peut décrire le bien-être, il n'en reste pas moins que le problème d'une exigence déraisonnable subsiste. En définissant le bien-être comme simple satisfaction des préférences, on demande à la préférence d'anticiper ce que sera la réalité : l'agent qui exprime la préférence est ainsi privé de la possibilité de se laisser surprendre par la réalité.

Deux problèmes remettent en cause l'idée selon laquelle nos préférences décrivent l'état social dont la réalisation serait un bien pour l'agent : le problème des croyances erronées et le problème de l'occurrence différée.

---

[302] SEN, Amartya, "The Impossibility of a Paretian Liberal", *Journal of Political Economy*, vol. 78, no 1, jan-mar 1970, pp 152-157.

[303] WITTGENSTEIN, Ludwig, Philosophische Untersuchungen, Blackwell Publishers, 1953. Traduction française par F. Dastur, M. Elie, J.-L. Gautero, D. Janicaud, E. Rigal, Recherches philosophiques, Paris : Gallimard, 2004. § 437.

Le problème des croyances erronées

Le premier problème est que nous pouvons préférer un état social alors que nous ne le connaissons pas assez pour savoir s'il va réaliser notre bien. Le philosophe et théoricien de la justice canadien Will Kymlicka donne l'exemple d'un groupe d'amis qui choisit de commander une pizza au lieu d'un plat chinois[304]. Mais ils ignoraient que la pizza était empoisonnée. Les personnes qui ont exprimé une préférence pour cette pizza ont fini la soirée aux urgences. Préféraient-elles l'état social qui a résulté de la réalisation de ladite préférence ? Probablement pas. Elles n'ont pas anticipé toutes les caractéristiques de l'état social qu'elles ont choisi. Par conséquent, dans la conception du bien-être comme satisfaction des préférences, il serait inexact de dire que le bien consiste dans la satisfaction de nos préférences, du moins des préférences actuelles.

[304]KYMLICKA, *Contemporary political philosophy: an Introduction,* Oxford University Press, 1990, 1999 ; traduction française Marc Saint-Upéry, *Les théories de la justice: une introduction,* Paris: La Découverte, 1999, p 25.

Le problème de l'occurrence différée de l'état social

On considère généralement que les préférences, lorsqu'elles sont satisfaites, le sont dans un laps de temps et dans un environnement suffisamment proches, de sorte que l'agent puisse prendre connaissance de la réalisation de sa préférence. Mais la conception du bien-être comme satisfaction des préférences fait l'économie de l'expérience correspondant à la satisfaction. De ce fait, il n'est plus nécessaire de faire l'expérience de la réalisation de nos préférences pour être satisfaits. Peut-on dire que la préférence décrit le bien de l'agent lorsque ce bien correspond à un état social que l'agent ne connaîtra jamais ? Voilà qui peut rendre perplexe.

Derek Parfit fait éprouver cette perplexité à ses lecteurs grâce à l'expérience de pensée suivante : supposons, dit Parfit, que je rencontre un étranger atteint d'une maladie que l'on croit incurable. J'éprouve de la compassion à son endroit, et je souhaite de tout mon cœur qu'il soit guéri. Nous ne nous revoyons plus jamais, mais, contre toute attente, l'étranger guérit sans que je le sache. D'après la conception du bien-être comme satisfaction des préférences, la guérison de l'étranger rend ma vie meilleure. Derek Parfit écrit que ce n'est pas plausible[305]. Or, si le bien-être de l'agent consiste dans la satisfaction de ses préférences, qu'il fasse ou non l'expérience de cette satisfaction, le cas de l'étranger ne remet pas en cause cette conception. Pourquoi alors ne serait-ce pas plausible ?

En fait, c'est le terme « bien-être » [*well-being*] qui suscite l'ambiguïté. Il laisse entendre que l'agent doit « être bien » grâce à la satisfaction de ses préférences. Mais, en ce cas, il n'y aurait pas de différence entre la conception du bien-être comme réalisation d'états mentaux désirés et celle d'états sociaux préférés. Or, l'intérêt de la conception du « bien-être » comme satisfaction des préférences est sa capacité d'accueillir des critères de détermination de la préférence autres que la recherche d'une expérience plaisante ou désirable. Il serait donc plus approprié de parler simplement de « bien » de l'agent. Or, on peut juger bon que l'étranger guérisse, tout en ignorant l'occurrence de cette guérison. En d'autres termes, la préférence coupée de l'expérience ne représente pas seulement le bien *pour* moi – le fait d' « être bien », mais également le bien *selon* moi – le fait de « juger bon » - [306]. La préférence peut exprimer l'intérêt comme le jugement moral.

On peut affiner l'analyse de l'exemple de Parfit en différenciant deux préférences :

---

[305]PARFIT, Derek, *Reasons and Persons*, Oxford: Clarendon Press, 1984, Appendix I.

[306] Les défenseurs d'une conception hédoniste de la préférence soutiendront que les états sociaux « bons » selon moi me procurent, indirectement, une expérience agréable (le plaisir de voir les autres heureux).

- la préférence pour ma présence dans un état social dans lequel l'étranger est guéri et
- la préférence pour un état social dans lequel il est guéri, que je sois présente ou non.

La question de l'occurrence différée est liée à une autre difficulté du préférentialisme : l'absence, au moment de l'évaluation et du choix collectif, des agents porteurs des intérêts en jeu.

b) Satisfaction des préférences et réalisation des intérêts

Nous avons vu que l'occurrence différée de l'état social a des implications insatisfaisantes si l'on juge que l'expérience de la réalisation d'un état social est requise pour que la préférence puisse être dite « satisfaite ». Mais l'écart temporel entre l'occurrence de la préférence et celle de l'état social préféré crée également des difficultés lorsque l'agent change de préférence entre-temps. Si on reprend l'exemple de Parfit, je peux ne plus être satisfaite par la guérison de l'étranger parce que mes préférences ne sont plus les mêmes qu'à l'époque où je l'avais souhaitée de tout mon cœur. Faut-il alors ne prendre en compte dans le choix social que les préférences réalisables dans un futur proche, puisque ce sont les seules dont la satisfaction correspondra toujours aux conceptions de l'agent ? Mais ne serait-ce pas arbitraire de restreindre les préférences admissibles de cette manière ? Après tout, les préférences ayant le plus d'importance pour les politiques publiques et les choix collectifs sont les préférences dont la réalisation est de longue haleine[307].

[307] Puisque nous envisageons toutes les possibilités logiques, remarquons cependant que les préférences ne peuvent pas avoir pour objet des états sociaux passés. Elles peuvent en revanche porter sur l'interprétation et la mémoire de ce passé.

Quelles préférences individuelles ?

D'après la conception du « bien-être » comme satisfaction des préférences, la préférence décrit l'état social dont l'existence serait un bien selon l'agent concerné, ou encore pour l'agent concerné. Lorsque les préférences décrivent le bien « pour » l'agent, elles peuvent être associées à l'idée d'intérêt. Lorsque les préférences décrivent le bien « selon » l'agent, elles relèvent de l'opinion[308]. Puisque l'écart entre l'occurrence de la préférence et celle de l'état social a des conséquences intuitivement gênantes, on peut être tenté de limiter la conception du bien-être comme satisfaction des préférences à la satisfaction des préférences qui concernent l'agent, c'est-à-dire de celles qui représentent son intérêt. Cette approche, selon laquelle le bien-être correspond à la satisfaction des préférences qui ne concernent que l'existence de l'agent est, dans la terminologie de Parfit, la *Success Theory*.

Mais rappelons que la distinction entre les préférences déterminées par la recherche de l'intérêt et d'autres préférences, n'est pas représentée dans le cadre formel de base de la théorie du choix social, comme le rappelle Arrow :

> *« D'un point de vue formel, on ne peut distinguer entre la gêne d'un individu ayant son jardin pollué par des fumées d'usine et son dégoût extrême du fétichisme en Afrique centrale. Un certain nombre d'Américains considéreraient probablement la première situation comme étrangère à la politique sociale, la seconde en relevant, quoique la majorité d'entre eux serait vraisemblablement d'un avis opposé. Je désire simplement souligner que l'on doit s'intéresser à la totalité du système de valeurs, y compris les jugements de valeur, pour trouver une théorie vraiment générale du bien-être collectif. »*[309]

La représentation formelle des préférences ne permet pas de différencier les préférences qui décrivent l'intérêt de l'agent, c'est-à-dire les modalités d'un état du monde qui l'affectent directement, des préférences qui décrivent des systèmes de valeur dont la réalisation n'affecte pas directement l'agent. L'éradication du fétichisme en Afrique Centrale n'affecte pas directement l'existence d'un Américain. En souhaitant que l'on s'intéresse « à la totalité du système de valeurs », Arrow fait référence à la justification normative de l'axiome de non-restriction du domaine des préférences admissibles : un choix collectif « authentique » ne peut exclure sur des bases arbitraires certaines préférences. Un choix collectif n'est pas *a priori*

---

[308] Cf. SEN, Amartya, « Social Choice Theory : A Reexamination", *Econometrica*, 45, Janvier 1977, pp 53-89. Rééd.*In Choice, Welfare and Measurement*, Oxford: Basic Blackwell Publishers, 1982. Dans cet article, Sen distingue deux interprétations des préférences, l'une en termes d'agrégation des interêts, l'autre d'agrégation des jugements.

[309] ARROW, Kenneth, *Social Choice and Individual Values*, New Haven, Yale university Press, 1951 (2ème édition révisée, 1961).Traduction française par l'Association de Traduction Economique de l'Université de Montpellier, Paris, Calmann-Levy, 1974, p 47.

seulement une agrégation d'intérêt ; il peut également être la mise en commun d'opinions sur des caractéristiques du monde qui ne concernent pas directement l'agent.

L'économie du critère de l'expérience rend la distinction entre ce qui m'affecte et ce qui ne m'affecte pas malaisée. Supposons, propose Parfit, qu'un nommé Derek se coupe la joue en se rasant : il s'agit d'un changement du monde qui le concerne. Mais, en un sens, ce changement concerne également Confucius, puisqu'il est devenu vrai que Confucius a vécu sur une planète où, quelques siècles plus tard, une joue aura été coupée. Confucius aurait pu avoir une préférence pour un monde dans lequel, quelques siècles plus tard, la joue de Derek n'est pas coupée par le rasage.

L'exemple de Parfit est un cas limite ; si nous résolvons la difficulté qu'il soulève, nous serons *a fortiori* armés pour résoudre les difficultés plus concrètes que signale Arrow dans les domaines que sont les politiques environnementales ou encore les relations entre des cultures différentes. La définition de ce qui concerne les agents ne dépendant plus de l'état mental de ce dernier, il apparaît dès lors qu'une description plus précise des états sociaux seuls permet de distinguer les intérêts bien compris de l'agent. Ainsi, il faudrait analyser plus précisément dans quelle mesure l'accident de rasage survenu au XXe siècle pourrait affecter Confucius. Il n'affecte pas les conditions de vie ou de survie de Confucius, son état de santé ou son espérance de vie. En revanche, si Confucius avait consacré toute son existence à la lutte contre les accidents de rasage, alors, en effet, ce changement le concerne. En résumé, nous ne pouvons savoir si une préférence pour un état social est une préférence qui « concerne » l'agent que si nous en savons davantage sur les états sociaux, leur fonctionnement et leur relation aux préférences des agents. La manière dont on décrit un état social sera donc de toute première importance.

La distinction entre les préférences « qui me concernent » et celles « qui ne me concernent pas » pourrait cependant être traitée du point de vue de la défense des libertés individuelles. Pensons aux préférences intrusives du libertin et du prude dans le paradoxe de Sen : chacun veut forcer l'autre à lire ou ne pas lire certains livres[310]. Si seules les préférences qui « concernent » l'agent qui exprime ces préférences sont prises en compte, ni le prude ni le libertin n'auront leur mot à dire à propos des lectures d'autrui. Mais en ce cas les préférences

[310]SEN, Amartya, "The Impossibility of a Paretian Liberal", *Journal of Political Economy*, vol. 78, no 1, jan-mar 1970, pp 152-157.

motivées par des principes moraux ou altruistes seraient également exclues. Si un agent préfère que l'on guérisse les malades du SIDA sans être elle-même atteinte ou susceptible de l'être, est-ce une préférence intrusive ?

Ici encore, il faut décrire avec plus de précision les états sociaux préférés, afin de comprendre ce qui fait qu'ils sont préférés. La préférence du libertin consiste à « forcer » le prude à lire l'*Amant de Lady Chatterley*. La réalisation de cette préférence consisterait donc en un accroissement du plaisir du libertin et une privation de liberté du prude, et peut-être également une privation de bien-être. Il serait ainsi intéressant de décomposer la description de l'état social. Le libertin souhaite-t-il le plaisir de voir le prude souffrir ? Ou pense-t-il que ce dernier se sentira libéré après avoir lu l'*Amant de Lady Chatterley* ? La préférence pour l'état social correspondant à la description suivante : « le prude lit l'*Amant de Lady Chatterley* » peut correspondre en fait à *deux* préférences distinctes, visant chacun un état social différent. L'une d'elle rechercherait l'état où « le libertin savoure le plaisir de voir le prude choqué par le contenu du livre *l'Amant de Lady Chatterley* » ; l'autre concevrait son but plutôt dans les termes suivants : « le prude lit l'*Amant de Lady Chatterley* et prend connaissance de la possibilité de vivre différemment. ». En affinant la description des états sociaux, il apparaît que la préférence intrusive peut également concerner l'agent – le libertin éprouve du plaisir dans le cas où le prude est forcé à lire l'*Amant de Lady Chatterley*. Distinguer les préférences qui concernent l'agent de celles qui ne le concerne pas ne résout donc pas le problème des préférences intrusives. Ce n'est pas cette distinction qui permet d'exclure certaines préférences. Nous avons besoin soit d'un autre principe supplémentaire, soit d'une description plus précise des états sociaux préférés.

La préférence définie indépendamment de l'expérience doit donc décrire l'état social préféré de telle sorte que la position de l'agent par rapport à cet état social soit spécifiée. Ainsi, nous sommes en mesure de dire si cet état « concerne » à proprement parler l'agent.

Quelles préférences collectives ?

Quelles préférences collectives décrivent le bien de la collectivité ? En parlant de préférences collectives, on peut en fait évoquer deux évaluations différentes. La préférence collective peut être la préférence d'une collectivité d'agents : il s'agit alors de la préférence de la collectivité qui l'exprime. Mais elle peut également être la préférence d'une collectivité de patients, c'est-à-dire de l'ensemble des personnes qui seront affectées par la réalisation de cette préférence. Les « patients » sont ici les « patients moraux », c'est-à-dire ceux qui sont affectés par des

décisions supposant des enjeux normatifs, sans pour autant être en mesure de participer activement à ces décisions et de contrôler leurs existences[311]

[311] VECA, Salvatore, "La Possession", in Collectif, *Quelle philosophie pour le XXIe siècle? L'Organon du nouveau siècle,* Paris : Gallimard/Editions du Centre Pompidou, 2001.

La collectivité d'agents

Si la préférence collective est celle de ceux qui expriment une préférence, cela présuppose la capacité d'exprimer cette préférence. Cette capacité relève du fait, mais aussi du droit.

Concernant la capacité de fait, peuvent participer à la construction d'une préférence collective ceux qui remplissent les deux conditions suivantes : ils sont présents (ils existent ici et maintenant) et ils ont les aptitudes minimales requises pour participer à un processus de décision collective (la capacité de délibérer et de déterminer l'état social le meilleur selon les critères retenus pour la délibération, ainsi que la capacité de réfléchir calmement, sans être sous l'emprise d'émotions, de passions ou encore d'obsessions).

La frontière entre ceux qui jouissent de cette capacité et les autres peut être floue dans les faits : d'une part, les critères de présence sont spatio-temporels, mais des personnes apparaissent et disparaissent de la collectivité à tous instants. La délimitation d'un territoire et d'une période susceptibles de définir la collectivité semble donc arbitraire *a priori*. Toutefois, ces critères de présence sont ceux qui sont souvent implicitement retenus pour définir la collectivité.

D'autre part, la définition des aptitudes à participer à la prise de décision collective fait appel à des critères du développement psychologiques qui fixent des seuils également aveugles eu égard au caractère flou des frontières entre ceux qui disposent de toute leur raison et ceux qui en manquent quelquefois ou souvent. Le simple fait de fixer l'âge minimum pour voter à 18 ans est arbitraire en un sens : pourquoi ne pas le fixer à 17 ans et 364 jours ?

Concernant la capacité de droit de la collectivité d'être l'agent exprimant une préférence, elle implique que la collectivité se voit reconnaître le droit de faire valoir cette préférence, à ses membres comme à une autre collectivité. Mais comment légitimer ce droit ?

On peut concevoir des critères d'inclusion extérieurs à la décision collective. Les critères identitaires, quoique flous, sont généralement pris comme constitutifs de la collectivité qu'est l'Etat-Nation : l'appartenance à un territoire donné en un temps donné justifierait ici le droit de faire valoir ses préférences dans le processus de choix collectif. Cette « appartenance » peut à son tour être légitimée de deux manières. Soit elle relève de la filiation naturelle – et correspond alors au « droit du sang » -, soit elle relève d'une présence suffisamment durable pour qu'elle puisse être reconnue – et cela correspond donc au « droit du sol ». L'argument de la filiation naturelle porte son attention sur l'identité de l'agent qui énonce la préférence collective. Sa capacité à énoncer cette dernière dépend en effet non pas des conséquences du choix collectif, mais de la nature même de l'agent. Cet argument a des implications absurdes :

en effet, il nous faudrait, pour déterminer les choix collectifs légitimes pour un territoire, le territoire français, par exemple, retrouver les descendants de ses tous premiers habitants. Or, pour ces derniers, il est fort possible que les décisions concernant la France ne reflètent ni leurs intérêts ni leurs valeurs[312]. L'argument de la reconnaissance de la durée de présence pose à nouveau les questions de définition des frontières.

Une seconde manière de légitimer le droit de la collectivité-agent à faire valoir sa préférence consisterait à faire de la définition même de la collectivité un objet de choix collectif. Sans aller jusqu'à la question de la justification originelle, l'extension du suffrage peut déjà faire l'objet d'un vote. De nombreuses entités politiques ont exclu ou excluent certaines catégories de population du processus démocratique : Athènes ne considérait pas les femmes, les étrangers ou les esclaves comme des citoyens. En France comme dans de nombreux pays, les mineurs, les citoyens étrangers et certains auteurs de crime (pour une durée limitée) sont privés de leurs droits civiques. Deux questions se posent alors. Ces exclus doivent-ils participer à la décision de les inclure ou non dans la collectivité ? Et, d'autre part, choisiront-ils nécessairement d'être inclus dans les processus de décision ? Ne préfèreront-ils pas conserver leur statut de « patients » subissant les effets du choix collectif ? Si seuls les agents participant actuellement aux décisions collectives ont le droit de décider qui peut être inclus ou non dans la collectivité, ceux qui n'en font pas encore partie pourront craindre de continuer d'être exclus. Par ailleurs, cela entraîne un problème de régression à l'infini des justifications. Remarquons que les personnes qui sont exclues du processus de décision collective peuvent être contre leur inclusion dans ce même processus. Ainsi, avant l'octroi du droit de vote aux femmes en Suisse (en 1979), il a existé une « Fédération des femmes suisses contre le droit de vote des femmes » créée en 1959[313].

Jon Elster et Aanund Hylland proposent l'argumentation suivante pour justifier les frontières de la collectivité par un choix collectif, en prenant l'exemple de l'âge du vote : à chaque classe d'âge *n*, on demande si les personnes âgées d'au moins *n* années devraient pouvoir participer à la décision collective[314]. On demande à ceux âgés de 21 ans et plus si l'âge minimum de vote devrait être fixé à 21 ans ; puis on demande à ceux de 20 ans et plus si l'âge

[312] A propos du caractère construit des « identités nationales, cf. THIESSE, Anne-Marie, *La création des identités nationales*, Paris : Seuil, 1999. Sur la notion d'identité en philosophie, cf. BENMAKHLOUF, Ali, *L'identité : une fable philosophique*, Paris : Presses Universitaires de France, 2011.

[313] Nous devons cet exemple au cours donné par Jon Elster : ELSTER, Jon, « Les décisions collectives », cours donné au Collège de France, Paris, 15 janvier-7 mai 2009. Retranscription audio : http://www.college-de-france.fr/site/jon-elster/#|q=/site/jon-elster//_audiovideos.jsp| (consulté le 20 décembre 2012).

[314] ELSTER, Jon and HYLLAND, Aanund (eds), *Foundations of social choice theory*, Cambridge University Press, 1986, p 8.

minimum de votre devrait être fixé à 20 ans ; et on répète l'opération pour les âges de 19 ans, 18 ans, 17 ans, jusqu'à ce qu'une majorité simple décide qu'il conviendrait mieux de ne pas attribuer le droit de vote en-deçà d'un âge donné. Si, par exemple, la majorité des personnes âgées de 16 ans et plus juge que le vote devrait être interdit aux personnes âgées de 16 ans et moins, tandis que la majorité des personnes âgées de 17 ans et plus juge que ceux qui ont 17 ans révolus devraient pouvoir participer à la décision collective, alors l'âge minimum de vote sera fixé à 17 ans. Cette procédure présente l'avantage d'inclure les potentiels exclus dans la décision d'exclusion. Cet argument suppose que les citoyens sont globalement assurés de faire partie successivement de chaque classe d'âge. Même si les personnes âgées de 16 ans sont privées du droit de vote, elles savent que c'est temporaire.

En revanche, lorsque la population est divisée non pas selon son âge, mais selon sa région d'origine, par exemple, les minorités peuvent se trouver spoliées de leurs droits avec une telle méthode. Supposons un grand Etat-Nation, la Tomanie. La Tomanie a annexé il y a quelques années l'Österlich, qui est une région petite mais riche. Au sein même de l'Österlich, existe une communauté religieuse, les Pastafari, qui ont été persécutés par les Österlichiens en raison de leur refus de la religion dominante. Le gouvernement tomanien leur a offert sa protection. Si la décision de définir la collectivité est prise à l'échelle de la Tomanie, les Tomaniens choisiront d'inclure les Österlichiens. Mais, dans Österlich, il y a une majorité en faveur de l'indépendance, afin de profiter au mieux de leurs richesses. En revanche, les Pastafari s'accordent pour désirer demeurer au sein de la Tomanie. Contrairement à l'exemple précédent, personne ne peut s'attendre à être successivement un Pastafari, un Österlichien et un Tomanien. Les minorités impliquées verraient donc leur droit à déterminer par elles-mêmes leur existence violés lorsque la définition de l'appartenance à la collectivité est laissée aux soins de la majorité.

En résumé, si la préférence décrit le bien de l'agent comme un état social correspondant, ce n'est peut-être pas une bonne idée que de comprendre la préférence collective seulement comme la préférence exprimée par la collectivité. En effet, la réalisation d'une préférence collective aboutit à l'existence d'un état social dans lequel seront incluses des personnes qui n'auront peut-être pas eu accès au processus de choix collectif les concernant.

La collectivité de patients

Ne prendre en compte que les préférences collectives d'un groupe disposant de la capacité de les faire valoir activement est donc arbitraire. Une solution alternative consiste à considérer comme une préférence collective la préférence de tous ceux qui sont concernés par ses effets.

En théorie de la démocratie, cela correspond au « principe d'inclusion de tous les affectés » (*all-affected principle*) [315].

Dans un premier temps, on peut se demander ce que signifie « être affecté ». D'après la conception hédoniste de la préférence, le fait d'être affecté consiste en une modification des expériences pertinentes pour la préférence. L'hédonisme étroit y verra une augmentation ou une diminution de plaisir ou de peine. L'hédonisme des désirs décrira le phénomène en termes de réalisation des expériences désirées, ou bien de frustration.

Mais ce sont les préférences conçues indépendamment de l'expérience qui sont notre objet d'études dans cette section. Nous avons vu que ces préférences, lorsqu'elles sont correctes du point de vue cognitif, décrivent l'état social dont la réalisation est conforme aux intérêts ou aux valeurs de l'agent. Dans les paragraphes précédents, nous avons défendu l'hypothèse selon laquelle il convient de préciser la description de cet état social afin de déterminer comment l'agent s'y situerait. Les caractéristiques de l'état social sont relatives à celles de l'agent, dépendent de ses motivations. Si on applique cet argument aux préférences collectives, cela signifie que la satisfaction de la préférence attribuée à une collectivité correspond à la réalisation d'un état social avec lequel cette collectivité est en relation : soit l'intérêt de cette collectivité correspond à certaines caractéristiques de cet état social, soit ses valeurs ne peuvent être représentées que par la conformité de cet état social à certains critères normatifs.

Par conséquent, nous devons ici également nous intéresser à la description des états sociaux correspondant à la satisfaction d'une préférence collective. Or, les choix effectués par une collectivité-agent aboutissent fréquemment à la réalisation d'états sociaux affectant des membres extérieurs à cette collectivité. Il serait donc légitime, lors du processus de choix collectif, de prendre en compte les préférences des personnes affectées, y compris lorsqu'elles ne sont pas aptes à les exprimer. Si la préférence collective a pour objet un état social dont la description montre que des agents ne participant pas à son élaboration ont été affectés, alors

[315]WHELAN, F. G.'Prologue: Democratic Theory and the Boundary Problem', in R. J. Pennock and J.W. Chapman (eds), *Liberal Democracy*. NewYork: NewYork University Press, 1983, pp. 13–47 ; NÄSSTROM, Sofia, "The Challenge of the All-Affected Principle", *Political Studies*, vol. 59, pp 116-134, 2011.Il faut distinguer le principe d'inclusion de tous les assujettis du principe d'inclusion de tous les affectés. Le premier stipule que ceux qui sont sujets de la loi, ceux qui sont tenus de la respecter, devraient aussi en être les auteurs. Il s'adosse généralement à une entité politique déjà existante. Le second est plus radical, en ce qu'il ne tient pas l'existence d'une entité politique comme l'Etat-Nation pour acquise : il stipule que tous ceux dont l'existence est affectée par des décisions devraient avoir leur mot à dire concernant ces décisions. Ce principe est à la fois très exigeant et très souple : très exigeant, en ce qu'une décision politique peut affecter un nombre incalculable de personnes ; très souple, en ce qu'il peut s'adapter à des contextes de décision démocratique très divers, tels que les institutions internationales, les collectivités locales, les organismes spécifiques tels que les associations, les entreprises, les familles…

elle n'est pas justifiée et ne peut être à proprement parler qualifiée de « préférence collective » : en bonne logique, il s'agira plutôt de la préférence d'un groupe pour la collectivité – un cas d'oligarchie -[316] .

Deux exemples permettent de montrer les exigences qu'impliquerait une telle conception de la préférence collective. Les décisions d'un Etat-nation en matière de politique extérieure affectent les peuples d'autres Etats. Il en va de même pour le choix de politiques protectionnistes dans une économie globalisée. Par conséquent, si la réalisation d'un état social affecte les agents qui sont liés à ce dernier, alors les modes de prises de décision devraient prendre en compte les préférences actuelles ou probables des agents concernés. Le processus de choix social ne devrait pas correspondre à une entité localisable, mais être fonction des personnes affectées, à moins que le fait de se trouver sur un territoire implique que l'on sera affecté. Mais la localisation est subordonnée au principe d'inclusion de tous les affectés.

Notons que l'inclusion des préférences des personnes appartenant aux générations futures soulève des difficultés propres. Tout d'abord, l'existence de ces personnes n'est pas certaine et dépend de décisions de procréer de la génération présente. C'est ce que Derek Parfit appelle le problème de la non-existence. Ensuite, l'identité de ces personnes dépend également de nos actions présentes. Il s'agit du problème de non-identité chez Parfit. Enfin, même si l'identité des personnes appartenant aux générations futures était déterminée, celle de leurs préférences dépend également de nos actions présentes, notamment des choix en matière d'éducation[317].

[316] Nous nous fondons sur les propriétés logiques du cadre conceptuel de la théorie du choix social pour affirmer cela. Autrement dit, ce sont des thèses valables *a priori*, et qui excluent artificiellement une réflexion sur la forme des institutions. Notre travail porte sur les concepts de « choix collectif » et de « préférence collective », conçus comme des concepts logico-normatifs et non comme des concepts empiriques et descriptifs. De ce fait, il est possible que ce que l'on qualifie de « démocratie » en vertu de sa structure institutionnelle, ou de nos habitudes descriptives, puisse parfaitement être une « dictature » ou une « oligarchie » lorsque les exigences logiques et normatives très fortes de la théorie du choix social lui sont appliquées. Tout au long de nos recherches, nous avons été guidée par l'intuition que ce décalage entre la perspective logico-normative de la théorie du choix social et les descriptions historiques et empiriques des systèmes politiques que nous connaissons pouvait faire de la première un outil critique puissant.

[317] Sur ce sujet, cf. BYKVIST, Krister, « Preference formation and intergenerational justice », in GOSSERIES, Axel and MEYER, Lukas, *Intergenerational Justice*, Oxford University Press, 2009, pp 301-322. Bykvist s'appuie sur les implications contradictoires de l'attribution de préférences "comparatives" aux générations futures pour défendre une conception de la préférence comme désir, c'est-à-dire comme attirance positive pour un état de choses, qu'il y ait comparaison ou non. Il y défend également une filtration de ces désirs.

c) Préférences et faisabilité

Les sections a) et b) viennent de mettre en évidence l'importance de la connaissance des états sociaux préférés pour la conception du bien-être comme satisfaction des préférences. La section a) montre qu'une préférence ne peut décrire le bien-être de l'agent, ou le bien selon lui, que si elle témoigne d'une connaissance appropriée des états sociaux correspondants. La section b) souligne l'importance de la description de ces états sociaux pour comprendre ce qui motive l'agent à la préférer (ou non), et en quel sens ces états sociaux les concernent. Mais la connaissance des états sociaux ne se limite pas à appréhender leur contenu ; elle concerne aussi l'ensemble des états sociaux faisables. Or, la perception de ce qui est faisable et de ce qui ne l'est pas ne correspond pas toujours à la réalité. Nous pouvons croire qu'un changement n'est pas faisable, alors qu'il l'est. Lorsque les agents croient que l'état social préféré n'est pas réalisable, ils modifient leur préférence en fonction, et désignent comme « bon » un état social moins satisfaisant pour eux mais également moins ambitieux. Il s'agit du problème des préférences adaptatives : ces préférences sont moins fiables parce qu'elles décrivent non pas l'état social que l'agent juge réellement bon, mais celui qu'il s'est résigné à choisir[318]. Les préférences adaptatives mettent au défi la conception du bien-être comme satisfaction des préférences de décrire correctement le bien des agents. En effet, il semble que cette conception suppose que c'est le fait qu'un bien est préféré qui lui donne sa valeur, alors qu'intuitivement, nous dirions plutôt que c'est parce qu'un bien a de la valeur que nous le préférons.

Les préférences adaptatives sont définies comme des préférences construites en réponse à la limitation de l'ensemble d'options réalisables. Cet ensemble d'options réalisables est également appelé « environnement de choix » dans la littérature de la décision. La préférence adaptative est donc un ajustement des aspirations aux possibilités, comme l'illustre la fable du renard et des raisins que Jon Elster évoque dans le chapitre qu'il consacre à cette question : un renard voulait manger des raisins, mais, ne pouvant les attraper, il décréta qu'ils étaient trop verts[319]. En ce sens, une préférence adaptative est l'opposé de l' « attrait du fruit défendu ». Elster souligne le caractère problématique des préférences adaptatives pour l'utilitarisme, les

[318]Cf. entre autres, ELSTER, Jon, "Sour grapes – utilitarianism and the genesis of wants", *in* Sen, Amartya Kumar and Williams, Bernard (eds), *Utilitarianism and beyond,* Cambridge: University Press; 1982, SEN, Amartya , *Commodities and Capabilities,* Amsterdam: North-Holland, 1985, Appendix B; BOVENS, Luc, « Sour Grapes and Character Planning », *The Journal of Philosophy,* Vol. 89, No. 2, feb. 1992, pp 57-78 ;
NUSSBAUM, Martha, « Adaptive preferences and women's options », *Economics and Philosophy*, vol. 17, Issue 01, 2001, pp 67-88.

[319]ELSTER, *Ibidem.*

théories de la justice qui font appel aux préférences ainsi que pour la théorie du choix social : en effet, les préférences adaptatives étant construites en réaction à la restriction des options et non pas à partir des valeurs et des intérêts réels des agents, elles ne constituent pas une base d'information et d'évaluation fiable. John Stuart Mill illustrait ce phénomène au XIXe siècle par l'exemple de la préférence des femmes mormones pour la polygamie, préférence qui, selon Mill, allait à l'encontre de leur liberté comme de leur bien-être :

> *« Personne ne désapprouve plus profondément que moi cette institution mormone [la polygamie], pour ces mêmes raisons et aussi parce que, loin d'être comprise dans le principe de liberté, elle constitue une infraction directe à celui-ci, rivant simplement les chaînes d'une moitié de la communauté, et dispensant l'autre moitié de toute réciprocité d'obligation envers la première. Quoi qu'il en soit, il faut rappeler que, de la part des femmes concernées qui en paraissent les victimes, cette relation est tout aussi volontaire que dans toute autre forme d'institution matrimoniale. Et aussi surprenant que ce fait puisse paraître, il s'explique par les idées communes et les habitudes du monde : on apprend aux femmes que le mariage est la seule chose nécessaire pour elles ; ce qui explique dès lors que beaucoup d'entre elles préfèrent épouser un homme qui a beaucoup d'autres femmes à ne pas se marier du tout. »*

Cependant, cette définition des préférences adaptatives inclut des processus de formation ou de changement des préférences qui, intuitivement, ne semblent pas préjudiciables pour les personnes concernées. Adapter ses préférences aux possibilités existantes peut être tout à fait raisonnable : Nussbaum rappelle ainsi que ceux qui, enfants, ont rêvé de voler comme des oiseaux, ont bien fait de renoncer à ce souhait[320] [321]. Nous avons besoin de critères permettant de distinguer les préférences adaptatives des préférences décrivant de manière fiable le bien-être des agents. Et, tout d'abord, qu'entend-on exactement par « ajuster ses aspirations à ses possibilités » ?

[320]NUSSBAUM, Martha, *Women and human Development: the capabilities Approach*, Cambridge: Cambridge University Press, 1999. *Femmes et développement humain,* traduit de l'anglais (USA) par Camille Chaplain, Paris, Edition Des Femmes-Antoinette Fouque, 2008, p 201.

[321] Pour une série d'exemples de préférences adaptatives qui, intuitivement, semblent « bonnes », cf. BRUCKNER, Donald W., "In defense of adaptive preferences", *Philos Stud,* 142, 2007, pp 307-324.

Le faisable et le possible

Les préférences adaptatives ne sont pas une source d'information fiable sur le bien-être de l'agent si l'ensemble des options perçues comme faisables par l'agent est plus restreint qu'il ne l'est réellement. La conception du bien-être comme satisfaction des préférences stipule que le bien-être de l'agent correspond à l'état social que cette dernière préfère. Afin que l'information sur les préférences donnée par le jugement de l'agent soit fiable, il est crucial que cet agent ait une perception adéquate des possibilités qui s'offrent à lui[322]. Mais les analyses des préférences adaptatives ont tendance à se focaliser sur les motivations et les désirs de l'agent, faisant comme si les questions de faisabilité ou de réalisabilité n'étaient pas ce qui est en jeu. Or, pour diagnostiquer une préférence adaptative, il faut commencer par identifier les frontières de l'environnement de choix « réalisable »[323].

La préférence adaptative est un ajustement des préférences aux « possibilités ». Ces « possibilités » peuvent être comprises en trois sens différents :

(1) Peut être possible ce qui est perçu comme faisable, étant données les croyances actuelles de l'agent[324]. Dans l'exemple des femmes mormones, l'environnement de choix « perçu comme faisable » inclurait l'union polygame et le célibat, ce dernier étant mal vu dans leur milieu social. Les femmes mormones préfèreraient donc « épouser un homme qui a beaucoup d'autres femmes à ne pas se marier du tout ».

(2) Peut être défini comme possible ce qui est réellement faisable, y compris si l'agent ne le sait pas, ou pas encore. Ainsi, si les femmes mormones avaient davantage de moyens pour négocier les modalités de la vie conjugale sur un pied d'égalité, elles pourraient envisager d'inclure dans l'ensemble d'options la possibilité de ne consentir qu'à des unions monogames, ou encore celle d'un célibat qui ne soit pas dévalorisé[325].

---

[322] La restriction perçue des options peut influencer les préférences, mais également la perception du contenu de ces options. Elster (*op. cit.*,p 225) cite l'exemple suivant : « Si je n'obtiens pas la promotion que j'espérais, alors je peux rationaliser cet échec soit en disant que « mes supérieurs ont peur de mes capacités » (défaut de perception de la situation) , soit que « l'emploi supérieur ne vaut pas la peine de toute manière » (défaut de formation de la préférence). ». La première manière de rationaliser l'échec ne nous semble pas porter préjudice à l'agent, et ne concerne pas l'énoncé de ses préférences (à moins que l'agent ne juge préférable de ne pas obtenir la promotion désirée *parce que* ses supérieurs auraient peur de ses capacités).

[323] Nous adoptons cette perspective parce que notre analyse des préférences conçues indépendamment des expériences de plaisir a montré, dans les sections ci-dessus, que ce type de préférence dépend en fait étroitement de la connaissance des états sociaux par l'agent. Il est donc crucial de déterminer dans quelle mesure l'agent jouit d'une connaissance appropriée des moyens de réalisation des états sociaux si l'on veut décider d'une préférence si elle est adaptative ou non.

[324] Dans le cas de l' « attrait du fruit défendu », ou encore des préférences contre-adaptatives, l'agent préfère justement ce qu'elle ne peut pas obtenir. Dans ce cas, l'ensemble des options perçues comme faisables excède celui des options réellement faisables.

[325] Cette émancipation dépend de l'éducation, de l'acquisition du sens de sa propre valeur à travers la reconnaissance sociale (dans des rôles autres que ceux d'épouse et de mère), et également de l'association des personnes opprimées entre elles.

(3) Enfin, peut être conçu comme possible ce qui l'est logiquement, mais n'est pas à court terme à la portée des agents et ne peut donc être qualifié de « faisable ». Au moment où Mill écrit sur la polygamie chez les Mormons, il n'existe pas aux Etats-Unis d'interdiction légale de celle-ci : les Mormons sont exceptionnellement dispensés de suivre la règle de la monogamie. Mill se prononce d'ailleurs contre une telle interdiction. Fait donc partie de l'ensemble d'options possibles mais actuellement non faisables le choix alternatif pour les femmes mormones d'avoir recours à une loi les protégeant de ces unions et d'en appeler au juge. L'étendue des options possibles sans être faisables est vaste : elle n'exclut en fait que des alternatives impossibles telles que « être immortel », « voler comme un oiseau »[326]…

Les préférences adaptatives témoignent d'un écart entre la faisabilité perçue et la faisabilité réelle.

[326] Si l'on désire être exhaustif, on peut souligner que les progrès technologiques repoussent les limites du possible, en incluant des options inenvisageables jusqu'à il y a peu, telles que « aller sur la Lune », « cloner un être humain »…

Le processus d'ajustement

Le processus d'ajustement qui caractérise les préférences adaptatives diffère d'autres processus de changement ou de formation des préférences. Ainsi, Jon Elster distingue la formation des préférences adaptatives de l'apprentissage. Dans notre vocabulaire, l'apprentissage correspond à une amélioration de la perception de la faisabilité, tandis que l'adaptation repose au contraire sur une perception faussée. D'autre part, une préférence adaptative est construite en réponse à la perception restreinte de la faisabilité, et non au choix délibéré de restreindre cette faisabilité : par exemple, en prenant des engagements ou en respectant une promesse ; une personne qui s'engage à respecter un contrat sait qu'elle peut le rompre, mais se refuse à le faire non parce qu'elle croit que c'est impossible, mais parce qu'elle a choisi de ne pas le faire. Les préférences adaptatives diffèrent d'un changement positif de valeurs, correspondant à une étape de la « construction de la personnalité » (*character planning*) dans la terminologie d'Elster. Ainsi, on peut cesser de valoriser une carrière ambitieuse mais inaccessible, non pas en faisant comme si tout à coup cette carrière n'avait plus de valeur pour nous, mais en valorisant positivement d'autres activités, tel qu'apprendre à jouer d'un instrument de musique ou voyager, activités qui n'ont été possibles que parce que l'on a renoncé à ladite carrière.

Quel critère permet d'exclure de la base d'information les préférences adaptatives tout en retenant des phénomènes comme l'apprentissage, les engagements ou la construction de la personnalité ?

Dans l'article intitulé « Sour Grapes and Character Planning », le philosophe Luc Bovens propose un critère de rationalité. Selon lui, les préférences adaptatives sont irrationnelles, contrairement aux autres phénomènes de formation ou de changement de préférence, notamment ceux relatifs au *character planning* et à la construction de la personnalité. Est rationnelle une préférence qui résulte d'un processus d'arbitrage des différents critères de jugement. L'apprentissage, les engagements, la construction de la personnalité et la modification réfléchie du caractère résultent d'une réévaluation du poids des raisons et d'un changement de priorités. En revanche, une préférence adaptative est irrationnelle : par conséquent, elle ne résulte pas d'un réexamen des raisons ; ces dernières restent inchangées en dépit de la modification de la préférence.

A quoi renvoient ces « raisons » ? De la réponse à cette question dépend la pertinence du critère proposé par Bovens. Dans l'exemple des femmes mormones, nous pourrions supposer que Mary, l'une d'elle s'est, dans son adolescence, promis de ne jamais contracter une union

polygame. Mais, peu à peu, elle constate que toutes ses amies se sont mariées et qu'une vie solitaire et pauvre en activités l'attend. Elle réévalue donc ses raisons de refuser la polygamie et accepte finalement de devenir la troisième épouse d'un membre de la communauté. Il s'agit d'une préférence adaptative, puisque Mary a ajusté ses aspirations à ses possibilités. Intuitivement, cette préférence n'est pas fiable pour l'évaluation sociale, le choix collectif ou les théories de la justice. Et, cependant, d'après le critère proposé par Bovens, elle est tout à fait rationnelle. Ne peut-on pas concevoir un critère plus satisfaisant ?

Martha Nussbaum propose quant à elle « une théorie substantielle de la justice et des biens essentiels » afin de diagnostiquer les préférences adaptatives préjudiciables pour les agents[327].Elle oppose à une approche fondée sur les préférences une approche normative basée sur le fonctionnement et la capabilité, c'est-à-dire sur des réalisations qui témoignent de la condition objective de l'agent, que les préférences de celle-ci décrivent ou non correctement cette condition. Ce critère exclut donc non seulement les préférences adaptatives, mais en fait la notion même de préférence, que Nussbaum remplace par celle de « désir éclairé ». Le critère de Nussbaum ne distingue donc pas les préférences non-adaptatives des préférences adaptatives, puisqu'il exclut d'emblée les préférences. Nous discuterons en détail la proposition de Nussbaum dans la section 7.1.3.. Jusqu'ici, nous tentons de « sauver les préférences » afin de les concilier aux capabilités de Sen ; nous ne nous tournerons donc vers la proposition de Nussbaum que si la cause des préférences s'avère désespérée.

Reste le critère proposé par Elster lui-même : est adaptative une préférence qui n'a pas été formée de manière autonome par l'agent. Elster met en relation la liberté et la satisfaction des préférences de deux manières, qui correspondent aux deux étapes que sont la formation et la satisfaction des préférences. La liberté ne consiste pas seulement à satisfaire toutes les préférences que l'agent pourrait avoir ; encore faut-il que ces préférences aient été générées de manière autonome, en d'autres termes que l'agent ait préféré avoir ces préférences plutôt que d'autres. Néanmoins, de son propre aveu, Elster ne peut offrir de critère permettant de distinguer les désirs autonomes des désirs hétéronomes. Nous avons donc un problème de régression à l'infini : l'agent devrait à nouveau avoir des préférences autonomes pour ses préférences pour ses préférences…

[327]NUSSBAUM, Martha, *Women and human Development: the capabilities Approach*, Cambridge: Cambridge University Press, 1999. *Femmes et développement humain*, traduit de l'anglais (USA) par Camille Chaplain, Paris, Edition Des Femmes-Antoinette Fouque, 2008, p 203.

Une solution : considérer les préférences irréalisables

Elster propose donc une autre formulation du critère d'autonomie de formation des préférences. Cette formulation alternative est intéressante en ce qu'elle ne repose pas sur une analyse des motivations ou des intentions de l'agent, mais sur les effets de la modification de l'environnement de choix sur les préférences.

> *« Une preuve alternative que mon désir* [want] *de faire* x *n'est pas modelé par le manque d'options serait que je ne suis pas libre de faire* x. »[328]

Si nous reprenons notre catégorisation des environnements de choix, il est, d'après cette preuve, un cas dans lequel nous pouvons être certains de ne pas avoir affaire à des préférences adaptatives : c'est lorsque les préférences répondent non pas à un environnement de choix perçu comme faisable, ni même réellement faisable, mais à l'environnement de choix plus large encore qui inclut toutes les options *possibles.*

> *« […] (i) plus l'ensemble faisable est large et (ii) plus vos désirs vont au-delà de celui-ci, plus la probabilité que vos désirs sont modelés par cet ensemble est petite. »*[329]

Les préférences autonomes ne dépendent pas des variations d'étendue de l'environnement de choix, que ces variations soient réelles ou perçues. Par conséquent, l'état social classé au sommet de l'ordre de préférence devrait conserver sa place, y compris lorsqu'il ne fait pas partie de l'environnement de choix faisable, ou perçu comme faisable. Elster donne l'exemple d'un choix entre les trois options suivantes : être un citoyen libre (*x*), être un prisonnier de camp de concentration (*y*), et être un gardien de camp de concentration (*z*). Lorsque nous sommes libres, nous préférons en principe d'abord être libres, et nous ne voudrions sans doute pas jouer le rôle de tortionnaire associé à la fonction de gardien : notre ordre de préférence est donc (*x, y, z*). Mais le prisonnier des camps qui endure des privations et des tortures terribles préférera probablement *z* à *y*. La restriction de ses opportunités correspond donc à un remaniement de l'ordre des préférences, et ce remaniement est un symptôme d'adaptation et donc d'absence d'autonomie. Elster formule donc une condition d'autonomie des préférences :

> *« Si* S_1 *et* S_2 *sont deux ensembles faisables, entraînant des structures de préférences* R_1 *et* R_2*, alors, ni pour x ni pour y (appartenant toutes deux à l'ensemble global), il ne devrait être le cas que* xP_1y *et* yP_2x. » [330]

[328] ELSTER, *op. cit.*,p 228. Traduction libre..
[329] *Ibidem,* p 229.
[330] *Ibidem.*

Cette condition d'autonomie montre que l'ordre de préférence formé de manière autonome devrait être construit sur un « ensemble global » d'options, correspondant dans notre terminologie à l'ensemble « possible ». Le critère de distinction entre préférences adaptatives et préférences non-adaptatives se situe donc non pas du côté des états mentaux, des motivations ou des intentions, mais de celui de la définition de l'environnement de choix.

Dans un autre contexte, Marc Fleurbaey a également défendu la prise en compte des préférences sur des environnements de choix plus étendus que le seul ensemble faisable comme base d'information pertinente pour l'évaluation sociale[331]. Le problème est que l'une des conditions de la règle de choix collectif dans le théorème d'impossibilité d'Arrow exclut les préférences construites sur un ensemble d'options qui excède l'environnement de choix actuellement faisable et proposé à l'évaluation sociale. Il s'agit de la condition d'Indépendance des Alternatives Non Pertinentes, dont nous rappelons ici la formulation :

> *« Soit R et R' les relations de préférence sociales binaires déterminées par la règle de choix collectif f correspondant respectivement à deux ensembles de préférences individuelles, (R_1,....,R_n) et (R'_1,…..,R'_n). Si, pour toutes les paires d'options x, y dans un sous-ensemble S de X, x R_i y ↔ x R'_i y, pour tout i, alors les fonctions de choix C(S, R) et C(S, R') sont les mêmes. »*[332]

En des termes plus intuitifs, cette condition stipule que l'ordre de préférence collective énoncée sur un ensemble d'options données comme faisables doit dépendre exclusivement de l'ordre des préférences de chacun des membres de la collectivité sur ce même ensemble d'option. Sen a vu que cette condition exclut les comparaisons interpersonnelles, puisque la préférence collective ne dépend que d'une base d'information constituée d'ordres de préférences. Mais, comme le souligne Fleurbaey, cette condition exclut également les préférences sur les options situées hors de l'environnement de choix ; or, une telle information peut être pertinente du point de vue normatif.

Fleurbaey utilise en effet la possibilité de relaxer la condition d'indépendance afin d'inclure les préférences des agents pour les ressources des autres personnes, y compris lorsque ces ressources sont hors de leur portée. De telles préférences sont techniquement caractérisées comme des cas d' « envie »[333]. Cela permet donc d'introduire des considérations d'équité

---

[331] FLEURBAEY, Marc, "Social Choice and Just Institutions", *Economics and Philosophy*, 23, 2007, p. 15-43.

[332] Il s'agit de la formulation de Sen :SEN, Amartya, *Collective Choice and Social Welfare*, San Francisco, Holden day, 1970, p 41.

[333] Dans ce contexte, l'envie est simplement la préférence pour le lot d'une autre personne, du fait de la valeur de ce lot, et non pas en raison de la psychologie particulière de l'agent. La présence d'envie signale une forme d'inégalité économique et son diagnostic ne relève en aucun cas du jugement moral ou du blâme.

dans le cadre conceptuel de la théorie du choix social. Le propos de Fleurbaey n'est donc pas de contrer les phénomènes d'adaptation des préférences, mais de proposer une base informationnelle viable pour transposer des théories de la justice basées sur des préférences pour des ressources, telles que celle utilisées dans les théories de Rawls ou de Dworkin, dans un cadre axiomatique afin de spécifier des allocations optimales dans un contexte de mise en œuvre de politiques publiques caractérisé par des contraintes diverses comme la limitation de l'information ou les problèmes d'incitation [*second-best context*].

Fleurbaey souligne l'intérêt d'une solution au problème du choix social qui ne fasse pas appel à une notion d'utilité définie en dehors du cadre conceptuel de base. En effet, cette solution est remarquable d'économie, puisqu'elle ne mobilise pas d'autres concepts que ceux que la théorie du choix social utilise déjà : il n'est pas question de plaisir, d'expérience, de désir, etc. De la même manière, nous avons proposé un critère permettant de distinguer les préférences non-adaptatives des préférences adaptatives en nous fondant en premier lieu sur la structure même des préférences. Une préférence dont le contenu est l'évaluation de tous les états sociaux *possibles,* et pas seulement des états faisables ou perçus comme faisables ne sera pas adaptative dans la mesure où il ne s'agit pas d'une réaction à la perception des limites de la faisabilité. Le phénomène d'adaptation ne peut avoir lieu que lorsque les considérations de faisabilité sont prises en compte dans l'élaboration des préférences. Ce qui pose problème, ce n'est donc pas la préférence en tant que telle, mais les considérations de faisabilité[334].

En conclusion, la rupture entre la conception hédoniste du bien-être et celle correspondant à la satisfaction des préférences apparaît comme majeure. Il ne s'agit pas exclusivement d'un problème de représentabilité des états mentaux ; il s'agit d'un changement de perspective, du passage d'une approche mentaliste, axée sur les motivations subjectives de l'agent, à une approche réaliste, dans laquelle la préférence dépend de la connaissance d'un environnement de choix extérieur à l'agent, d'un monde d'options que celui-ci découvre à travers la délibération et la construction des préférences. Pour que l'agent puisse donc se reconnaître dans ces dernières, un long processus d'apprentissage, de réflexion, de comparaisons et de dissipation des illusions est en fait nécessaire. Nous sommes bien loin d'une vision des préférences les assimilant à des caprices.

[334] Nous ne voulons pas dire ici que les préférences prises en compte dans le choix social, qu'il s'agisse des votes ou des intérêts, doivent viser des fins irréalisables. Une société qui ne poursuivrait que l'immortalité n'est pas souhaitable. Mais le fait de considérer les préférences indépendamment de la question de la faisabilité, quitte à réintroduire celle-ci *ensuite*, permet d'une part d'éviter que les citoyens ne limitent leurs revendications sous prétexte que celles-ci ne seraient pas « raisonnables », et d'autre part d'enrichir la base informationnelle en évitant que celle-ci ne soit biaisée par les considérations de faisabilité.

Mais est-ce à dire que la préférence peut décrire le « bien objectif » de l'individu ? La connaissance des états sociaux est-elle suffisante ? Ne devrait-on pas imposer des conditions de rationalité supplémentaires à l'agent ?

5.2.3. Préférences rationnelles et objectivité du bien

Certaines préférences actuelles semblent aller à l'encontre du bien de l'agent, y compris lorsque celle-ci connaît parfaitement les états sociaux[335].Lorsqu'on reconnaît que les préférences actuelles ne peuvent pas décrire le bien de l'agent, alors on souscrit à une conception objective du bien-être ou du bien. Dans les lignes qui suivent, nous ne distingueront pas le « bien-être objectif » du « bien objectif » de l'agent. Certes, le terme « bien-être » connote fortement la subjectivité et la sensation dans la langue française : parler de « bien-être objectif » relève quasiment de l'oxymore. Mais les théories du bien-être dont le contenu consiste en une liste de biens[336] font appel à des justifications qui dépassent à la fois la recherche d'une expérience agréable ou les conceptions subjectives et arbitraires des individus.

Nous expliquerons dans un premier temps en quel sens les préférences de l'économie normative et de la théorie du choix social peuvent décrire le bien objectif (a). Nous étudierons ensuite deux compréhensions distinctes du rapport entre préférence rationnelle et bien objectif. La première compréhension stipule que la préférence rationnelle *révèle* le bien (b). La seconde compréhension conçoit le bien comme *déterminé* par la préférence rationnelle (c).

[335]GRIFFIN, J.P., *Well-Being*, Oxford: Clarendon Press, 1986.

[336] Parfit et Griffin parlent d' *Objective List Theories.*

a) Des rapports entre bien objectif et préférences

A quelles conditions une préférence peut-elle décrire le bien objectif de l'agent ? Il semble que ce soit à la condition d'être « rationnelle », en un sens qu'il nous faudra expliciter. Mais les défenseurs d'une conception objective du bien tendent à penser que toute préférence énoncée n'est pas rationnelle. Autrement, ils ne verraient pas de différence entre leur approche et l'approche qui identifie le bien à la satisfaction des préférences, et que nous avons analysée dans la section précédente. Sera donc prise en compte non pas la préférence actuelle de l'agent, mais une préférence hypothétique énoncée dans certaines conditions d'information, de raisonnement et de disposition affective jugées adéquates.

Dans la section précédente, il est apparu que les préférences actuelles requièrent une information suffisante pour faire en sorte que la réalisation de l'état social préféré corresponde effectivement aux finalités de l'agent. La conception du bien-être ou du bien comme réalisation des préférences rationnelles ajoute des conditions supplémentaires : sera rationnelle la préférence qui, non seulement témoignera d'une parfaite connaissance des états sociaux, mais sera également le produit d'un processus de réflexion qu'il reste à spécifier.

Les préférences rationnelles seraient alors des préférences hypothétiques. Nous avons admis qu'elles décrivent le bien-être objectif de l'agent. La distinction entre bien-être (*well-being*) objectif et bien-être subjectif élaborée par Scanlon constitue un bon point de départ :

> *« Par* critère subjectif *[de bien-être], je veux dire un critère selon lequel le niveau de bien-être dont jouit une personne dans des circonstances matérielles données ou l'importance pour cette personne d'un bénéfice donné ou d'un sacrifice doit être estimée en évaluant ces circonstances matérielles, ce bénéfice ou ce sacrifice uniquement du point de vue des goûts et des intérêts de cette personne. »*
>
> *« Par* critère objectif, *je veux dire un critère qui fournit une base pour l'évaluation du niveau de bien-être d'une personne, base indépendante des goûts et des intérêts de cette personne, ce qui rend ainsi possible qu'une telle évaluation puisse être correcte même si elle s'oppose aux préférences de l'individu en question, non seulement telles qu'il croit qu'elles sont, mais même comme elles seraient si elles étaient corrigées afin d'être cohérentes, expurgées d'erreurs factuelles, etc. »*[337]

Le caractère objectif du bien-être est donc d'abord défini négativement, par son indépendance à l'égard de la préférence. Néanmoins, Scanlon reconnaît que les liens entre préférence et bien objectif sont plus complexes :

[337]SCANLON, Thomas, « Preference and Urgency », *Journal of Philosophy* 72 (19), 1975, pp 655-669. Traduction libre.

> *« Une valeur objective élevée peut être donnée au fait d'offrir les conditions qui sont nécessaires pour permettre aux individus de développer leurs propres préférences et intérêts et de faire en sorte que celles-ci soient prises en compte dans la détermination de la politique sociale. »*[338]

La capacité d'énoncer des préférences peut faire partie d'une liste de biens objectifs. Mais ce n'est pas tout : la distinction de Scanlon repose sur une conception de la préférence beaucoup plus restrictive que celle que nous défendons dans ce travail, comme en témoigne l'utilisation des termes « goûts » et « intérêts » pour qualifier celles-ci. Or, en nous référant en premier lieu au cadre conceptuel de la théorie du choix social et des fonctions de bien-être collectif, nous avons porté notre attention sur une étendue d'interprétations de la « préférence » beaucoup plus vaste. Notre approche de la préférence est donc structurelle et non pas substantielle. On peut dès lors se demander si les jugements de valeur déterminés par une conception objective du bien-être ne sont pas susceptibles d'être représentés sous la forme de classements de préférences. Cela a de fait déjà été proposé : par exemple, Sen représente des évaluations de l'accomplissement des fonctionnements d'une personne par la relation d'ordre de préférences P[339]. Or, les fonctionnements sont des états ou des activités pouvant être décrits indépendamment de la perception de l'agent qui réalise ces fonctionnements. Comme le soulignent Grüne-Yanoff et Hansson dans une introduction sur les préférences :

> *« Dans les usages les plus courants en sciences sociales, les jugements de préférence représentent des jugements subjectifs. Cependant, une interprétation alternative en termes de bien comparé objectif* [objective betterness] *est compatible avec la structure. »*[340]

Cette structure comparative de la relation de préférence est utilisée par Broome pour son analyse du bien[341]. Par « bien », Broome ne fait pas référence à une théorie morale substantielle et il n'assigne pas de contenu à cette notion. Il s'intéresse plutôt à la structure des jugements de valeur comparatifs établis selon le critère du meilleur. Broome fait donc une analyse de jugements de valeur ayant la structure suivante :

' — est meilleur que — '

Néanmoins, Broome affirme que ces jugements de valeur selon le « meilleur » ne correspondent pas à des préférences, car lui aussi prend acte du fait que, parfois, les préférences actuelles d'une personne semblent contredire son bien objectif. En revanche,

---

[338] *Ibidem.*

[339] SEN, Amartya ,*Commodities and Capabilities,* Amsterdam: North-Holland, 1985, chap. V.

[340] GRÜNE-YANOFF, Till, et HANSSON, Sven Ove, *Preference Change. Approaches from philosophy, economics and psychology*, Springer, 2009, p 8. Traduction libre.

[341] BROOME, John, *Weighing Goods*, Oxford : Blackwell, 1991.

Broome concède qu'une théorie du bien-être comme satisfaction des préférences hypothétiques *rationnelles* pourrait palier les lacunes de la préférence actuelle lorsque celle-ci doit décrire le bien de l'agent[342]. Cette théorie pourrait même fonder les jugements comparatifs selon le critère du meilleur. Broome affirme donc que le « meilleur » est plus général que le « préférable », mais il n'exclut pas la possibilité que des préférences rationnelles puissent en fait décrire le bien objectif et donc coïncider systématiquement avec les évaluations selon le meilleur.

Lorsque nous disons que la « préférence rationnelle » de l'agent – individuel ou collectif – décrit son bien-être objectif ou même le bien objectif, cette affirmation peut être comprise en deux sens différents, dont chacun a des implications normatives susceptibles de diverger fortement.

La première interprétation de cette affirmation consiste à considérer que l'objectivité d'une conception du bien est garantie par son indépendance à l'égard des préférences. Le bien objectif pour un agent pourrait être identifié en l'absence de toute énonciation de préférence de sa part.

Mais cette affirmation peut également sous-tendre une conception de l'objectivité qui consisterait dans les conditions d'élaboration de la « préférence », ou encore dans le caractère rationnel de celle-ci. Est objectif ce qui est l'objet d'une préférence « rationnelle ». Ici, il n'y a pas d'identification possible du bien en l'absence de ladite préférence, préférence dont les conditions de rationalité restent à déterminer.

Chacune de ces conceptions mérite ainsi une analyse plus fine, à la lumière de notre problématique générale, qui est de déterminer s'il existe une conception de la préférence compatible avec le cadre de la théorie du choix social et les capabilités.

[342] *Ibidem*, p 133.

b) L'indépendance de la conception objective du bien à l'égard de la préférence

Le bien « objectif » des personnes peut être objectif au sens où il ne dépend pas de leur subjectivité. Dans ce cas, la coïncidence entre le contenu de la préférence et le bien objectif est fortuite. Le bien objectif a de la valeur par lui-même, indépendamment du fait d'être préféré. Lorsque la préférence en question satisfait certaines conditions de rationalité, elle est plus à même de décrire ce bien objectif. Selon qu'on définit la « rationalité » de la préférence en accord avec cette conception objective du bien, ou non, la préférence rationnelle décrira nécessairement ce bien, ou non. Si, par «préférence rationnelle », on entend exclusivement une préférence déterminée par la recherche du bien objectif, alors toute préférence rationnelle décrit nécessairement le bien objectif de l'agent. Si, en revanche, par « préférence rationnelle », on désigne simplement certaines conditions de délibération, telles que le fait de réfléchir au calme ou celui d'envisager les perspectives à long terme, alors la préférence rationnelle sera plus à même de décrire le bien objectif de l'agent, mais ce ne sera pas une vérité nécessaire[343].

[343] Nous devons cette clarification à John Broome, *op. cit.*,p 133.

Du choix collectif au bien commun, la préférence comme révélation du bien objectif

Si ce n'est la préférence qui fait qu'un état ou un objet est bon, qu'est-ce que c'est ? Le bien objectif doit découler de principes qui ne dépendent pas de la préférence, y compris dans son interprétation large. Cette première interprétation de la relation entre préférence rationnelle et bien objectif suppose que le bien existe indépendamment de nos jugements ou de nos préférences. En ce sens, lorsque la préférence de l'agent décrit son bien objectif, elle ne le *détermine* pas, mais le *révèle*. Il est concevable que le bien objectif puisse subsister sans qu'aucune préférence, aucun jugement de valeur ne vienne le décrire[344].

Quels seraient les fondements d'une telle conception objective ? Les réponses possibles à cette question sont multiples, mais elles peuvent être classées en deux grandes familles, la première comprenant les réponses naturalistes et la seconde les réponses conventionnalistes[345].

La réponse naturaliste établit une conception du bien dérivée de la nature même de l'être dont il est question. Du fait de sa nature, l'être humain devrait rechercher un bien qui lui est propre, et le contenu de la liste de biens objectifs qu'il recherche est déterminé par ce « bien propre ». Cette conception du bien est donc déterminée indépendamment de toute activité évaluative effectuée par les individus ou les collectivités particulières. En revanche, une activité cognitive peut être mise en œuvre afin de mieux connaître cette conception du bien et de découvrir le contenu de la liste de biens objectifs constitutive du bien humain.

La réponse naturaliste définit un bien universel car déterminé par la nature humaine : ce bien est universel au sens où il s'agit du bien pour tous les êtres appartenant au genre humain. En revanche, la réponse conventionnaliste l'est moins : il lui suffit que la conception du bien soit acceptée et établie dans un contexte social ou historique donné.

Si les préférences décrivent un bien objectif, alors la signification même du processus de choix collectif s'en trouve profondément modifiée. En effet, lorsque les préférences individuelles décrivent un bien subjectif, le choix collectif est déterminé par ces préférences.

[344] Il est toutefois possible que la satisfaction des désirs de la personne soit intégrée à une liste de biens objectifs. Mais cela aurait des implications problématiques : que faire lorsque la satisfaction d'un de ses désirs empêche l'agent d'accéder à un autre bien objectif ? Cf. ARNESON, Richard, "Desire Formation and Human Good", in OLSARETTI, Serena (ed.), *Preferences and Well-Being*, Cambridge University Press, 2006.

[345] Cf. SCANLON, *op. cit.*,qui utilise cette distinction afin de caractériser les justifications morales de l'appel à la notion d'urgence, justifications qui ne reposent ni sur la force des désirs, ni sur un accord unanime.

En revanche, si elles décrivent un bien objectif, celui-ci existe indépendamment d'elles. Le processus de choix collectif a dès lors pour fonction de révéler ce bien.

C'est en ce sens que l'on peut comprendre le rôle des élections et du vote dans le *Contrat Social* de Rousseau : ce n'est pas par n'importe quel vote que les citoyens accèdent à la volonté générale. Certains votes, au contraire, manquent cette dernière.

> *« […] la volonté générale est toujours droite et tend toujours à l'utilité publique : mais […] les délibérations du peuple [n'ont pas] toujours la même rectitude. On veut toujours son bien, mais on ne le voit pas toujours. »*[346]

On ne peut, dit Rousseau, soupçonner le peuple de ne pas vouloir ce que veut la volonté générale, c'est-à-dire l'intérêt de la collectivité et ce qui la maintient en vie. Et cependant, les opinions qu'expriment les citoyens à travers leurs votes, les « préférences » qu'ils révèlent en somme, ne correspondent pas toujours à l'objet par définition de cette volonté générale. Ce n'est pas parce que la volonté du peuple ne correspond pas à la volonté que celle-ci se trompe : ce serait une contradiction pour Rousseau, parce que la volonté générale est la volonté propre du peuple, et ne peut, par définition, que viser le bien commun, c'est-à-dire tout ce qui contribue à la protection de la personne et des biens de chaque citoyen, ainsi qu'à l'existence du corps politique par lequel la volonté générale peut s'exprimer. Si la volonté générale « erre », si les choix collectifs satisfont les intérêts privés et ne réalisent pas le bien commun, c'est parce que les conditions d'énonciation de la volonté générale ne sont pas bonnes. Des groupes d'intérêts font des alliances, et font passer leur volonté particulière pour la volonté générale ; ou parfois, les citoyens sont mal informés, et se trompent. Mais, malgré cela, la volonté générale subsiste :

> *« Enfin quand l'Etat près de sa ruine ne subsiste plus que par une forme illusoire et vaine, que le lien social est rompu dans tous les cœurs, que le plus vil intérêt se pare effrontément du nom sacré du bien public, alors la volonté générale devient muette ; tous guidés par des motifs secrets n'opinent pas plus comme citoyens que si l'Etat n'eût jamais existé ; et l'on fait passer uniquement sous le nom de lois des décrets iniques qui n'ont pour but que l'intérêt particulier.*
>
> *S'ensuit-il de là que la volonté générale soit anéantie ou corrompue ? Non, elle est toujours constante, inaltérable et pure ; mais elle est subordonnée à d'autres qui l'emportent sur elle. […] Ainsi, la loi de l'ordre public dans les assemblées n'est pas tant d'y maintenir la volonté générale que de faire qu'elle soit toujours interrogée et qu'elle réponde toujours. »*[347]

[346]ROUSSEAU, Jean-Jacques, *Du Contrat Social*, Livre II, Chapitre III.
[347]*Ibidem*, Livre IV, chapitre I.

Le fait que, selon Rousseau, la volonté générale subsiste, même lorsque le résultat des délibérations actuelles ne coïncide plus avec celle-ci, montre que le rapport entre choix collectif et volonté générale n'est pas un rapport de détermination de celle-ci par celui-là. On attend bel et bien ici du processus d'agrégation des « préférences »[348] de nous faire connaître, de nous révéler une volonté générale qui existe du fait même de l'existence d'un peuple souverain et libre, ou potentiellement libre, volonté qui ne peut, par sa nature même, que viser le bien commun. Le théorème d'impossibilité d'Arrow pourrait ainsi être réinterprété dans la perspective de Rousseau : pour Rousseau, l'impossibilité de dériver une préférence collective de l'agrégation des préférences individuelles serait due à un vice de procédure de la décision en assemblée. Celle-ci ne se serait pas déroulée de façon à générer une décision déterminée par la volonté générale, mais de façon à promouvoir certains intérêts particuliers. Autrement dit, une bonne procédure de décision collective serait une procédure limitant le domaine des préférences particulières admissibles de sorte que, d'une part, la décision collective soit possible, et d'autre part, elle corresponde au bien commun déterminé par la volonté générale[349].

La conception objective du bien discrimine donc les préférences selon leur propension à révéler ou non ce bien. Elle fournit un critère d'élimination des préférences qui ne décrivent pas le bien de la personne ou le bien général…Ce critère est relatif au contenu de la préférence ; les conditions d'énonciation de la préférence rationnelle sont ensuite définies en fonction de leur capacité à garantir l'accès à la conception objective du bien par la préférence. Ainsi, Rousseau propose de tenir des scrutins secrets afin d'éviter que la volonté générale ne soit accaparée par les alliances d'intérêt ; il souligne également que des citoyens bien informés sont plus à même d'émettre des opinions proches du bien commun que devrait viser par définition la volonté générale.

[348] Nous prenons quelques libertés avec le texte de Rousseau en le traduisant dans le vocabulaire technique de la théorie du choix social. Mais notre propos n'est pas de mettre en avant la spécificité de la pensée de Rousseau – ce qui ferait l'objet d'un travail de recherche propre - , mais plutôt d'en dégager une caractéristique commune aux approches de l'évaluation sociale fondées sur une conception objective du bien *révélée* par le choix collectif. D'autre part, nous espérons avoir suffisamment mis en avant la souplesse interprétative autorisée par le formalisme de la théorie du choix social pour pouvoir prendre la liberté de l'appliquer à la réflexion de Rousseau sur les délibérations et les décisions en assemblée.

[349] Il ne faut pas sous-estimer malgré cela le problème de l'impossibilité du choix social. La double exigence qu'imposerait la conception de la décision collective imposée par Rousseau aux procédures de vote en assemblée est très importante, et risque de ne pouvoir être satisfaite qu'exceptionnellement (et Rousseau semble conscient de cette difficulté).

Les limites de la description du bien objectif par la préférence

La conception objective du bien ne dépend pas des préférences. Mais on peut se demander si les préférences telles que les représente et les conçoit la théorie du choix social sont l'outil le plus approprié pour dévoiler la liste de biens objectivement bons. Rappelons que ces préférences sont caractérisées par une structure comparative. Les préférences ordinales consistent à comparer par paire des états sociaux afin de les classer selon des critères évaluatifs. Si ces préférences ordinales sont supposées décrire le bien objectif des agents, alors elles classeront les états sociaux en fonction des biens objectifs qui s'y trouvent. Mais la représentation de jugements de valeur par un classement n'est pas particulièrement appropriée aux évaluations déterminées par une conception objective du bien, et ce, que cette conception corresponde à un bien unique ou à une liste de biens.

Supposons qu'il n'y ait qu'un bien. Si ce bien est quantifiable, il reste possible de classer les états sociaux en fonction de la quantité de ce bien dont on dispose. Si, par exemple, ce bien est la connaissance, alors les états sociaux seront classés en fonction de la quantité de connaissance et du degré d'éducation des agents. Mais que faire si ce bien est la justice ? Il n'existe pas d'intermédiaire entre le juste et l'injuste : on ne peut pas classer les états sociaux en fonction de leur caractère « plus ou moins juste » : soit ils sont justes, soit ils ne le sont pas. Certes, on peut « compter » le nombre d'injustices dans chaque état social. Mais il suffit d'une injustice pour qu'un état social puisse être qualifié d' « injuste ».

Supposons qu'il y ait plusieurs biens : la liste pourrait inclure la connaissance et la richesse matérielle, par exemple. Dans ce cas, les états sociaux devraient être classés en fonction de la quantité de connaissance et de richesse matérielle qu'ils garantissent aux personnes. Un état social apportant beaucoup des deux sera bien sûr idéalement reconnu meilleur qu'un état apportant peu ou pas de connaissance et de richesse matérielle. Mais comment déterminer la valeur respective d'un état apportant beaucoup de connaissance et peu de richesses et celle d'un état garantissant peu de connaissance mais beaucoup de richesses? La valeur intrinsèque de ces biens devrait à elle seule décider de leur position dans le classement de préférences. En somme, la conception objective du bien indépendamment de toute préférence requiert une théorie suffisamment riche des biens pour anticiper tous les cas de figure de comparaison et d'arbitrage qui caractérisent les activités d'évaluation et de décision dans le monde réel.

c) Le bien-être comme objet déterminé par une préférence rationnelle

La deuxième voie permettant de définir une conception objective du bien consiste à définir celle-ci comme ce qui serait l'objet d'une préférence rationnelle. Ici, la conception du bien n'est pas tout à fait indépendante des préférences. Mais le processus de détermination de la conception objective du bien impose des conditions de rationalité à la préférence.

Quelles conditions supplémentaires de la préférence permettent de la caractériser comme « rationnelle » ? La préférence rationnelle est sans doute une préférence informée. Toute préférence rationnelle est donc une préférence informée, mais toute préférence informée n'est pas nécessairement rationnelle. La rationalité de la préférence peut requérir des conditions supplémentaires qui sont autant de filtres des préférences.

Les conditions de rationalité de la préférence

Dans le cadre arrovien, la rationalité correspond à des propriétés formelles. La théorie du choix social et l'économie normative identifient la rationalité de la préférence aux propriétés de la relation d'ordre qui représente ladite préférence, à savoir la réflexivité, la transitivité et la complétude. En ce sens, une préférence serait intrinsèquement rationnelle. Mais une conception plus substantielle de la rationalité peut être requise si l'on attend de la préférence qu'elle décrive le bien objectif de l'agent et ne soit pas le simple produit de la réflexion d'un « idiot rationnel », pour reprendre l'expression de Sen[350]

Les filtres appliqués aux préférences permettent d'exclure les préférences irrationnelles. Deux types de filtres existent : le premier type de filtre s'applique aux motivations à l'œuvre dans la préférence et donc au contenu des raisons de préférer ; le second type de filtre concerne les procédures de raisonnement déterminant la préférence, c'est-à-dire la manière dont la préférence est produite. Schématiquement, le premier filtre garantit la rationalité du contenu des préférences, tandis que le second identifie une préférence rationnelle à une préférence dont le processus d'élaboration est rationnel.

[350] SEN, Amartya, *On Ethics and Economics*, Blackwell Publishers, Oxford, 1991; traduction française par Sophie Marnat : *Ethique et Economie*, Paris, Presses Universitaires de France, Quadrige, 1993, 2008.

La rationalité comme poursuite de l'intérêt personnel

Quelles sortes de motivations déterminent les préférences rationnelles ? Un candidat potentiel est l'intérêt personnel. Au XIXe siècle, Edgeworth fait de l'hypothèse selon laquelle un agent est mû par son propre intérêt un principe premier de la science économique[351]. Lorsque John Broome présente la théorie du bien comme « satisfaction des préférences déterminées par l'intérêt personnel », théorie selon laquelle un état social est bon pour une personne si la préférence de cette personne pour cet état social est bien informée, rationnelle et déterminée par l'intérêt personnel, il ajoute que « cette théorie n'est pas très loin de la vérité »[352].

Mais Derek Parfit a montré que la théorie selon laquelle la rationalité correspond à la poursuite de l'intérêt personnel est défectueuse. Son argumentation procède ainsi : il formule tout d'abord la thèse centrale de la théorie selon laquelle est rationnelle une préférence ou une action en vue de l'intérêt personnel, qu'il appelle S1 :

> *« (S1) Pour chaque personne, il y a un but ultime par excellence rationnel : ce but est que sa vie se déroule, pour elle, aussi bien que possible. »*[353]

Parfit soutient que S1 est intrinsèquement défectueuse [*self-defeating*]. Une théorie de la rationalité est défectueuse si, lorsqu'une personne essaie d'accomplir les buts définis par cette théorie, elle se trouve plus mal lotie qu'au départ, et ce sans avoir fait d'erreur au regard de la théorie en question. La théorie est défectueuse en raison de la disposition dans lesquelles les personnes devraient s'y trouver selon elle. Parfit propose une critique de S1 basée sur la possibilité du cas suivant : supposons, dit-il, que je sois parfaitement « transparente », et ne puisse mentir sans me trahir, en rougissant ou en modifiant le ton de ma voix. Je tombe en panne au milieu du désert. Un étranger passe en voiture, je l'arrête, et lui promets une récompense s'il me raccompagne chez moi. Il est dans mon intérêt qu'il accepte ; mais il est également dans mon intérêt de conserver la somme que je lui ai promise. En lui faisant cette promesse, je compte bien ne pas la tenir ; mais, étant incapable de mentir de façon convaincante, l'étranger comprend qu'il sera trompé et refuse de me raccompagner. La situation qui résulte est la pire pour moi, et cependant elle résulte d'actions entreprises en vue de la réalisation de mon intérêt personnel.

La critique parfitéenne de la théorie de l'intérêt personnel ne vise que la cohérence de cette théorie. Elle reste muette sur la signification de son contenu. Amartya Sen a adressé une

---

[351]EDGEWORTH, *Mathematical Psychics: an Essay on the Application of Mathematics to the Moral Sciences*, London: Kegan Paul, 1881, p 16.

[352]BROOME, John, *Weighing Goods*, Oxford : Blackwell, 1991, p 134. Traduction libre..

[353] PARFIT, Derek, *Reasons and Persons*, Oxford :Clarendon Press, 1984, chap. I, §1.Traduction libre.

critique plus fondamentale à cette théorie et notamment à son usage dans la science économique[354]. Il a mis en évidence l'existence de deux types de comportement qui ne correspondent pas à la description d'un choix dicté par l'intérêt personnel. Il s'agit de la compassion et de l'engagement. Tous deux sont déterminés par des motifs autres que la poursuite de l'intérêt personnel.

Compatir, étymologiquement, c'est « pâtir avec » : la compassion consiste à éprouver des affects similaires à ceux d'autrui, et à agir en conséquence. Lorsqu'une autre personne souffre, sa souffrance suscite en moi une souffrance similaire, et je souhaite donc annuler ou limiter cette souffrance. Dans l'exemple de Parfit, l'étranger agirait par compassion s'il acceptait de raccompagner le chauffeur malchanceux par souci du bien-être de celui-ci et non pas parce qu'il espérait une récompense.

Contrairement à la compassion, le point de départ de l'engagement n'est pas un affect. L'engagement fait d'abord appel à un discours : s'engager, c'est communiquer, à soi ou aux autres, une volonté d'agir selon certains principes. On peut s'engager envers soi-même à terminer une tâche, à ne pas mentir, etc. On peut s'engager envers autrui à tenir une promesse. Dans l'exemple de Parfit ; le chauffeur malchanceux tiendrait ses engagements s'il donnait effectivement la récompense promise une fois rentré chez lui.

Théoriquement, la compassion peut être réduite à la poursuite de l'intérêt personnel : si la vue d'autrui souffrant cause en moi une souffrance, alors c'est en vue de mon bien-être que je dois aider autrui. Néanmoins, cette interprétation présuppose une priorité du souci de son propre bien-être sur celui d'autrui ; or, cette priorité doit être démontrée. Autrement dit, cette lecture égoïste de la compassion n'est pas incohérente, mais elle n'est qu'une lecture possible et ne prouve rien. En ce qui concerne l'engagement, il est beaucoup plus difficile de le réduire à un comportement égoïste. Sen explique cette difficulté par

> *« le fait que l'engagement établit une distance entre choix personnel et bien-être personnel. »*[355]

L'engagement met entre parenthèse l'effet escompté des actions entreprises afin de le remplir sur la situation de l'agent. L'agent qui s'engage est indifférent aux effets qu'aura son acte sur son bien-être, tandis que l'agent qui compatit sait que son comportement altruiste lui procurera un certain bien-être. Or, l'engagement décrit des comportements qu'il serait *a priori* arbitraire de qualifier d' « irrationnels » : ces comportements sont fréquents, ils ne sont dans

[354] SEN, Amartya, « Rational Fools : A Critique of the Behavioural Foundations of Economic Theory", in *Choice, Welfare and Measurement*, Cambridge, MA: the MIT Press, 1982, p 84-106; traduction française par Sophie Marnat, « Des idiots rationnels », *Ethique et Economie*, Paris, Presses Universitaires de France, Quadrige, 1993, 2008

[355] *Ibidem.*

l'ensemble pas désapprouvés par le sens commun, et ils contribuent au bon fonctionnement de la société…

Définir une préférence rationnelle exclusivement comme une préférence déterminée par la poursuite de l'intérêt personnel n'est pas faux, mais restrictif. D'autre part, la définition de ce qui satisfait la poursuite de l'intérêt personnel dépend de la conception du bien que l'on défend. L'intérêt personnel n'est pas défini de la même manière d'après la conception hédoniste du bien-être que selon celle qui vise la satisfaction de toute préférence ou celle qui fait correspondre au bien-être une notion objective du bien. Dans la conception hédoniste, est dans l'intérêt de l'agent la réalisation d'une expérience de plaisir, ou la disparition d'une expérience de peine ; cette conception intègre donc les actes effectués par compassion. Selon la conception du bien-être comme satisfaction des préférences, l'intérêt de l'agent consiste à réaliser l'état social préféré par celui-ci, y compris lorsque cela ne suppose aucune expérience : dans ce cas, si l'agent est altruiste ou au contraire veut à tout prix nuire à une personne, sans que cela ne lui procure aucun plaisir, alors son « intérêt personnel » pourrait correspondre à des finalités qui ne sont ni égoïstes, ni relevant de la maximisation du plaisir. Selon la conception objective du bien, définie indépendamment de la préférence, l'intérêt de l'agent peut être décrit sans faire référence aux opinions, perceptions ou sensations particulières de celle-ci. Si, par exemple, on suppose que l'être humain est naturellement porté à vivre en société et en harmonie avec les siens, la définition de son intérêt selon la conception objective du bien ne correspond pas à une poursuite égoïste. Selon la conception du bien-être que l'on défend, l'intérêt personnel peut être exclusif et égoïste, mais pas nécessairement.

Identifier la préférence rationnelle à la préférence déterminée par l'intérêt personnel est donc incohérent – *dixit* Parfit –, contraire à nos intuitions de ce qui est rationnel et de ce qui ne l'est pas – d'après la critique de Sen –, et trop indéterminé du point de vue de la définition de l'intérêt. De fait, si la préférence est rationnelle parce qu'elle est déterminée par la poursuite de l'intérêt personnel, cela signifie qu'elle est rationnelle parce qu'elle réalise l'une des trois conceptions du bien-être que nous avons analysées, et donc que sa définition dépend de l'une de ces conceptions. Or, notre propos ici est au contraire de prouver que c'est la préférence rationnelle qui détermine le contenu du bien. Il est donc insatisfaisant d'associer la rationalité de la préférence à la poursuite de l'intérêt personnel.

Rationalité et blanchiment des préférences

Une raison de passer des préférences actuelles aux préférences rationnelles hypothétiques est que les premières sont parfois « antisociales » ou « perverses ». En témoignent les préférences intrusives du libertin et du prude que décrit Sen dans son conflit libéral-parétien[356]. En témoignent aussi les préférences sadiques du Docteur Fischer de Genève, personnage créé par Graham Greene qui prend plaisir à humilier ses riches invités[357]. En témoignent encore les préférences de la Cousine Bette. Ces figures littéraires évoquent des expériences vécues ou possibles.

Utilisant la métaphore de la lessive, le philosophe et politologue Robert Goodin propose de « blanchir » ces préférences. Des préférences « blanchies » seraient des préférences socialement acceptables, ne portant pas préjudice à autrui. Mais comment justifier la sentence qui qualifie les préférences « antisociales » ou « perverses » d'irrationnelles ? Si Goodin ne fait pas appel à la notion de rationalité, il semble toutefois considérer comme acquis que les préférences orientées par le désir de nuire à autrui sont intuitivement inacceptables. Si, par « rationnelles », sont désignées des préférences exclusivement déterminées par le plaisir ou la satisfaction des désirs de l'agent, aussi arbitraires soient-ils, alors ces préférences perverses seront bel et bien rationnelles. Mais Goodin aborde le problème d'un autre angle et considère le point de vue de l'intérêt collectif. Il y a, dit-il, une « logique interne à l'agrégation des préférences » :

> *« respecter les choix des personnes nous engage implicitement à respecter leur dignité, et cela implique certaines limitations claires des choix que nous pouvons être obligés de respecter »*[358]

Les préférences perverses sont collectivement irrationnelles parce qu'elles remettent en cause la possibilité même du choix social. Cela ne signifie certes pas que les personnes renoncent à ces préférences sitôt qu'elles s'engagent dans un processus de décision collective. Néanmoins, Goodin montre que le problème des préférences perverses est moins fréquent dans les faits que leur possibilité logique ne pouvait le laisser présager. Il cite deux cas à l'appui.

Le premier cas fait appel à des travaux en anthropologie. Des anthropologues ont en effet observé que les discussions précédant les votes dans des villages Thaï produisaient des

[356]SEN, Amartya, "The Impossibility of a Paretian Liberal", *Journal of Political Economy*, vol. 78, no 1, jan-mar 1970, pp 152-157.

[357] GREENE, Graham, *Dr Fischer of Geneva or The Bomb Party*, London: Penguin, 1992. Cet exemple est cité par Brian Barry.

[358]*Ibidem*, p 85. Traduction libre..

changements significatifs de préférences aboutissant à des votes davantage orientés vers l'intérêt de la communauté.

Le second cas met en évidence le décalage entre les comportements de vote et les « découvertes » de la théorie du choix rationnel. En effet, l'une de ces découvertes, le « paradoxe du vote », affirme qu'il n'est pas dans l'intérêt des citoyens de faire l'effort d'aller voter dans la mesure où leur vote n'a qu'une influence négligeable sur l'élection[359]. Mais ce paradoxe n'en est pas un si on considère qu'une préférence rationnelle peut être orientée vers l'intérêt de la collectivité. Or, le fait même de s'engager dans un processus de choix collectif et de vote présuppose de la part des citoyens un intérêt pour l'avenir de cette collectivité, en plus du leur[360].

Si les préférences antisociales sont collectivement irrationnelles, il est justifié de les exclure du processus de choix collectif visant à déterminer une préférence rationnelle pour la collectivité. Certes, une alternative consisterait à garantir des droits qui retirent les options préjudiciables à autrui de l'ensemble des alternatives soumises au choix collectif. Mais cette perspective des droits requiert une justification fondée sur le respect des personnes ; or, Goodin affirme que le respect des personnes passe par l'intérêt pour leurs attitudes et leurs préférences, plutôt que par l'annulation de la possibilité de voir une préférence réalisée. Plutôt que de filtrer « à la sortie » le produit des préférences [*output filter*], Goodin préconise donc de filtrer « à l'entrée » la production des préférences elle-même [*input filter*]. L'argument du respect peut être complété par deux autres arguments en faveur d'un filtre en amont : d'une part, ce type de filtre est plus durable, puisque, si les préférences changent, les personnes perdront l'intérêt d'essayer de transgresser les droits des autres en réalisant les états sociaux interdits en dépit de la loi – ces droits exigeant des politiques de garantie parfois coûteuses –; d'autre part, même un filtre en amont doit être mis en place au moyen d'une décision collective, à tout le moins dans une démocratie : autrement dit, l'établissement de droits garantis suppose d'ores et déjà un lessivage minimal des préférences.

Comment fonctionnerait ce filtre à l'entrée ? Goodin distingue deux étapes. En un premier temps, les citoyens eux-mêmes ignoreraient leurs préférences collectivement irrationnelles au

[359] Ce paradoxe était encore cité par le blog « Democracy in America », hébergé par *The Economist*, à la veille des élections américaines, en novembre 2012. Cf. *Democracy in Ameria*, "Is it irrational to vote?", billet publié le 23 octobre 2012:

http://www.economist.com/blogs/democracyinamerica/2012/10/presidential-election-0. Consulté le 25 octobre 2012

[360] De ce point de vue, le phénomène de l'abstention pourrait être compris comme une certaine négligence des intérêts de la collectivité. Il nous semble cependant que les choses sont plus compliquées.

moment de s'engager dans une procédure de vote ou d'évaluation collective. Ces phénomènes correspondent aux cas du village Thaï ou du paradoxe du vote. Mais cet « auto-lessivage » des préférences peut s'avérer insuffisant : Goodin propose donc que les décideurs remédient aux préférences perverses. Il admet que la démarche cela suppose un certain paternalisme. Mais le paternalisme n'est pas incompatible avec une certaine souveraineté des personnes sur leurs préférences et leurs opinions si, d'une part, il n'exclut pas toutes les préférences des personnes, mais seulement quelques-unes, et d'autre part il se justifie par la garantie de la prise en compte des préférences des autres personnes : celles et ceux qui subissent les effets des choix cruels des autres méritent également d'être entendus. Cependant, cette position est difficile à maintenir dans la mesure où le point de vue du décideur, qui est celui de la procédure du choix collectif, devrait être neutre eu égard au contenu des préférences. Exclure certaines préférences en raison de leur contenu suppose une norme qui vient s'ajouter à celle définie par le choix collectif. Or, quel serait le fondement de cette norme ?

Les autres arguments de Goodin sont plus prometteurs. Le premier consiste à affirmer qu'une préférence perverse actuelle pourrait cesser d'être celle de l'agent après « lessivage ». Le filtrage des préférences aurait alors une fonction de révélateur : les personnes ayant des préférences antisociales comprendraient qu'elles ne peuvent à proprement parler s'identifier à de telles préférences : on ne serait méchant que par ignorance. Le second argument défend le filtrage en amont des préférences en ajoutant que les responsables de cette politique seraient toujours tenus de rendre des comptes à la collectivité, ce qui revient à faire du filtrage un choix collectif, si ce n'est que l'occurrence de l'état social collectivement préféré précède la production de la préférence elle-même.

La proposition de Goodin met en évidence l'importance d'imposer des conditions de rationalité à la préférence qui soient propres au contexte d'un choix collectif. Dans un tel contexte, englobant à la fois l'évaluation des configurations sociales et économiques, d'une part, et les décisions qui en découlent, d'autre part, les restrictions imposées aux préférences des membres de la collectivité ne peuvent être justifiées que lorsque lesdites préférences sont susceptibles d'annuler la possibilité même d'une procédure de choix collectif. Notons qu'il faut distinguer cette possibilité de la possibilité du choix collectif lui-même : même si une délibération et un vote peuvent avoir lieu, cela n'implique pas qu'ils puissent générer un choix collectif satisfaisant et cohérent.

La difficulté est que les filtres imposés de l'extérieur aux préférences risquent de ne pas être efficaces s'ils ne sont pas complétés par des filtres « internes ».

Rationalité et thérapie des préférences

Le philosophe utilitariste Richard Brandt définit la préférence pour le bien comme une préférence rationnelle[361]. Qu'est-ce qu'une préférence rationnelle, et en quoi se distingue-t-elle de préférences irrationnelles ? Une préférence rationnelle est déterminée d'une part par des facteurs cognitifs, et d'autre part par des attirances et des aversions – ce que l'on nomme généralement désir ou volonté, mais Brandt privilégie le terme « valence », emprunté à la psychologie. De fait, Brandt utilise certains résultats de la psychologie moderne afin d'établir les principes permettant de distinguer les préférences rationnelles de celles qui ne le sont pas. Le diagnostic des préférences irrationnelles repose donc sur l'identification d'erreurs d'information ou de raisonnement, et sur les travaux psychologiques montrant dans quelle mesure les valences sont conditionnées. La thèse de Brandt est qu'une préférence rationnelle est une préférence qui peut « survivre » à une psychothérapie cognitive.

Les préférences rationnelles résultent donc d'une démarche critique en deux étapes. Tout d'abord, les valences sont considérées comme données et la critique ne porte que sur les facteurs cognitifs. Brandt constate que les préférences actuelles peuvent être mal informées. L'information est pour Brandt le privilège de la préférence rationnelle[362].Mais *A Theory of the Good and the Right* distingue également d'autres facteurs cognitifs, qui sont autant de déterminations susceptibles de distinguer notre préférence rationnelle des préférences actuelles informées par définition. Brandt met notamment l'accent sur l'importance de l'attention[363]. Un manque d'attention peut nous amener à négliger certaines options, ou à ne pas nous représenter complètement les détails d'une nouvelle situation. Supposons par exemple un scientifique européen auquel est proposé un poste de chercheur en Amérique du Nord, dans lequel il pourra réaliser un projet qui lui tient à cœur. Il accepte cette opportunité, mais sans songer aux changements de mode de vie que ce déménagement implique. Or, il ne sait pas conduire, et n'a pas songé aux conséquences du manque de transports en commun dans la nouvelle ville où il va vivre. Une fois installé, il regrette son choix. Il ne s'agit pas à d'un problème d'information, mais d'une erreur de raisonnement : cet homme n'a pas songé à vérifier que sa nouvelle situation réalise ses préférences dans tous les domaines. Les autres sources d'erreur possibles citées par Brandt sont l'humeur, l'occurrence d'émotions incontrôlables, la force exagérée d'un désir : des personnes songent au suicide lors d'un accès

[361]BRANDT, Richard, *A Theory of the Good and the Right,* Oxford : Oxford University Press, 1979.

[362] Donc, de toute préférence, puisque nous ne voyons pas très bien ce que serait une préférence qui subsisterait indépendamment d'un agent individuel ou collectif pour l'énoncer à un moment donné.

[363] BRANDT, *op. cit.*,chapitre IV.

de mélancolie, et cependant y renoncent aussitôt revenues dans un état d'esprit plus calme. Brandt souligne également la tendance à négliger les désirs futurs : c'est la cigale de la fable qui, l'été, oublie de songer à sa faim une fois l'hiver venu.

La seconde étape définie par Brandt consiste en une thérapie des valences elles-mêmes. Brandt définit une valence comme le fait d'être dans un état motivationnel. Une personne veut un objet O, ou un objet O déclenche une valence chez cette personne, si sa disposition est telle que, si une action susceptible de faire advenir O lui semble possible, alors cette personne la tendance de cette personne à effectuer cette action sera accrue. Brandt définit cette tendance à agir en termes de probabilités : la tendance à effectuer une action est fonction à la fois de la probabilité de son succès et de la valence de l'agent pour cette action.

Y a-t-il des valences irrationnelles, des dispositions que l'agent a mais qu'elle cesserait d'avoir dans des certaines conditions hypothétiques idéales ? Si oui, cela implique qu'une personne rationnelle ne ferait pas l'expérience de certains désirs. Nous avons écarté les conceptions du bien-être qui assignent un contenu donné aux désirs : la conception hédoniste supposait que le désir est recherche de plaisir, la conception du bien-être comme satisfaction des préférences ne faisait aucune distinction entre les désirs qu'il faudrait avoir et ceux qu'il faudrait éviter, et la conception objective du bien stipulait que la rationalité du désir dépend de son objet, et que ce dernier devait être déterminé indépendamment, par la nature ou par une convention par exemple. La conception du bien-être et du bien comme satisfaction des préférences rationnelles fait entièrement reposer la rationalité du désir sur le processus de génération de ce dernier.

Le problème de cette conception est qu'elle implique qu'il faudrait agir comme si nous avions des désirs que nous n'avons pas[364]. Comment est-ce possible ? Brandt offre une solution basée sur une conception cognitive de la rationalité des désirs : sont irrationnels les désirs auxquels nous renoncerions si nous avions eu davantage de savoir ; ainsi, un désir exagéré de réputation est basé sur une erreur cognitive, ou encore il nous arrive de désirer des choses auxquelles nous renoncerions si nous savions que nous ne pouvons pas les obtenir[365]. La thèse selon laquelle certains désirs sont rationnels et d'autres ne le sont pas présuppose que les facteurs cognitifs, croyances et opinions, jouent un rôle déterminant dans la formation desdits

[364] BRANDT, *op. cit.*, chapitre V.

[365] On remarquera que la conception de la préférence actuelle que nous avons défendue exclut d'emblée ce type de désir, puisqu'on ne peut logiquement attribuer à un agent une préférence pour un état social qu'à condition que les caractéristiques de cet état social déterminent la préférence. Or, le fait qu'un état social ne puisse pas arriver devrait être pris en compte dans la préférence.

désirs. Cette thèse correspond à la conception fonctionnelle de la préférence que nous avons défendue dans notre Chapitre 4, conception selon laquelle la préférence dispose à agir, mais en conjonction avec d'autres attitudes telles que les croyances et les opinions. Brandt articule cette conception du désir avec les résultats de recherches menées en psychologie et propose donc une analyse exploratoire de la genèse des désirs afin de déterminer le rôle des facteurs cognitifs dans leur formation[366].

La thèse de Brandt est donc que certains désirs, certaines attirances ou certaines aversions apparaitraient ou disparaitraient si les dispositions de la personne étaient nourries de l'information appropriée. La « psychothérapie cognitive » que propose Brandt afin de rendre les préférences rationnelles consiste à confronter les désirs avec une information pertinente, au moment approprié et en requérant l'attention appropriée de la part de l'agent. Brandt offre ensuite des cas d'erreurs dans la formation des désirs qui pourraient être « guéris » par une thérapie des préférences[367]. Les erreurs dans la formation des désirs que cite Brandt sont les suivantes :

- la dépendance des désirs à l'égard de fausses croyances, le conditionnement culturel ou social des désirs : par exemple, un désir exacerbé de réussite peut trouver ses origines dans l'imitation d'un modèle familial ;
- la généralisation à partir d'exemples peu représentatifs : Brandt cite le cas d'une personne effrayée par les chiens après avoir été attaquée dans son enfance ;
- un désir exagérément intense dû à une privation originelle, tel que l'avarice d'une personne qui aurait souffert de la pauvreté dans son enfance. Danse ce dernier cas d'erreurs, la thérapie cognitive consisterait à faire apparaître à cette personne que ses chances de se retrouver dans cette situation de pauvreté sont infimes, et que l'état de confort et de plaisir résultant de l'achat de certains biens lui procurerait davantage de bien-être.

Les conditions de rationalité de la préférence décrites par Brandt sont *in fine* des conditions cognitives. Même lors de la deuxième étape de la psychothérapie cognitive, qui concerne le contenu des désirs de l'agent, Brandt ne vise pas à définir le contenu de ces désirs en les articulant à une certaine conception du bien-être ou du bien, mais simplement à améliorer l'acquisition et l'usage de la connaissance des différents états sociaux dans lesquels se

[366] BRANDT, *ibidem.*

[367] On remarquera que l'on glisse du vocabulaire éthique vers une terminologie médicale, ce qui laisserait entendre que la préférence « rationnelle » est la préférence « normale », et l'irrationalité une maladie ; mais le problème est peut-être davantage celui d'une conception trop médicalisée de la psychologie elle-même, que celui de l'utilisation de la psychologie pour valider des hypothèses en éthique et en philosophie morale.

trouverait l'agent une fois ses préférences réalisées. En ce sens, la préférence rationnelle s'inscrit dans la continuité de la préférence actuelle déterminée par une connaissance suffisante des états sociaux. La psychothérapie cognitive consiste essentiellement à corriger les erreurs de perception et d'évaluation de l'importance des caractéristiques des états sociaux afin de faire en sorte que l'agent puisse se reconnaître dans sa préférence une fois celle-ci réalisée.

Les analyses de Goodin et de Brandt illustrent la manière dont on peut passer des préférences actuelles aux préférences rationnelles. De fait, notre conception des préférences actuelles suppose d'ores et déjà un effort d'acquisition des informations ; la rationalisation des préférences consisterait alors à effectuer un bon usage de ces informations.

D'autre part, notre analyse de l'évaluation des préférences d'un agent à partir d'une position extérieure a mis en évidence l'importance du retour réflexif de l'agent sur ses préférences passées. En effet, une personne peut cesser de se reconnaître dans les préférences qu'elle avait énoncées auparavant. Le phénomène du regret permet de caractériser la préférence regrettée comme irrationnelle, puisque la personne cesse de trouver celle-ci justifiée. Il y aurait donc, du moins en théorie, une justification suffisante pour réaliser en lieu et place d'une préférence actuelle erronée une préférence future que l'agent reconnaîtrait comme sienne.

Les arguments en faveur de la conception du bien-être comme objet de préférences rationnelles

Les développements de Goodin et de Brandt ont mis en évidence l'intérêt de privilégier les préférences rationnelles hypothétiques par rapport aux préférences actuelles dans un contexte de choix social. Tous deux jugent que le rapport entre préférences rationnelles et bien de l'agent doit être le suivant : les préférences rationnelles déterminent le contenu du bien de l'agent, et ne révèlent pas un bien déjà existant. Le bien reste indéterminé tant que les préférences n'ont pas été construites. Nous récapitulons ici brièvement les arguments en faveur de cette indétermination du bien. Tout d'abord, la conception du choix collectif comme simple révélation d'un bien indépendant occulte la dimension même de choix dans le choix collectif, en le réduisant à un simple processus de connaissance ; d'autre part, les conditions de rationalité imposées aux préférences sont telles qu'elles garantissent une certaine sécurité eu égard à l'objectif de choisir le meilleur ; enfin, le fait de recourir aux préférences préserve le caractère authentiquement démocratique du choix social. Ce souci de démocratie s'inscrit

dans notre proposition d'articuler les capabilités d'Amartya Sen au choix social et à une approche démocratique de l'évaluation.

Quelle conception de la préférence pour le choix collectif ?

Dans le chapitre précédent, nous avons montré que la préférence correspond à un certain type de jugement de valeur. Mais elle n'est pas purement descriptive, puisqu'elle dispose l'agent à choisir, en conjonction avec d'autres attitudes et d'autres états tels que les croyances. Or, la conception de la préférence comme révélation d'un bien objectif existant indépendamment de celle-ci ignore cette fonction de la préférence qui consiste à déterminer ou influencer un choix. S'il fait sens de parler de choix collectif, c'est-à-dire de détermination d'un état social plutôt que d'un autre, alors la préférence collective et les préférences individuelles agrégées qu'elle représente doivent être conçues comme engageant ce type de détermination.

Sécurité des préférences et conception du bien

En comparant la conception du bien comme satisfaction des préférences rationnelles et la conception du bien indépendante des préférences mais pouvant être décrites par celles-ci, John Broome s'inquiète de ce que la première ne suffirait pas à garantir le bien des personnes.

> *« On pourrait être tenté de réfuter la théorie du bien comme satisfaction des préférences rationnelles de la même façon que j'ai réfuté la théorie du bien comme satisfaction des préférences actuelles : en mettant en évidence un fait indéniable. Le fait que j'ai à l'esprit est que des personnes rationnelles et bien informées préfèrent quelquefois, face à une alternative, l'option la pire pour elles, parce qu'elle est meilleure pour quelqu'un d'autre et qu'elles veulent promouvoir le bien de cette personne. Ce fait, toutefois, n'est évidemment pas indéniable, puisque souvent, il n'advient pas. Donc la réfutation échouera.*
>
> *Pour le moment, malgré cela, je n'ai pas besoin de rejeter la théorie de la satisfaction des préférences rationnelles. Dans l'esprit de neutralité que j'ai adopté dans la section 2.4., je n'y toucherai pas. Cependant, je ne l'accepterai pas non plus. Bien que le fait que j'ai mentionné ne soit pas indéniable, il est du moins plausible. Par conséquent la théorie de la satisfaction des préférences rationnelles ne nous laisse pas en sécurité. Il est possible d'en faire le fondement d'un argument en faveur de la cohérence du bien, mais je ne le ferai pas. »*[368]

Broome est peut-être trop prudent. En effet, la théorie de la satisfaction des préférences rationnelles peut suffire à justifier les évaluations comparatives en termes de bien dont il analyse la structure dans son livre *Weighing Goods,* d'autant qu'il se défend d'adopter une conception substantielle du Bien :

> *« Une partie de l'éthique est concernée par le bien. Ce livre est sur cette partie. Spécifiquement, il est sur la structure du bien.*
>
> *Plus exactement, il est sur la relation « meilleur que » [*betterness*]. Certaines choses sont meilleures que d'autres, de telle sorte qu'une relation de ce type existe entre elles. Ce livre est sur la structure de cette relation. Même si je devrais utiliser le mot « bien » aussi bien que le mot « meilleur », je ne me réfère pas à une entité métaphysique mystérieuse quelconque, mais seulement au problème ordinaire de savoir quelles choses sont meilleures ou pires que d'autres. »*[369]

Broome juge que les réalités décrites par les jugements selon lesquels une chose est meilleure qu'une autre ne correspondent pas de manière certaine à celles décrites par des préférences rationnelles et bien informées. Son argument est que les préférences rationnelles visent parfois le bien d'une personne autre que l'agent. Cela implique que les jugements de valeur

[368]BROOME, John, *Weighing Goods*, Oxford : Blackwell, 1991, p 133. Traduction libre.
[369]*Ibidem,* p 1.

selon le critère comparatif du bien sont nécessairement déterminés par le bien de la personne qui énonce un tel jugement (en supposant bien entendu que la personne n'affirme pas explicitement que *x* est meilleur que *y* pour un autre qu'elle-même). Broome impose ici une restriction inutile aux évaluations comparatives selon le critère du bien. En effet, à partir du moment où l'on a renoncé à identifier l'intérêt de la personne à l'effet produit sur elle par un état social, il est tout à fait cohérent de dire que son intérêt correspond au bien d'un autre, dans la mesure où sa disposition est rationnelle.

Broome doute la capacité des préférences rationnelles à décrire le bien. Notre analyse a pourtant mis en évidence que les seules préférences, correctement définies, devraient être déjà très exigeantes, puisqu'elles requièrent une connaissance suffisante des états sociaux pour se décider. Les préférences rationnelles ajoutent encore des restrictions supplémentaires aux fins admissibles pour la collectivité comme pour les personnes individuelles à long terme. Comme l'écrit Daniel Hausman :

> *« Contrairement aux désirs, qui viennent souvent à l'improviste et n'ont pas besoin d'être compatibles avec les autres désirs et valeurs d'un individu, les préférences reposent sur l'évaluation pondérée de considérations concurrentes et dépendent de jugements concernant les mérites des alternatives qui sont comparées. Quelquefois les comparaisons sont triviales. Jill a soif. Elle voit une fontaine d'eau. Elle a de bonnes raisons de boire et aucune raison de ne pas boire. Alors elle préfère boire à ne pas boire. »*[370]

Mais lorsque les comparaisons ne sont pas triviales, les préférences exigent un processus de réflexion plus complexe. Hausman donne l'exemple d'une comparaison entre la surdité et le handicap moteur : de telles évaluations sont pertinentes dans un contexte de choix collectif d'allocation des ressources publiques. Il montre que de telles évaluations ne devraient pas faire l'objet d'une réponse immédiate à une question, lors de sondages par exemple.

> *« Si, au lieu de répondre à un sondage, il y avait une vraie question, ce serait une des plus importantes décisions que l'on devrait avoir jamais à faire, et la plupart des gens réfléchiraient longuement avant de répondre. Ils chercheraient des informations concernant ce que c'est que d'être incapable d'entendre ou de marcher. Ils essaieraient d'imaginer concrètement à quoi ressembleraient leurs vies. Ils voudraient savoir quelles carrières ils peuvent poursuivre, quels loisirs ils pourraient pratiquer, comment leur famille et leurs amis réagiraient. Une bonne part de connaissance de soi*

[370] HAUSMAN, Daniel, *Preference, Value, Choice and Welfare*, Cambridge: Cambridge University Press, 2012, p 120.Traduction libre.

serait aussi requise. Je me demanderais moi-même ce à quoi j'attache le plus d'importance et à quelles choses je pourrais le plus facilement m'adapter. »[371]

Enoncer une véritable préférence ne se réduit pas à répondre au hasard à un sondage ou à une question de jeu télévisé. Lorsque la préférence implique une réalisation, le processus de comparaison engage une réflexion approfondie sur les caractéristiques factuelles des états sociaux, les conséquences de ces derniers, la manière dont ils affectent l'agent concerné, ceux et celles dont il s'inquiète, ses relations avec autrui, ses valeurs éthiques, son identité et la reconnaissance dont il jouit. Cette réflexion requiert des informations, un raisonnement, de l'imagination afin de concevoir les différentes possibilités, la capacité de se représenter dans différentes situations, de savoir quelles seront nos réactions émotionnelles future.

Peut-être que les évaluations comparatives déterminées par un critère objectif du bien requièrent encore plus. Mais sont-elles alors seulement faisables ? Les conditions de rationalité substantielle des préférences décrites par Goodin, Brandt et Hausman paraissent d'ores et déjà imposer des contraintes exigeantes sur les préférences. L'insécurité soulignée par Broome à l'égard de la détermination du bien par les préférences rationnelles semble donc une préoccupation exagérée.

[371] *Ibidem, p* 121.

L'autorité de l'agent concernant son bien

Le dernier argument en faveur de la détermination du bien par les préférences rationnelles est le suivant : si une conception du bien déterminant les existences des personnes doit être élaborée indépendamment des opinions, valeurs et jugements de ces dernières, qui a autorité pour le faire ? Il est difficile de faire appel à une quelconque transcendance dès lors que l'on présuppose la possibilité et surtout l'importance du choix collectif. On ne se débarrasse pas si facilement de la préférence une fois éliminées les interprétations les plus réductrices de cette dernière. Il n'est pas absurde ni incohérent d'affirmer que la conception du bien commun devrait ignorer les évaluations des membres de la collectivité. Mais, en ce cas, le rôle du choix collectif se trouverait réduit à rien ; il s'agirait d'un paradigme politique qui n'est pas celui de la démocratie. Or, l'étude prolongée des règles de choix social suppose déjà, en amont, le pari de la rationalité collective et de l'espoir de déterminer certains états sociaux à partir des évaluations et des attentes des membres de la société.

IIIe partie : du choix social à l'évaluation des capabilités

A ce stade de notre réflexion, rappelons que les motivations de l'étude du cadre conceptuel de la théorie du choix social consistaient à intégrer l'évaluation des capabilités dans celui-ci, faute d'une théorie de la justice satisfaisante des capabilités.

Mais un tel projet se heurte à première vue à la spécificité des capabilités. Comme l'écrit Mozaffar Qizilbash, tandis que la théorie du choix social concerne les évaluations sociales, les capabilités relèvent plutôt de l'évaluation de situations individuelles :

> *« A première vue, les domaines du choix social et de l'approche par les capabilités semblent tout à fait distincts. Dans sa conférence Nobel, Amartya Sen écrit que le 'sujet de la théorie du « choix social » inclut, à l'intérieur de son vaste champ, des problèmes divers dont la caractéristique commune est de mettre en relation des jugements sociaux et des décisions collectives avec les opinions et les intérêts des individus qui composent la société ou le groupe.' En revanche, l'approche par les capabilités de Sen concerne les états d'une* personne. *La 'capabilité' d'une personne désigne les différentes collections de* n-tuple *ou d' 'être' et de 'faire' ou de 'fonctionnements' qui constituent les vies parmi lesquelles elle peut en choisir une. Elle reflète les* libertés *ou les* opportunités *de la personne. Dans cette mesure, l'approche par les capabilités semble se concentrer sur l'individu plutôt que sur la société. »*[372]

En substance, Qizilbash attribue à la théorie du choix social l'objet « société » et réserve aux capabilités l'objet « individu ». Si donc la théorie du choix social et les capabilités ne se distinguaient que par leur objet, on pourrait envisager une complémentarité des deux. Mais, d'une part, notre travail a montré que cette distinction est superficielle. La base d'information du choix social peut être constituée de préférences individuelles – même si nous avons soutenu qu'il convenait mieux de parler de préférences positionnelles –. Quant aux capabilités, elles mettent en relation les différences de niveau et de qualité de vie entre individus avec des déterminants sociaux. De plus, on peut attribuer des capabilités à un groupe, même si Sen et la plupart des théoriciens des capabilités ne revendiquent pas cet

[372] QIZILBASH, Mozaffar, "Social choice and individual capabilities", *Politics Philosophy Economics* 6,2007, pp 169-192. Traduction libre.. Sur le problème de la compatibilité de la théorie du choix social et des capabilités, cf. également la thèse de Muriel Gilardone, *Contexte, sens et portée de l'approche par les capabilités de Amartya Kumar Sen*, Thèse de doctorat, Université Lumière Lyon 2, Faculté de sciences économiques et de gestion, 2007, qui défend une vue proche de la nôtre.

usage du concept ; à tout le moins peut-on dire que plusieurs individus partageant une même position partageront également une même capabilité.

D'autre part, l'association de la théorie du choix social avec l'approche par les capabilités soulève un problème beaucoup plus fondamental, que Qizilbash met également en évidence :

> *« Une opposition centrale entre la théorie moderne du choix social et l'approche par les capabilités tient au rôle crucial des* préférences *dans la première. La focalisation sur les préférences est claire dans la formulation d'Arrow. Sen lui-même a soutenu que la base du choix social faite de préférences individuelles pouvait être vue comme un avantage pour la théorie du choix social, à cause des nombreuses façons d'interpréter la notion de préférence. Toutefois, la conception fine de Sen, qui est constituée par des arguments en faveur des capabilités et des fonctionnements comme espaces d'évaluation, est basée sur une critique de certaines représentations de la qualité de vie qui conçoivent cette dernière en termes d''utilité', comprise comme la satisfaction des désirs ou le bonheur. Ce point est important à cause du problème bien connu des préférences ou des désirs adaptatifs. »*[373]

Les capabilités sont en bonne partie une réponse au problème des préférences adaptatives et peuvent être vues comme une critique radicale du welfarisme qui caractérise la théorie du choix social et les fonctions de bien-être social. Mais, curieusement, les travaux de Sen en théorie du choix social n'ont pas remis en cause l'usage d'une base d'information préférentialiste, mais seulement leur interprétation.

> *« [Sen] suit Arrow en soulignant que dans [le contexte de la mesure de la liberté et des opportunités] il utilise les préférences au sens de valeurs plutôt que de goûts, de satisfaction mentale ou de désir. Pour autant que ce sens soit distinct de la satisfaction mentale ou des désirs, et dans la mesure où les valeurs sont moins malléables que les états mentaux ou les désirs, ce déplacement semblerait limiter le problème de l'adaptation pour l'utilisation des préférences en théorie du choix social. Le fait que les valeurs elles-mêmes puissent s'adapter et être socialement conditionnées subsiste, néanmoins, et c'est sans doute une des raisons pour lesquelles Sen suggère qu'il serait bon pour la théorie du choix social d'étudier plus en profondeur la formation des préférences. »*[374]

A présent, la portée d'une analyse conceptuelle fouillée des présupposés individualistes et welfaristes de la théorie du choix social apparaît plus clairement. Si on veut articuler théorie du choix social et approche par les capabilités, il est crucial de dissocier la préférence du contenu descriptif qui lui est généralement assigné, à savoir une expérience mentale de satisfaction moralement arbitraire et influencée par l'environnement. Il est également crucial de ne pas présupposer que les préférences sont livrées à l'arbitre de l'individu : au contraire,

[373] *Ibidem.*
[374] *Ibidem*

elles dépendent pour une bonne part de la position économique et sociale de cet individu, position qu'il peut partager avec d'autres individus, de même que les capabilités qui décrivent les privations et les possibilités caractéristiques de cette position.

Un autre facteur d'incompatibilité possible entre théorie du choix social et approche par les capabilités concerne le mode de description des objets du choix.

> *« Dans la proposition classique d'Arrow, les objets du choix social sont des 'états sociaux', tandis que dans l'approche par les capabilités, les objets du choix sont des collections de fonctionnements qui sont des états de la personne. »*[375]

Nous avons vu dans le chapitre précédent que l'étendue et la précision de la description d'un état social peut être cruciale pour la détermination d'une préférence. La question est à présent la suivante : peut-on faire varier le cadre conceptuel de la théorie du choix social de sorte qu'il soit possible de substituer à une description des options en termes d'états – une description plutôt statique, par conséquent – une description en termes de fonctionnements - de caractère davantage dynamique, dans la mesure où les fonctionnements décrivent des processus et des activités, et pas seulement des états d'une personne ? Si la théorie du choix social doit nécessairement présupposer une certaine conception de la réalité et de ce fait restreindre les possibilités descriptives de cette réalité, par exemple en revendiquant l'association inéluctable d'une proposition à un fait statique, alors notre défi d'associer choix social et capabilité est perdu d'avance. Mais pourquoi le devrait-elle ? Et si elle ne le doit pas, n'y a-t-il pas ici le point de départ d'une solution au problème des préférences adaptatives – dans la mesure où une description des états sociaux en termes dynamiques est plus riche, donc plus généreuse en informations, et de ce fait plus susceptible de déterminer une préférence réfléchie - ?

Dans cette dernière partie, nous proposerons une conceptualisation de la sélection et de l'évaluation des capabilités importantes pour la qualité de vie. A l'instar de Sen, nous ne concevons une telle procédure de sélection que de manière démocratique. Il faudrait éviter que la liste des capabilités importantes pour la qualité de vie soit imposée d' « en haut ». Le point de départ de cette conceptualisation de la sélection démocratique des capabilités est que cette sélection peut être basée sur les « préférences » et donc s'articuler avec une interprétation donnée du cadre conceptuel de la théorie du choix social. L'interprétation des préférences que nous utilisons ici est donc celle mise en place dans la partie précédente : les préférences sont comprises comme des jugements de valeur comparatifs, dont la fonction est

[375] *Ibidem*

de déterminer un choix, et qui sont susceptibles d'être filtrées ou ajustées par des conditions de rationalité.

Le Chapitre 6 expose la définition des capabilités et l'avantage que cet outil conceptuel d'évaluation des inégalités et de la pauvreté présente par rapport à d'autres outils. Le Chapitre 7 est consacré à la question de la sélection démocratique des capabilités importantes pour la qualité de vie.

Chapitre 6 : Les capabilités : conséquentialisme et liberté réelle

Une théorie de la justice ou, plus généralement, une approche donnée de l'évaluation sociale, se distingue par sa métrique et par son critère de répartition. La métrique, c'est la variable de la justice, ce que l'évaluation sociale mesure et ce dont on vise la redistribution. Le critère de répartition définit la manière dont la substance mesurée à l'aide de la métrique est allouée entre les différents porteurs de revendications – des individus lorsque la théorie de la justice endosse un individualisme éthique, c'est-à-dire lorsqu'elle fait du bien de l'individu la finalité des actions qu'elle préconise[376][377][378].

L'approche par les capabilités revendique, comme son nom l'indique, la capabilité comme métrique la plus appropriée pour évaluer la pauvreté et la richesse, définir les attentes des membres de la société et orienter les actions sociales. En revanche, elle reste muette au sujet du critère de répartition. Or, il est malaisé de décider de la répartition d'une variable multidimensionnelle comme les capabilités. Si on les réduit à une seule dimension, on court le risque de perdre ce qui fait leur intérêt même, à savoir leur richesse informationnelle.

L'approche par les capabilités est donc intrinsèquement pluraliste. Or, ce pluralisme se retrouve dans le caractère hétérogène assumé des collectivités que décrit le cadre conceptuel de la théorie du choix social. C'est ici que nos travaux d'analyse historique et conceptuelle de la théorie du choix social prennent sens. En effet, en montrant que l'interprétation standard du cadre welfariste de la théorie du choix social n'englobe pas tous les usages possibles de ce cadre, nous avons ouvert la porte à une articulation possible entre capabilités et choix social. Plus exactement, la théorie du choix social pourrait jouer le rôle consistant à définir le critère de répartition des capabilités.

Dans ce chapitre, nous récapitulerons les arguments en faveur d'une métrique conséquentialiste pour l'évaluation sociale (6.1.), avant d'exposer plus précisément les raisons

---

[376] Cf.VAN PARIJS, Philippe, *Qu'est-ce qu'une société juste ?*, Paris, Le Seuil, 1991, chapitre 10.

[377] Il existe également des théories de la justice qui refusent explicitement d'adopter des critères de répartition [*pattern*]. Ainsi, Nozick préfère définir en amont les procédures légitimes d'acquisition des ressources, puis de laisser les interactions économiques faire le reste. Cf. NOZICK, Robert, *Anarchy, State and Utopia*, New York, Basic Books, 1977; traduction française par Evelyne d'Auzac de Lamartine, *Anarchie, Etat et Utopie*, Paris, Presse Universitaires de France, 1988, 2003, 2008.

[378] A l'exception des théories communautariennes, la plupart des théories de la justice contemporaine revendiquent cet individualisme éthique. En revanche, elles n'endossent pas toutes un individualisme descriptif ou encore explicatif, et sont loin de s'accorder pour expliquer les inégalités par des causes individuelles.

d'adopter une perspective centrée sur les opportunités et les capabilités plutôt que sur les résultats (6.2.).

6.1. Conséquentialisme et mesure des résultats

L'économie normative et les théories de la justice sociale partagent une même préoccupation : comment décrire de la manière la plus adéquate la situation d'un individu ? Comment savoir si une personne est heureuse, si ses besoins sont satisfaits ? Ses libertés de base sont-elles garanties? Nous présenterons les principaux arguments en faveur d'une conception conséquentialiste de la justice sociale (6.1.1.), avant d'analyser les deux classes principales de métriques conséquentialistes, les utilités et préférences d'une part (6.1.2.), et les opportunités de bien-être et les capabilités d'autre part (6.1.3.).

6.1.1. Pour une conception conséquentialiste de l'évaluation sociale

Il existe de nombreuses méthodes de classification des métriques de la justice[379]. Nous proposons ici une méthode basée sur la distinction entre métriques déontologiques et métriques conséquentialistes. La distinction entre conception déontologique et conception conséquentialiste de la justice sociale est définie analytiquement par Thomas Pogge :

> *« […] une conception de la justice est déontologique si et seulement si elle soutient* à la fois *que (A) tous les bénéfices et tous les coûts qu'un schème institutionnel apporte sont toujours plus importants que les biens et les maux qu'il laisse simplement arriver (de sorte que ces derniers puissent figurer au plus comme une question subsidiaire dans l'évaluation des institutions) et que (B) à l'intérieur du domaine de ce qu'un schème institutionnel provoque, les coûts et les bénéfices qu'il établit ont parfois plus de poids que les coûts et bénéfices que l'on prévoit qu'il engendrera. »*[380]

Pogge définit la théorie de la justice de Nozick comme déontologique. La théorie de Rawls est semi-conséquentialiste. En effet, Nozick et Rawls s'accordent sur la proposition (A): les institutions ne sont pas supposées corriger les effets d'événements qu'elles ne provoquent pas elles-mêmes. De là vient la distinction rawlsienne entre biens premiers naturels et biens premiers sociaux, les seconds seuls pouvant faire l'objet d'une redistribution ou d'une

[379] Cf. FLEURBAEY, Marc, *Théories économiques de la justice*, Paris, Economica, 1999, § 1.6., KYMLICKA, *Contemporary political philosophy: an Introduction*, Oxford University Press, 1990, 1999 ; traduction française Marc Saint-Upéry, *Les théories de la justice: une introduction*, Paris: La Découverte, 1999 ; ROEMER, John, *Theories of Distributive Justice*, Cambridge (Massachussets), 1995, chapter 5 ; VAN PARIJS, Philippe, *Qu'est-ce qu'une société juste ?*, Paris, Le Seuil, 1991, chapitre 10.

[380] POGGE, Thomas, *Realizing Rawls*, Ithaca: Cornell University Press, 1989, 3.5..Traduction libre.

compensation. De ce fait, les handicaps physiques ou les problèmes de santé ne relèvent pas d'une prise en charge institutionnelle[381]. En revanche, Pogge n'attribue la proposition (B) qu'à Nozick. Rawls se soucie en effet des coûts et bénéfices que l'on peut prévoir, et pas seulement de ceux qui sont établis *a priori.* Pour comprendre cela, rappelons que l'objet de la théorie de la justice rawlsienne, ce ne sont pas les institutions évaluées du point de vue de leurs caractéristiques intrinsèques ; il s'agit de la structure de base, c'est-à-dire de la manière dont les institutions affectent les existences individuelles et la vie sociale. Supposons que les institutions ne prennent pas en charge les congés parentaux : il est vraisemblable que les femmes cesseront de travailler lorsqu'elles auront des enfants. Cela peut affecter la perception sociale des rôles sociaux assignés aux femmes. Une perspective déontologique ignore cela, tandis que le semi-conséquentialisme rawlsien prendrait ce type de faits en compte.

Les conceptions conséquentialistes de la justice nient à la fois (A) et (B) : ces conceptions évaluent les actions sociales, les politiques publiques et les configurations socio-économiques en général du point de vue des biens identifiables qui viennent à exister. Peu importe que la structure institutionnelle ait provoqué directement l'existence de ces biens, ou qu'elle ait simplement laissé ces derniers advenir ou subsister. Le conséquentialisme évalue les règles et les normes exclusivement du point de vue de leurs effets, et non pas de leur constitution ou de celle du système dont elles découlent. Kymlicka identifie trois arguments en faveur d'une conception conséquentialiste de la justice, arguments que nous reprenons ici[382].

Le premier argument est que le conséquentialisme repose sur une description objective des biens générés par l'acte ou la règle que l'on évalue. C'est-à-dire qu'avant même d'énoncer un jugement de valeur, un conséquentialiste s'assure que nous puissions nous accorder sur les caractéristiques et les composantes de l'état résultant de l'acte ou de la règle évalués. Kymlicka illustre cet argument par les exemples suivants et en conclut que le conséquentialisme évite de ce fait l'arbitraire moral :

> *« Nous avons tous eu affaire à des gens qui prétendent que telle ou telle activité – l'homosexualité, par exemple, ou le jeu, la danse, la boisson, le fait de jurer, etc. – est moralement condamnable, tout en étant incapables de signaler les conséquences censément négatives qui en découleraient. Le conséquentialisme nous défend de formuler de tels interdits, qui ont toute l'apparence de l'arbitraire moral. Il exige que quiconque condamne un comportement donné démontre que* quelqu'un d'autre est

[381] A moins que l'on n'explique ces problèmes de santé par des facteurs économiques et sociaux. En ce cas, la critique adressée à Rawls et à Nozick serait qu'ils ont une mauvaise interprétation de la réalité, et non pas que l'approche déontologique est par elle-même défectueuse.

[382]KYMLICKA, *Contemporary political philosophy: an Introduction,* Oxford University Press, 1990, 1999 ; traduction française Marc Saint-Upéry, *Les théories de la justice: une introduction,* Paris: La Découverte, 1999, p 18-20.

lésé *par ce comportement, qu'il a des conséquences nuisibles pour la vie d'un tiers. De même, du point de vue conséquentialiste, un comportement n'est moralement louable que s'il bénéficie à quelqu'un d'autre. »*

Il est moins arbitraire de s'accorder sur la description d'un fait qui subsiste que sur les qualités intrinsèques de règles morales variables selon les communautés ou les individus. Cet argument de Kymlicka met en évidence le présupposé réaliste du conséquentialisme : les conséquences des actes ont une réalité indépendante de la perception ou des idées subjectives. L'argument présuppose quant à lui une conception non-autoritaire de la norme : il n'y a pas lieu d'intervenir au nom de cette norme et de décider de la manière dont les personnes devraient vivre tant que ce mode de vie n'affecte pas les autres personnes.

Le deuxième argument de Kymlicka est que le conséquentialisme correspond à une distinction bien définie entre les jugements de valeur moraux et les jugements de valeur d'ordre esthétique ou religieux, par exemple. Le conséquentialisme ne s'intéresse qu'aux effets des actes ou des règles sur le bien-être ou l'intégrité des individus, et non aux propriétés intrinsèques de ces actes ou de ces règles. Un conséquentialiste se moque de savoir si un acte est beau, honorable, sacré, convenable, conforme aux traditions…Cela permet de distinguer plus aisément la sphère de la moralité d'autres sphères dans lesquelles sont énoncés des jugements de valeur.

Le troisième argument est que le conséquentialisme permet d'éviter les conservatismes. En effet, n'accordant pas de valeur en propre aux institutions, aux règles ou aux normes, il peut être à le point de départ d'une critique sociale radicale. Le conséquentialisme d'un Bentham ou d'un Mill leur a ainsi permis d'afficher en leur temps des positions très progressistes sur la condition des femmes.

Nous défendrons donc ici une conception conséquentialiste de la justice sociale. Le lecteur familier des travaux de Rawls s'étonnera alors peut-être de l'usage du terme « justice ». Dans la *Théorie de la Justice*, le « juste » correspond aux limites imposées par les principes de justice définis à partir de la procédure de raisonnement représentée par l'accord dans la position originelle. Il est distingué du « bien », qui décrit à la fois les finalités poursuivies en privé par les citoyens, les instruments qui servent à la réalisation de ces finalités, et enfin la société bien ordonnée, régie par les principes de justice, que les citoyens sont peu à peu amenés à aimer et désirer pour elle-même (Rawls renoncera à cette dernière manière de décrire le « bien » dans *Libéralisme politique,* craignant que la thèse de la justice comme bien soit incompatible avec le pluralisme des valeurs qui caractérise les sociétés modernes). Or,

nous adoptons dans notre raisonnement des conceptions de la « justice » qui correspondent à ce que Rawls décrirait en termes de « bien » et non de « juste ». Il convient donc de préciser que nous distinguons ici la « justice sociale » au sens large de la « justice sociale » au sens étroit, et que nous utilisons plutôt le premier sens du terme. La « justice sociale » décrit simplement une approche explicitement normative de la philosophie politique, dans la lignée des travaux de Rawls. La justice au sens étroit du terme correspond à la définition rawlsienne, c'est-à-dire à une norme distincte du bien, définie à l'aide du raisonnement représenté par l'accord dans la position originelle et dont l'objet correspond à la structure de base.

6.1.2. Le conséquentialisme utilitariste : de l'agrégation des utilités au vote sur la base de préférences

L'utilitarisme est une espèce de conséquentialisme. C'est un conséquentialisme « subjectif », au sens où la valeur des attributs pertinents pour l'évaluation sociale utilitariste dépend de l'appréciation subjective des individus ; en revanche, le point de vue de l'agrégation des utilités, qui correspond dans l'utilitarisme classique au point de vue (réel ou construit) du « législateur » ou du « décideur social », est un point de vue impartial et objectif[383].

[383] Le passage de ce point de vue subjectif au point de vue de l'agrégation est même au cœur des difficultés structurelles que rencontrent l'utilitarisme, l'économie du bien-être welfariste ou la théorie du choix social. En effet, cette approche de l'évaluation sociale se caractérise par le rejet de postulats tels que celui de la « main invisible » ou de l' « harmonisation spontanée des intérêts » qui permettraient d'éliminer la nécessité théorique du point de vue du législateur.. Cf. AUDARD, Catherine, *Anthologie historique et critique de l'utilitarisme,* Paris, PUF, 1999, Introduction.

a) L'agrégation utilitariste

Rappelons que les utilités décrivent la mesure cardinale de l'expérience subjective créée par une configuration du monde donnée. L'utilitarisme normatif prescrit la maximisation du bien-être collectif, c'est-à-dire de la somme des utilités des membres de la société considérée. L'évaluation utilitariste requiert tout d'abord d'assigner une mesure quantitative de ces états, en usant d'une échelle commune pour différents individus. Cette opération doit être effectuée pour chacune des décisions ou des états de chose que l'on compare. L'étape suivante consiste à agréger ces utilités en les additionnant. On retient alors l'état de choses ou la décision qui génère l'utilité la plus importante.

L'utilitarisme se heurte aux objections que nous avons explicitées dans la section 5.2.1. : d'une part, la mesure de l'utilité crée des problèmes méthodologiques importants, et une théorie de la justice susceptible de générer un accord entre différentes parties peut difficilement s'appuyer sur une échelle de mesure arbitraire ; d'autre part, faire de la maximisation de l'utilité l'unique objectif des actions individuelles comme des décisions sociales peut négliger certaines sphères importantes de l'existence humaine.

Outre le problème de la définition de l'utilité et de sa signification éthique, la maximisation de l'utilité apparaît également problématique. Le premier problème est la neutralité de la procédure de maximisation à l'égard des relations particulières entre les personnes. Le second problème concerne le rapport de l'utilitarisme à l'égalité.

L'utilitarisme est indifférent à l'égard des relations des individus entre eux. Chaque individu est considéré *a priori* comme une unité indépendante et interchangeable. L'utilitarisme fait donc abstraction de la réalité des relations privilégiées entre ces individus, qu'il s'agisse de l'affection ou de l'amour pour nos proches, ou encore des engagements que nous prenons et des promesses que nous faisons. Comme l'écrit Bernard Williams :

> *« si vous nourrissez authentiquement et sincèrement tous ces admirables engagements, aspirations et affections, vous ne pouvez pas en même temps refléter fidèlement dans votre action et votre raisonnement toutes les exigences de l'utilitarisme, et vous ne pouvez même pas souhaiter le faire. […] L'utilitarisme rejette nécessairement ou déprécie irrémédiablement la valeur de toutes ces dispositions, et s'appuie sur une image de l'homme que l'utilitarisme originel n'hésitait pas à défendre :*

un être doté idéalement d'aspirations purement privées, ou du moins aisées à sacrifier, et d'une seule véritable disposition morale, la bienveillance utilitariste. »[384]

Le fait que l'utilitarisme ignore les relations privilégiées résultant de l'amour pour la famille ou les proches, ou encore d'engagements et d'allégeances diverses, en fait une théorie morale difficile à concilier avec nos intuitions. On peut illustrer la bizarrerie de l'utilitarisme par la curieuse proposition de mariage que fait Mr Gradgrind, le personnage du roman de Dickens, *Les Temps difficiles,* à sa fille Louisa, âgée de vingt ans. Mr Gradgrind, fervent utilitariste, souhaite le mariage de celle-ci avec Mr Bounderby, son ami beaucoup plus vieux, comme parfaitement raisonnable, eu égard aux Faits, c'est-à-dire aux statistiques concernant les mariages d'hommes vieux avec des femmes jeunes, mariages parfaitement réussis. Lorsque Louisa demande à son père si elle doit aimer Mr Bounderby, Mr Gradgrind répond que non, mais qu'elle ne doit considérer la question que du point de vue des faits tangibles, sans se laisser embarrasser par des fantaisies. Cette histoire illustre l'étrangeté d'un raisonnement utilitariste appliqué à une relation privilégiée telle que le mariage.

Mais l'utilitarisme ne nous enseigne-t-il pas de nous méfier de nos intuitions et de nos préjugés ? S'il semble si exotique, c'est peut-être en raison de son caractère radical revendiqué. Ce qui ferait à la fois la limite et l'intérêt de l'utilitarisme, c'est ce refus d'être orienté par des intuitions peut-être aliénantes. Mais c'est également ce qui le rend difficile à appliquer.

Tournons-nous à présent vers les relations complexes qu'entretient le principe d'utilité avec celui d'égalité. L'utilitarisme est égalitaire dans la mesure où les intérêts de chaque individu pèsent pareillement lors de l'opération de sommation. Un article de philosophie politique, écrit par Jonathan Riley, formule ainsi l'interprétation utilitariste de l'égalité :

« L'interprétation utilitariste de l'égalité procédurale garantit que chaque citoyen se voit attribuer une opportunité égale d'influencer l'issue sociale au sens très limité où le bonheur de chaque personne a le même poids que celui d'une autre sur une base marginale ou par unité. »[385]

Cette égale considération des intérêts est un argument normatif en faveur de l'utilitarisme, même si un tel argument dépend d'un raisonnement non utilitariste – par exemple, d'un raisonnement s'appuyant sur un principe d'égale dignité des êtres humains ou des êtres

[384]WILLIAMS, Bernard, *Moral Luck,* Cambridge University Press, 1981. Traduction française *La Fortune morale: moralité et autres essais,* Paris : PUF, 1994.

[385]RILEY, Jonathan, « Utilitarian Ethics and Democratic Government », *Ethics,* Vol. 100, No. 2, Janv. 1990, pp 335-348. Traduction libre.

vivants doués de facultés d'éprouver du plaisir ou de la peine –. Mais l'utilitarisme n'est égalitariste que du point de vue de la recension des attentes, et non pas de celui de la satisfaction de celles-ci : en effet, si la propension des différents individus à être satisfaits diffère, alors l'objectif de maximisation du bien-être collectif prescrira la satisfaction de ceux qui sont le plus disposés à éprouver de la satisfaction en priorité.

Cela dit, il existe également un autre argument en faveur de l'égale considération des intérêts dans l'utilitarisme, selon lequel ce serait la meilleure méthode pour produire des états sociaux désirables car porteurs d'une utilité maximale. Ce second argument présuppose que les êtres doués de sensibilité sont tous dotés de capacités similaires d'éprouver du plaisir ou de la peine, et peuvent être substitués l'un à l'autre. Cet argument en faveur de l'égale considération des intérêts repose sur l'hypothèse de substituabilité des individus, conçus comme des réceptacles d'expériences plaisantes ou douloureuses. Mais cette hypothèse est-elle avérée ? Certaines personnes sont plongées dans un état végétatif et à notre connaissance incapables d'éprouver du plaisir ou de la peine. Et, en dehors de ces cas peu fréquents, nous rencontrons des personnes particulièrement optimistes, et d'autres plus mélancoliques. Ceux qui justifient l'égale considération des intérêts par la recherche d'états sociaux générant la plus grande satisfaction totale seraient prêts à supprimer comateux et mélancoliques si l'hypothèse de substituabilité des individus devait être réfutée. Même si l'utilitarisme met au défi la pertinence de nos intuitions morales, il devient difficile de défendre une théorie susceptible de prescrire notre propre suppression…La bienveillance utilitariste exige de nous un sens du sacrifice assez improbable.

b) Agrégation des préférences et égalité démocratique

Le passage de la maximisation du bien-être défini comme une utilité élevée à la maximisation du bien-être compris comme satisfaction des préférences permet de nuancer certaines des objections adressées à l'utilitarisme. En effet, les comparaisons interpersonnelles cessent d'être praticables lorsqu'on passe des utilités aux préférences. Or, comme le fait remarquer Jonathan Riley dans son article « Utilitarian Ethics and Democratic Government », les normes procédurales utilitaristes sont équivalentes aux normes démocratiques lorsque la base d'information est définie en termes de préférences et non d'utilités[386]. Il s'agit d'une équivalence logique : de même que le décideur social utilitariste, virtuel ou non, considère également tous les intérêts, de même le vote démocratique considère également toutes les

[386]RILEY, Jonathan, « Utilitarian Ethics and Democratic Government », *Ethics,* Vol. 100, No. 2, Janv. 1990, pp 335-348.

« voix », appelées préférences dans le jargon de l'économie du bien-être et de la théorie du choix social. En fait, Riley fait justement remarquer qu'en dépit des effets inégalitaires de l'agrégation utilitariste basée sur une information cardinale et comparable,

> *« il est fallacieux de conclure qu'il y a une disparité normative entre l'utilitarisme et la démocratie au sujet de l'égalité. »*

En réalité,

> *« La conception démocratique de l'égalité procédurale est ce qui pourrait être qualifié de conception purement ordinale-utilitariste. »*

Même s'il y a une équivalence logique entre l'agrégation des préférences ordinales et la collecte des votes, et si l'éthique utilitariste est un fondement possible de la démocratie, cela ne suffit pas à faire de la procédure d'agrégation un vote. En effet, l'utilitarisme classique comme les fonctions de bien-être social présupposent un calcul effectué du point de vue d'un décideur social. Mais les utilitaristes pourraient accepter de renoncer à ce procédé de calcul en faveur d'un vote en raison des coûts et des problèmes de praticabilité d'une telle opération. Les utilitaristes deviennent donc démocrates à cause des contraintes informationnelles.

Les autres raisons de passer de l'utilité, ou de la conception de la préférence comme état mental de satisfaction, aux ordres de préférences, sont exposées dans la section 5.2.2. de notre travail. Afin de garantir la sécurité informationnelle des préférences et leur capacité à représenter le bien-être de l'agent, il est souhaitable qu'elles soient informées (cf. section 5.2.2. de notre travail) et produites par une délibération rationnelle (cf. section 5.2.3. de notre travail).

6.2. De la mesure des résultats à l'évaluation des opportunités

Une métrique conséquentialiste de la justice sociale peut décrire les attributs des personnes de deux manières. Si elle correspond à un conséquentialisme étroit, seuls les résultats visibles seront pris en compte. Mais il n'est pas nécessaire de limiter la description des conséquences aux faits visibles ici et maintenant. Ainsi, Sen montre que la description des conséquences mérite d'être enrichie afin de rendre l'évaluation sociale plus pertinente :

> *« Un état de choses est riche en informations. Il n'y a pas de raison particulière d'insister en faveur d'une description appauvrie des états de choses lors de l'évaluation. Aussi, la portée du raisonnement*

en termes de conséquences peut incorporer les processus *de choix, et pas simplement les résultats étroitement définis. »*[387]

John Broome et Samuel Scheffler ont participé à cette discussion. Ils ont également critiqué la séparation stricte des processus et des résultats concernant la description des actions. Scheffler souligne ici que l'évaluation conséquentialiste peut parfaitement intégrer l'acte qui a produit le résultat évalué dans la description de la conséquence en question :

« Lorsque je dis que le conséquentialiste de l'acte requiert des agents qu'ils produisent dans l'ensemble le meilleur résultat ou le meilleur état de choses, je ne veux pas dire que le conséquentialiste de l'acte divise ce qui arrive en deux, les actes d'un côté et les résultats de l'autre, et évalue seulement ces derniers dans son principe de classement d'ensemble. Au contraire, l'acte lui-même est initialement évalué comme partie intégrante du résultat ou de l'état de choses dans l'ensemble. Le conséquentialiste de l'acte classe d'abord les résultats dans l'ensemble, résultats qui sont compris, de cette manière générale, comme incluant les actes nécessaires pour les produire, et ensuite il prescrit à l'agent de produire le meilleur résultat disponible selon cette définition. »[388]

Cette conception élargie du conséquentialisme de l'acte inclut donc les processus ayant produit les résultats évalués. En philosophie morale, ces processus correspondent aux actions des agents, individuels ou collectifs. Dans le contexte de l'évaluation sociale et de la justice, les processus peuvent également être le fait d'institutions. L'évaluation conséquentialiste élargie inclura donc dans la description des conséquences les droits ou les dispositifs institutionnels qui permettent la réalisation de telle ou telle configuration sociale ou de telle distribution des ressources.

Une conception conséquentialiste de la justice sociale ne s'intéresse donc pas qu'aux utilités ou aux préférences, mais également aux processus de production de ces utilités et de ces préférences. Or, cela implique également de considérer que certains processus auraient pu produire d'autres utilités ou d'autres préférences. Le conséquentialisme élargi nous conduit donc peu à peu à appréhender la métrique de la justice sociale non pas selon la modalité du résultat, mais selon celle de la possibilité, ou plus précisément de l'opportunité, dans la mesure où la possibilité doit être réelle et pas seulement formelle (l'agent doit disposer de tous les moyens de réalisation d'un état, et pas seulement se voir garantir une absence de contraintes). Les opportunités sont donc des ensembles de résultats possibles.

---

[387]SEN, Amartya, "Consequential evaluation and practical reason", *The Journal of Philosophy*, vol. 97, no 9, sept. 2000, pp 477-502. Traduction libre..

[388]SCHEFFLER, Samuel, *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions,* Oxford University Press, 1982, pp 1-2. Traduction libre..

Pourquoi choisir la perspective des opportunités plutôt que celle des résultats ?

Premièrement, les préférences changent, même lorsqu'elles sont informées et rationnelles. Par exemple, une personne peut être amenée à changer de religion (ou devenir croyante, ou cesser de l'être). Or, si l'on nous offre plus d'une option – par exemple, plusieurs opportunités en matière de religion et de spiritualité - , il sera plus aisé d'envisager un changement de préférence. L'évaluation de la justice sociale basée exclusivement sur les résultats revient à attribuer un résultat à un agent, ce résultat étant supposé le seul préférable pour l'agent ; en revanche, l'évaluation basée sur les opportunités laisse toujours ouverte la possibilité de modifier nos trajectoires de vie, puisqu'elle attribue des ensembles de résultats possibles à l'agent, au lieu du résultat donné qu'il lui est arrivé de préférer[389][390].

Deuxièmement, si nous décrivons la situation de l'individu en termes d'opportunités, cela suppose que nous accordions de l'importance à la réflexion qui constitue la comparaison des différentes opportunités, et à l'acte de choix qui en découle. Les métriques basées sur la description des opportunités présupposent à la fois l'existence et la valeur du choix délibéré.

Troisièmement, les métriques basées sur les opportunités plutôt que sur les résultats permettent la prise en compte de la responsabilité individuelle dans l'évaluation sociale (même si elles ne l'entraînent pas nécessairement) :

> *« L'argument en faveur de l'égale opportunité plutôt que de l'égalité tout court est simplement qu'il est moralement approprié de tenir les individus responsables des conséquences prévisibles de leurs choix volontaires, et en particulier de la portion de ces conséquences qui implique leur propre accomplissement du bien-être ou leurs propres gains et pertes de ressources. »*[391]

La perspective des opportunités interprète généralement les processus d'accomplissement des résultats comme des actions libres. Si le fait que les actions sont libres implique que les agents en sont responsables, alors cette perspective revient en effet à assigner à l'agent la responsabilité de la réalisation de l'une ou l'autre des opportunités dont elle dispose. Cela entraîne une division du travail entre société et individu, la première étant responsable de la quantité et de la qualité des opportunités offertes à chaque individu, les seconds ayant à leur charge l'évaluation et le choix de chaque opportunité. On pourrait objecter à cette approche le

[389] Cet argument présuppose une certaine permanence de l'identité personnelle (de même que toutes les discussions sur les changements de préférence). En revanche, si l'on juge que l'agent "change" avec ses préférences (cf. PARFIT, Derek, *Reasons and Persons*, Oxford :Clarendon Press, 1984) , cet argument en faveur d'une métrique de type « opportunité » plutôt que d'une métrique de type « résultat » perd son intérêt.

[390]Cf. ARNESON, Richard, « Equality and equal opportunity for welfare », *Philosophical Studies*, May 1989, Volume 56, Issue 1, p 84-85.

[391]ARNESON, Richard, « Equality and equal opportunity for welfare », *Philosophical Studies*, May 1989, Volume 56, Issue 1, p 88.Traduction libre..

fait que les individus sont parfois mal informés, ou encore sous l'emprise d'affects, d'influences sociales ou de perceptions biaisées les empêchant de choisir réellement. Mais c'est oublier que les opportunités sont des opportunités réelles : si des facteurs économiques, sociaux ou psychologiques empêchent les individus de choisir, cela signifie simplement qu'ils ont moins d'opportunités, voire aucune[392].

[392] Une autre objection serait que la métrique des opportunités postule le libre arbitre des individus. Mais, dans une perspective d'inspiration spinoziste, on peut interpréter la réalisation d'une opportunité non pas comme la détermination spontanée de celle-ci plutôt que d'une autre, mais comme la connaissance du processus de réalisation de cette opportunité, la conscience que l'agent en a. La métrique de l'opportunité tirerait alors sa valeur du fait qu'elle rend cette connaissance possible.

6.2.1. Les opportunités de bien-être

Le passage d'une métrique de l'ordre du résultat à une métrique de l'ordre de la possibilité laisse ouverte la question de la nature des opportunités.

Arneson exclut d'emblée la métrique de l'opportunité d'obtenir des ressources. Son argument est que l'évaluation sociale basée sur l'opportunité des ressources peut conduire à l'esclavage des talentueux – surtout si le critère de répartition des opportunités d'obtenir des ressources est l'égalisation - . En effet, les talents et compétences étant des ressources en elles-mêmes, les plus talentueux devront contribuer de manière disproportionnée par rapport aux autres membres de la société.

Arneson a donc défendu les opportunités de bien-être comme métrique appropriée de la justice sociale. Une opportunité de bien-être est un ensemble d'options dont la valeur dépend des perspectives de satisfaction correspondant au choix de chacune de ces options. Arneson propose de représenter les opportunités de bien-être ainsi :

> *« Supposons deux personnes entrant dans leur majorité et confrontées à divers choix de vie, chaque action que l'un pourrait choisir étant associée à son résultat possible. Dans le cas le plus simple, imaginons que nous connaissons la probabilité de chaque résultat sous la condition que l'agent fasse le choix de l'action qui mènera au dit résultat. Etant donné que l'un ou l'autre choix est fait, et l'un ou l'autre des résultats réalisé, l'agent serait alors confronté à un autre ensemble de choix, puis un autre, et ainsi de suite. Nous construisons un arbre de décision qui donne les histoires de vie complètes possibles d'un individu. Nous ajoutons alors la satisfaction des préférences espérée pour chaque histoire de vie possible. En faisant cela nous prenons en compte les préférences que les personnes ont eu égard au fait d'être confrontées avec une série particulière d'options données à chaque nœud de décision. »*[393]

Arneson s'assure de la validité de sa conception du bien-être en le définissant comme la satisfaction de préférences idéalement considérées. Plus exactement, il s'agit de préférences idéalement considérées de second rang [*ideally considered second-best preferences*]. Ces préférences « idéales » ou « rationnelles » sont celles résultant d'une délibération approfondie, avec une information complète et pertinente, sans être sous l'effet d'émotions fortes mais passagères, et sans erreur de raisonnement. Elles correspondent donc peu ou prou aux préférences individuelles rationnelles que nous avons analysées dans la section 5.2.4. de notre travail. Soucieux de la faisabilité d'une telle délibération, Arneson précise qu'il s'agit de

[393]ARNESON, Richard, « Equality and equal opportunity for welfare », *Philosophical Studies*, May 1989, Volume 56, Issue 1, p 85.Traduction libre..

préférences de second-rang [*second-best*] : ce terme emprunté à la science économique signifie ici que les préférences ne sont pas élaborées dans un contexte informationnel parfait, de « premier rang », mais dans un contexte satisfaisant, dont la description intègre certaines contraintes afin d'être plus proche de la réalité. Selon Arneson, ces contraintes incluent la résistance d'une personne aux conseils, même bons, les coûts d'éducation nécessaires pour vaincre cette résistance et la probabilité qu'un tel programme sera implanté dans la société au cours de l'existence de cette personne. L'exemple de préférences rationnelles de second rang que propose Arneson est celui d'une personne qui tire plus de satisfaction d'un vin bon marché que de la lecture de la poésie de Shelley ; les coûts de changement de ces préférences seraient trop élevés au regard de la satisfaction qu'en retire la personne, et de ce fait la réalisation du bien-être social peut ne requérir que des préférences de second rang.

L'argument en faveur des préférences rationnelles élaborées dans un contexte informationnel de second rang repose sur l'hypothèse selon laquelle ces préférences dépendent pour une bonne part du caractère de la personne. Il y aurait en certaines personnes une tendance spontanée et inexplicable à rechercher le vin médiocre plutôt que la poésie. Ces personnes seraient donc suffisamment heureuses de se voir dotées de vin et il serait trop coûteux de tenter de les faire changer d'avis. Si cet argument peut convenir au stade de l'application d'une théorie de la justice, il n'est ni nécessaire ni particulièrement bienvenu au stade de son élaboration. Une théorie philosophique de la justice vise à définir la structure de base ou la configuration économique et sociale la meilleure, celle de « premier rang ». Si les philosophes ne construisent pas des utopies, personne d'autre ne le fera. La réflexion sur la faisabilité intervient ensuite et appartient à la science politique, au droit, à la sociologie et à l'économie.

En revanche, l'exemple d'Arneson pourrait être réemployé afin de défendre la valeur d'une certaine diversité des goûts. Faire en sorte que chacun préfère la poésie de Shelley peut conduire à l'imposition d'une culture uniformisée à la négation des traditions minoritaires ou des mouvements *underground.* Mais la distinction pertinente n'est pas celle entre préférences de second rang et préférences de premier rang, mais entre des préférences résultant d'un programme éducationnel qui définit en amont les contenus préférables et équivaut à un endoctrinement – une certaine culture, par exemple - , et un programme éducationnel qui promeut au contraire la liberté dans la formation des goûts, tout en offrant aux personnes les outils et les savoirs nécessaires pour rechercher ce qui les épanouira le plus.

Par ailleurs, une perspective sociologique ne peut ignorer que la différence entre des préférences de second rang pour des satisfactions considérées comme moins élevées et des

préférences de second rang pour des plaisirs élitistes résulte d'inégalités sociales, des inégalités de ressources mais aussi des inégalités de positions et de reconnaissance. Et cette différence entretient également ces inégalités sociales[394].

Il serait donc plus judicieux de définir les opportunités de bien-être comme des opportunités de satisfaction de préférences idéalement considérées dans un contexte informationnel de premier rang, du moins au stade de l'élaboration de la théorie philosophique de la justice. Néanmoins, la discussion à propos des préférences rationnelles élaborées dans un contexte informationnel de second rang met en évidence certaines difficultés inhérentes à la description des options en termes de satisfaction subjective, et l'intérêt des capabilités développées par Amartya Sen et Martha Nussbaum par rapport aux opportunités de bien-être.

[394] Cf. BOURDIEU, Pierre, *La distinction, critique sociale du jugement*, Paris : Editions de Minuit, 1979.

6.2.2. Capabilités et fonctionnements

Les opportunités de bien-être et les capabilités partagent la critique des approches ressourcistes de l'évaluation sociale en général, et de l'indice rawlsien des biens premiers en particulier. Arneson, comme Sen et les théoriciens des capabilités, jugent que la valeur des ressources dépend de la capacité des agents à les convertir en outils d'accomplissement du bien-être, étant donnés leurs goûts, leurs intérêts, mais aussi leur condition objective, parfois caractérisée par des handicaps ou des défauts de connaissance ou de perception.

a) Capabilités ou opportunités de bien-être ?

En un sens spécifique du « bien-être », les capabilités peuvent être conçues comme des opportunités de bien-être. Ainsi, Peter Vallentyne défend la thèse selon laquelle la version la plus plausible des capabilités correspond à une généralisation des opportunités de bien-être[395]. Les opportunités sont ici comprises comme des libertés effectives : il ne suffit pas d'être formellement autorisé ou de ne pas être empêché de réaliser les fonctionnements pertinents pour le bien-être pour jouir des opportunités de bien-être. L'évaluation des opportunités inclut leurs conditions de réalisation, telles que la capacité de choisir avec discernement, l'absence de peur ou de honte, même irrationnelles, l'absence de domination...[396]. Le bien-être correspond ici à une conception qui n'est pas exclusivement subjective, et relève de la « qualité de vie » plutôt que du « plaisir ».

En effet, la différence essentielle entre capabilités et opportunités de bien-être est que les capabilités ne sont pas que des opportunités de bien-être au sens subjectif du terme [*welfare*], mais de bien-être au sens objectif du terme [*well-being*]. Les capabilités sont des opportunités de fonctionnements, et non pas de satisfaction de préférences. Or, non seulement le bien-être retiré de la satisfaction des préférences, même idéalement considérées, dépend en partie de la subjectivité de l'agent, tandis que le bien-être issu de l'exercice de fonctionnements, c'est-à-dire d'activités au sens large (aristotélicien) du terme – ces activités incluant davantage de dimensions de l'existence humaine que les actions, puisque le terme décrit aussi les «états » de la personne ou encore le fonctionnement physiologique, tel que l'assimilation de nutriment. - . La relation entre le bien-être [*welfare*] et la préférence n'est pas définie de la même façon que la relation entre le bien-être [*well-being*] et les fonctionnements.

[395] VALLENTYNE, Peter, "Capabilities VS Opportunities for Well-Being",Journal of Political Philosophy Volume 13, Issue 3, pages 359–371, Septembre 2005

[396] C'est parce que le terme "opportunité" est ambigu à cet égard que G.A. Cohen préfère celui d'"accès".

Amartya Sen définit la connexion entre bien-être et préférence comme une connexion empirique : l'état de bien-être [*welfare*] existe par lui-même, et sa relation avec l'objet de la préférence est arbitraire[397]. Supposons par exemple deux personnes, de même âge, de même sexe, de même taille et de même condition physique et psychologique, ayant reçu la même éducation, ayant besoin des mêmes soins médicaux, occupant une position sociale équivalente. Ces personnes pourraient, en dépit de cette similarité de condition, avoir des préférences différentes, et l'une pourrait se trouver satisfaite d'un état social, tandis que l'autre non.

En revanche, la valeur du bien-être [*well-being*] est conçue comme « survenante » sur la description des fonctionnements. La relation de survenance est définie ainsi par Pascal Engel :

> *« Une propriété M d'un objet survient ou est survenante sur une propriété P de cet objet, s'il ne peut y avoir de changement de M dans cet objet sans qu'il y ait un changement de P dans cet objet. »*[398]

Le bien-être est donc survenant sur un fonctionnement, s'il ne peut y avoir un accroissement ou une diminution de bien-être sans qu'il y ait un changement de fonctionnement. Par exemple, si le fonctionnement consistant à « assimiler les nutriments » cesse, le bien-être de la personne concernée est automatiquement affecté. En d'autres termes, « la survenance affirme que les propriétés évaluatives dépendent des propriétés non-évaluatives »[399]. Le jugement de bien-être dépend de ce qui se produit réellement, et que l'on décrit en termes de fonctionnements. En somme, la relation entre le fait, c'est-à-dire le fonctionnement, et la valeur, ici le bien-être, n'est pas accidentelle. Donc, par hypothèse, deux fonctionnements identiques correspondraient à deux états de bien-être identiques (mais on peut douter de la possibilité de deux fonctionnements identiques ; parions que la similitude devrait suffire ici).

La relation entre état et activités de la personne, d'une part, et le jugement évaluant son bien-être, d'autre part, est ce qui permet donc de distinguer l'approche de Sen de celle d'Arneson. Ce dernier affirme que le bien-être dépend des préférences de l'individu, tandis que Sen choisit de définir le bien-être d'un individu en fonction de la description indépendante de son état et de ses activités. Mais, avant de progresser plus avant dans notre réflexion, remarquons que les choses ne sont pas si simples : Arneson s'attache bien à préciser que ses préférences sont établies dans des conditions cognitives idéales, que nous avons rappelées dans la section précédente. Du côté de Sen, les fonctionnements constitutifs du bien-être font l'objet d'un

---

[397]SEN, Amartya ,*Commodities and Capabilities*, Amsterdam: North-Holland, 1985, p 37.

[398]ENGEL Pascal, *Introduction à la philosophie de l'esprit*, Paris, La Découverte, 1994, p. 38.

[399] ZUPPINGER, Thibaud, "La survenance, entre axiologie et anthropologie philosophique », *Raison Publique*, 27 avril 2010, http://www.raison-publique.fr/article293.html; Consulté le 31 août 2012.

jugement qui se doit d'être parfaitement informé et pertinent. Mais les conditions d'un tel jugement restent ambiguës, et cette ambiguïté peut être mise au compte du « flou conceptuel dans la relation de survenance » que signale Thibaud Zuppinger :

> *« Il existe une relation de dépendance sans réduction entre les faits et les propriétés axiologiques qui sont exprimés dans le jugement d'un sujet. Mais ce sujet est-il l'agent moral ou l'observateur extérieur ? Qui est tenu d'afficher une cohérence dans ses jugements ? Pour sortir de cette interrogation, peut-être est-il utile de penser* une *survenance* par *sujet impliqué. Les propriétés axiologiques exprimées peuvent varier d'une personne à l'autre, mais chacune est tenue de conserver une forme de cohérence pour les cas similaires. Sauver une personne est bien, jeter une vieille dame par terre est mal. En pratique, nous serions alors confrontés à des cas d'asymétrie morale. Untel peut agir moralement en écartant une personne du trajet d'une voiture, mais sera néanmoins jugé comme malveillant par une personne qui n'aura pas perçu le danger et ne percevra qu'une personne en bousculant une autre. La survenance n'assure pas la symétrie des jugements moraux. Elle devra donc être complétée par l'idéal régulateur de l'observateur omniscient. Cependant cet ajout n'est pas contenu analytiquement dans la relation de survenance et peut supporter une conception perspectiviste de la morale. »*[400]

De fait, il est extrêmement difficile de trouver dans la nature deux fonctionnements identiques : l'assimilation des nutriments par une personne ne sera jamais exactement celle d'une autre, et les médecins ne se lassent pas de découvrir la singularité des cas. Le risque est alors de tomber dans un relativisme analogue au relativisme préférentialiste, sauf qu'au lieu de dépendre de l'arbitraire de la préférence, il sera la résultante de la particularité du fait décrit. De là émerge spontanément la question : qui décrit ? Si la particularité de la description doit être attribuée au point de vue de celui qui décrit, nous ne sommes pas sortis du subjectivisme des préférences que l'approche par les capabilités prétend surmonter. Car, l'assimilation des nutriments est d'abord décrite par l'intermédiaire du patient, qui rend compte de sa digestion, de son sommeil et de son état de fatigue. Elle dépend du choix des examens à effectuer, choix susceptible de varier quelque peu selon le médecin, sa formation, sa perception des besoins du patient…C'est ainsi que T. Zuppinger peut écrire : « la survenance n'assure pas la symétrie des jugements moraux ». Et il suggère de la compléter soit par le point de vue d'un observateur omniscient, tout en précisant que la survenance n'implique pas ce point de vue. Sans doute est-ce trop demander que d'être omniscient pour diagnostiquer les fonctionnements et le bien-être qui en dépend. Mais des conditions cognitives adéquates peuvent suffire. Or, ces conditions pourraient *bel et bien être celles*

[400] *Ibidem.*

qu'Arneson impose à ses préférences idéalement considérées. En d'autres termes, la distinction entre l'approche de Sen et celle d'Arneson dépend crucialement de la nature des conditions cognitives de l'évaluation du bien-être des personnes, conditions qui restent à être identifiées en ce qui concerne l'approche de Sen.

Mais l'approche de Sen ne se limite pas aux fonctionnements, de même que celle d'Arneson ne se limite pas à la satisfaction des préférences informées. Tous deux adoptent une perspective basée sur les opportunités, les capabilités étant des ensembles d'opportunités de fonctionnements.

b) *Les fondements des capabilités : économie normative et néo-aristotélisme*

L'approche par les capabilités conceptualise l'avantage individuel. Remarquons cependant que certaines capabilités peuvent être pertinemment attribuées à des groupes plutôt qu'à des individus, notamment lorsqu'elles dépendent des droits dont ils jouissent ou de l'appartenance à une classe sociale, par exemple. D'autre part, comme le souligne justement Ingrid Robeyns, il faut distinguer individualisme éthique et individualisme méthodologique[401]. L'approche par les capabilités est individualiste d'un point de vue éthique, dans la mesure où, si elle conceptualise l'avantage individuel, c'est en vue de l'amélioration de celui-ci. Mais elle n'est en revanche pas individualiste d'un point de vue explicatif, puisqu'elle ne prétend pas que les individus sont la cause de leur situation actuelle ; au contraire, les capabilités relèvent d'une conception de la liberté comme liberté réelle et non comme liberté formelle, la première prenant en compte les conditions concrètes de réalisation des fins des agents, tandis que la seconde ne s'intéresse qu'à l'absence d'entraves ou d'interdictions, position qui se justifierait précisément par l'hypothèse selon laquelle l'individu contrôle l'ensemble des facteurs de réalisation de ses fins.

Si le cadre conceptuel normatif constitué par l'approche des capabilités est aujourd'hui largement utilisé tant en philosophie politique que dans les études de terrain en matière de développement humain, une certaine ambiguïté subsiste quant à ses fondements philosophiques. En effet, les capabilités sont nées des travaux d'Amartya Sen en économie du bien-être, travaux dont le point de départ est le cadre conceptuel commun à celle-ci et à la

[401] ROBEYNS, Ingrid, « The Capability Approach : A Theoretical Survey », *Journal of Human Development*, 6 : 1, 2005, p 107.

théorie du choix social, caractérisé par une structure d'agrégation de classements évaluatifs et l'utilisation du formalisme, dont l'avantage est de préciser le propos.

Mais elles ont également émergées des travaux de Martha Nussbaum, qui défend une anthropologie philosophique néo-aristotélicienne[402]. On peut, en s'appuyant sur les textes d'Aristote et l'interprétation réactualisée qu'en propose Nussbaum, tâcher de reconstruire les fondements aristotéliciens des capabilités. L'ontologie d'Aristote est une ontologie du mouvement. Chaque être appartenant au monde sublunaire – à la nature – tend spontanément vers une finalité, à sa voir sa propre réalisation[403]. Les êtres humains, comme les autres êtres, sont en mouvement, évoluent, en vue de la réalisation de leur fin propre, à savoir de leur accomplissement en tant qu'êtres humains. Ces mouvements, ces activités, sont la « tâche » propre de l'être humain, sa « fonction propre » (*ergon*).

> *« De même, en effet, que dans le cas d'un joueur de flûte, d'un statuaire, ou d'un artiste quelconque, et en général pour tous ceux qui ont une fonction ou une activité déterminée, c'est dans la fonction que réside, selon l'opinion courante, le bien, le « réussi », on peut penser qu'il en est ainsi pour l'homme, s'il est vrai qu'il y ait une certaine fonction spéciale à l'homme. »*[404]

Nous retrouvons ici la notion de fonctionnement, à ceci près que, dans le système aristotélicien, c'est l'essence de l'être humain qui détermine les fonctionnements constitutifs de sa nature, tandis que, dans la perspective moderne des capabilités, les fonctionnements pertinents pour les êtres humains, avec leurs différences de goûts, de valeurs ou de besoins, restent à déterminer. En effet, la conception de la valeur chez Aristote est essentialiste : les bons « fonctionnements » sont ceux qui sont propres à l'être humain, et l'exercice de la fonction propre est ce qui constitue le bien de l'être humain et par là-même son bonheur.

L'activité de choix, la dimension du volontaire, sont également caractéristiques du fonctionnement humain[405]. Le caractère volontaire d'une activité permet de distinguer celles qui sont vertueuses de celles qui ne le sont pas. Nussbaum affirme donc qu'il y a, dans les textes mêmes d'Aristote, des arguments en faveur d'une évaluation des capabilités plutôt que des fonctionnements en ce qui concerne l'avantage humain. La conception des capabilités que défend Nussbaum prend donc pour point de départ un travail d'histoire des idées, et non pas une réforme du cadre conceptuel de l'économie normative. Il s'en suit une différence

---

402 NUSSBAUM, Martha C., "Nature, Function, and Capability", *Oxford Studies in Ancient Philosophy*, suppl. vol. 1 (1988), pp. 145–84

403 Cf. ARISTOTE, *Physique*, 192b21.

404 ARISTOTE, *Ethique à Nicomaque*, trad. Tricot, Paris : Vrin, 1990, 1097b25-30.

405 Cf. ARISTOTE, *Ethique à Nicomaque*, Livre III.

stylistique évidente : les écrits de Nussbaum sur les capabilités sont interprétatifs et narratifs, tandis que ceux de Sen font appel au formalisme et à des concepts issus de la microéconomie. Jusqu'ici, cela n'entraîne pas l'incompatibilité des deux approches, qui pourraient être deux voies menant au même but. Nous verrons cependant que ces deux approches seront amenées à diverger de manière fondamentale lorsque sera venue l'heure de sélectionner une liste de capabilités et de fonctionnements pertinents.

Pour l'heure, nous nous contenterons de préciser que nous adoptons la conception de Sen. Dans la mesure où ce travail vise à élargir le cadre conceptuel de la théorie du choix social afin d'y intégrer les capabilités, il est plus cohérent de partir d'une définition des capabilités construite à partir des outils conceptuels de l'économie normative. Les capabilités et les fonctionnements ont été analysés et formalisés par Sen dans la monographie *Commodities and Capabilities*[406].

Une capabilité est un ensemble de fonctionnements qu'un agent peut réellement accomplir ; « réellement », parce que les capabilités dépendent à la fois des ressources dont l'agent dispose et de ses « facteurs de conversion », c'est-à-dire de sa capacité de convertir ces ressources en fonctionnements. Les « ensembles capabilités » représentent donc la liberté réelle qu'a un agent de choisir entre plusieurs « vies » possibles, plusieurs manières de réaliser ses fins. Plus précisément, ces « vies » sont des « combinaisons », ou des « vecteurs » (c'est-à-dire des ensembles d'éléments ordonnés et de même type) de fonctionnements.

Un fonctionnement est un « faire » ou un « être », un « mouvement » effectué en vue de la réalisation d'une fin. Un fonctionnement est donc dynamique et non pas statique ; sa description comprend le processus de réalisation du résultat, et pas le résultat seul.

Sen a formalisé les concepts de ressources, de fonctionnements, de bien-être et de capabilité de la manière suivante :

- x_i: vecteur de ressources possédées par la personne i (par exemple: le pain)
- *c(.) : fonction convertissant un vecteur de ressources en un vecteur de caractéristiques* de ces ressources (par exemple: les qualités nutritionnelles du pain, ou bien sa symbolique sociale et culturelle…)

[406]SEN, Amartya ,*Commodities and Capabilities,* Amsterdam: North-Holland, 1985.

- $f_i(.)$: fonction d'utilisation personnelle de la personne i, reflétant un usage possible des ressources qu'elle a à sa disposition (par exemple, son usage du pain dépendra de sa capacité à l'ingérer, à la digérer, de son métabolisme, de ses besoins nutritionnels, dependant eux-mêmes de sa corpulence, de son état de santé, de son activité physique,, etc)
- F_i: ensemble des fonctions d'utilisation f_i, que la personne i peut effectivement choisir.

Lorsqu'une personne *i* décide de réaliser une fonction d'utilisation particulière d'une ressource qu'elle a à sa disposition, elle accomplit un fonctionnement décrit par le vecteur b_i.

- b_i: fonctionnements effectivement accomplis par la personne i (elle a mangé le pain, et celui-ci l'a nourrie) : $b_i = f_i(c(x_i))$.

b est le "*being*" de la personne (son "être"). Le "*well-being*" est l'évaluation de "*b*". Ce n'est donc pas un état de satisfaction indépendant de l'exercice des fonctionnements de la personne, il depend de ces derniers, ainsi que nous l'avons vu lors de notre analyse de la relation de survenance entre les deux.

Si le choix de vecteurs de ressources de la personne i est limité à l'ensemble X_i,alors les vecteurs de fonctionnements faisables pour cette personne sont donnés par l'ensemble $Q_i(X_i)$, qui correspond à ses CAPABILITES:

- $Q_i(X_i) = [b_i |\ b_i = f_i(c(x_i))$, pour $f_i(.) \in F_i$ et pour $x_i(.) \in X_i]$

Cette expression peut se lire ainsi: pour un vecteur *X* donné de ressources, la personne *i* dispose de l'ensemble de capabilité *Q* qui comprend les différentes combinaisons de fonctionnements possibles tels que, chaque fonctionnement est lui-même fonction des facteurs de conversion de la personne et des caractéristiques de la ressource utilisée.

c) Classifier les capabilités

La possession des ressources constitue l' « opulence », mais non pas l'avantage. Un coffre plein de bijoux sur une île déserte rendra le marin égaré opulent, mais ne l'empêchera pas de mourir de faim. La fonction d'utilisation de ces ressources est déterminante pour l'avantage, et dépend des facteurs de conversion des ressources en fonctionnements dont disposent les agents. La nature de ces facteurs de conversion varie, et pourrait constituer un critère de classement des capabilités. Un premier classement est d'ailleurs proposé par Nussbaum, qui propose de distinguer les capabilités internes des capabilités externes. Les capabilités internes dépendent de facteurs de conversion « internes » à la personne – qui ne peuvent exister sans elle - , tels que les capacités intellectuelles, les émotions, la condition physique. Ceux-ci

dépendent à leur tour de l'éducation, de la santé et l'entourage. Les capabilités externes dépendent quant à elles des circonstances extérieures, que Nussbaum assimile globalement aux conditions politiques. Nussbaum précise cependant que la frontière entre les deux ne doit pas être surestimée, dans la mesure où la tâche du législateur consiste aussi à offrir les moyens d'éducation aux enfants ; d'autre part, la participation à la vie politique contribue à l'épanouissement du respect de soi et donc à l'accroissement des capabilités internes, par exemple. Ingrid Robeyns propose quant à elle de distinguer les facteurs de conversion personnels, les facteurs sociaux et les facteurs environnementaux.

Nous proposons ici une classification plus fine des facteurs de conversion. Le but de cette classification est double : d'une part elle permet de saisir la richesse de l'approche par les capabilités et d'éviter de la réduire à l'une ou l'autre de ces dimensions ; d'autre part, elle permet d'envisager une division du travail entre les disciplines susceptibles de pourvoir le théoricien des capabilités en informations sur l'état actuel des personnes et les possibilités d'améliorer leur avantage. Ici encore, les frontières sont poreuses et les facteurs de conversion s'influencent mutuellement. Nous distinguerons les facteurs biologiques et médicaux, les facteurs psychologiques, les facteurs cognitifs, les facteurs sociaux, les facteurs économiques, les facteurs politiques et les facteurs environnementaux.

Les facteurs biologiques et médicaux correspondent à la condition physique de l'agent et à son état de santé. L'assimilation des nutriments d'une ressource telle que le pain dépend du métabolisme de la personne, de ses dépenses énergétiques (qui varient selon son âge, sa corpulence, le fait d'être enceinte), de la présence de parasites, etc. L'attention portée à la condition physique des agents est très présente dans les premiers écrits de Sen sur les capabilités et ses travaux sur la famine. C'est cette attention qui semble avoir motivé dans un premier temps la prise de distance à la fois avec les approches basées sur l'évaluation de paniers de biens ou du revenu, d'une part, et les approches utilitaristes ou welfaristes d'autre part.

Les facteurs psychologiques correspondent à la capacité de chacun à éprouver des émotions appropriées à la réalisation d'une opportunité, ou à les maîtriser. Des affects tels que le respect de soi ou la confiance sont cruciaux. Les travaux de Rawls sur l'acquisition progressive du sens de la justice, sur les bases sociales du respect de soi[407], ou de Martha

[407] RAWLS, John, *A Theory of Justice*, The Belknap Press of Harvard University Press, 1971; traduction française par Catherine Audard sur un texte révisé par l'auteur en 1975: *Théorie de la Justice*, Paris, Editions du Seuil, 1987, 1997. Préface de la traduction française par John Rawls. Cf. Livre III, § 70-75.

Nussbaum sur les émotions démocratiques, sont des exemples d'analyses du rôle joué par les émotions pour l'avantage individuel et son amélioration[408].

Les facteurs de conversion cognitifs sont cruciaux, pour les capabilités en général et pour la spécification de l'approche que nous allons proposons dans les sections suivantes. Ces facteurs incluent l'information factuelle dont disposent les personnes, mais également l'aptitude à délibérer, à calculer, à analyser et à synthétiser, à concevoir un éventail large de solutions, ainsi que la connaissance des ressources informationnelles et la capacité d'en user, compétences de première importance dans une société où la technologie est présente dans la plupart des dimensions de notre existence. La conception des préférences que nous défendons dépend en effet des capabilités cognitives, et à ce titre capabilités et préférences s'entretiennent réciproquement.

Les facteurs sociaux de conversion sont constitués par les rapports sociaux. Ces rapports peuvent être de coopération ou de conflit. Ils peuvent soit accroître les capabilités (solidarité, par exemple), soit les restreindre (domination…). La spécification des capabilités en termes de non-domination proposée par Philip Pettit concerne ainsi cet aspect particulier de la conversion des ressources en fonctionnements[409] : la capacité actuelle qu'a une personne d'interférer avec les choix d'une autre personne, et ce non pas en vue de contribuer à la réalisation de ses fins propres, peut empêcher cette personne de convertir ses ressources en fonctionnements[410]. La réalisation de nos fins ne devrait pas dépendre d'une faveur, mais doit pouvoir avoir lieu quelles que soient les volontés des autres personnes ou des autres groupes de personnes[411].

---

[408]NUSSBAUM, Martha, *Not for profit: Why Democracy needs the humanities*, Princeton University Press, 2010. Traduction française Solange Chavel, *Les émotions démocratiques*, Paris: Flammarion, 2011.

[409] Pettit rattache la conception de la liberté comme non-domination qu'il défend à la tradition républicaniste en philosophie politique. Cf. PETTIT, Philip, *Republicanism: A Theory of Freedom and Government*, Oxford: Clarendon Press, 1997. Trad. Fr. Patrick Savidan et J.-F. Spitz, *Républicanisme*, Paris: Gallimard, 2004.

[410]PETTIT, Philip, "Capability and Freedom: A Defence of Sen", *Economics and Philosophy* 17, 2001, p. 1-20.

[411]Sen refuse toutefois de remplacer les capabilités par la conception républicaniste de la liberté comme non-domination. Son argument est le suivant : la liberté comme non-domination discrimine certaines situations que l'approche par les capabilités considère comme équivalentes. Par exemple, une personne handicapée que des travailleurs sociaux aident à se déplacer jouit des mêmes capabilités qu'une personne qui a des employés assurant les mêmes tâches, mais, selon Sen, elle ne dispose pas du même degré de liberté comme non-domination, puisque sa liberté dépend de l'existence des aides sociales dans le premier cas, et non dans le second. L'exemple de Sen n'est pas très heureux ici : en effet, Pettit pourrait répondre que cette personne handicapée, dans la mesure où elle est également une citoyenne active et pouvant participer à la décision d'implanter un système d'aide sociale, soit directement, soit par l'intermédiaire de ses représentants, ne dépend pas du bon vouloir de ces travailleurs sociaux ou de ceux qui les emploient. Un exemple plus approprié pourrait être le suivant : la personne handicapée est aidée par des travailleurs embauchés par un millionnaire philanthrope. Dans ce cas, en effet, l'exercice du fonctionnement « se déplacer » dépend d'une faveur que la personne handicapée ne peut pas maîtriser. Mais a-t-elle pour autant le même niveau de capabilité que dans le cas où elle emploie elle-même lesdits travailleurs ? Rien n'est moins certain, puisque les rapports sociaux qui constituent les facteurs de conversion de ressources en fonctionnements ne sont pas les mêmes : dans un premier cas, la personne est dominée par l'opulence de son bienfaiteur, tandis que dans le second elle se trouve elle en

Les facteurs économiques de conversion des ressources en fonctionnements dépendent de l'accès à un marché, des prix fixés ou des opportunités d'échange. Ces facteurs sont mis en évidence dès les premiers travaux de Sen sur les famines, et ce avant même que le concept de capabilité n'apparaisse dans ses écrits. Sen proposait un effet de se concentrer sur les possibilités d'accès [*entitlements*] des personnes aux ressources, plutôt que sur ces ressources, afin de décrire et expliquer les famines :

> *« Dans une économie pleinement dirigée, chaque personne* i *peut seulement obtenir le panier de ressources particulier qui lui est attribué. Dans une certaine mesure cela se passe dans la plupart des économies, par exemple avec les résidents des maisons de retraite ou des hôpitaux psychiatriques. Toutefois, il y a généralement un menu – qui peut être large –dans lequel on peut choisir.* E_1 *est l'ensemble des possibilités d'accès d'une personne* i *dans une société donnée, dans une situation donnée, et il consiste en un ensemble de paniers de ressources alternatifs, la personne pouvant décider d'obtenir n'importe lequel. Dans une économie avec propriété privée et échange sous la forme de commerce (échange avec les autres) et de production (échange avec la nature),* E_1 *peut être caractérisé comme dépendant de deux paramètres, à savoir les* dotations *de la personne (le panier qu'elle possède) et l'*application des possibilités d'accès à l'échange *(la fonction qui spécifie l'ensemble des paniers de biens que la personne peut atteindre avec chaque dotation).Par exemple, un paysan a sa terre, sa force de travail, et quelques autres ressources, qui constituent sa dotation. En partant de cette dotation, il peut produire un lot de nourriture qui sera le sien. Ou, en vendant sa force de travail, il peut obtenir un salaire avec lequel il achètera des biens, incluant de la nourriture. Ou il peut faire pousser des cultures commerciales et les vendre pour acheter de la nourriture et d'autres biens. Il y a beaucoup d'autres possibilités. »*[412]

Les possibilités d'accès à l'échange ou de consommation simple de la production sont autant de facteurs de conversions, qui seront captés par la suite par le concept de capabilité. Il est intéressant de souligner que la perspective de l'économie normative ne valorise pas un modèle économique plutôt qu'un autre : l'économie planifiée et l'économie de marché, la propriété collective et la propriété privée, ne sont pas considérées en soi comme bonnes ou mauvaises, mais sont évaluées à l'aune de leur aptitude à accroître les possibilités de conversion de ressources en fonctionnements ou en consommation dont jouissent les agents.

Les facteurs politiques de conversion des ressources en fonctionnements comprennent la jouissance et l'exercice de droits comme le droit de vote, la liberté d'expression, d'opinion, le droit de grève, l'égale participation à la vie politique...Ici, l'approche par les capabilités est

---

position de domination économique et sociale. Nous voyons bien par ailleurs que l'ensemble de ses opportunités est plus limité dans le premier cas que dans le second, puisque c'est le bienfaiteur qui décide des modalités de l'aide qu'il fournit.

[412]SEN, Amartya, *Poverty and Famines*, Oxford University Press, 1981.Traduction libre..

beaucoup plus déterminante que dans le cas de l'économie (à moins que l'on ne manque de connaissances en économie pour être aussi déterminants ?) : Sen défend clairement la démocratie comme régime politique seul susceptible de garantir un niveau de capabilité acceptable. Cela tient à deux caractéristiques de la démocratie (qui s'entretiennent mutuellement) : d'une part, la détermination des actions publiques par la décision collective (directement ou indirectement, par l'élection de représentants), ce qui a pour effet de ne jamais faire oublier la question de l'amélioration de l'avantage de chacun dans le débat public ; d'autre part, la garantie de certains droits politiques permet d'éviter la restriction des opportunités dont jouissent les agents à une seule, aussi désirable soit-elle : les capabilités sont, ne l'oublions pas, des libertés. Néanmoins, le terme « démocratie » englobe un ensemble de définitions d'institutions et de pratiques politiques et sociales assez large, et il serait intéressant, dans un travail de recherche ultérieur, de spécifier la ou les versions de la démocratie les plus susceptibles de favoriser l'accroissement des capabilités en diversifiant les possibilités de conversion des ressources en fonctionnements. Parmi les paramètres de la démocratie que l'on peut faire varier, figurent l'échelle de la décision (globale, nationale ou locale ?), le caractère direct ou indirect de la décision (par l'intermédiaire de l'élection de représentants), le degré de participation des citoyens individuels à la vie politique (qui détermine le rôle joué par cette vie de citoyen dans l'épanouissement personnel, et le choix d'un modèle « ancien » participatif de démocratie, ou plutôt moderne et libéral), le type de collectivité où la prise de décision démocratique a lieu (s'agit-il d'une collectivité politique, des associations, des syndicats, des entreprises, de la famille, des communautés religieuses ou encore des jurys ?), la forme de la prise de décision (il faudra déterminer si l'on n'a recours qu'au vote, qu'à la délibération, ou aux deux conjointement).

Les facteurs environnementaux incluent le climat et la location géographique. Leur caractère favorable est très lié aux décisions sociales, politiques et économiques : les classer à part ne doit donc pas laisser penser qu'ils relèveraient exclusivement de hasard naturels et de phénomènes non maîtrisables.

Capabilités et acte de choix : fonctionnements affinés ou ensembles d'opportunités ?

La conversion des ressources en fonctionnement n'est qu'un versant de la capabilité. L'autre versant est le choix d'une combinaison de fonctionnements. Nous avons cité l'importance de l'acte de choix dans l'interprétation des capabilités que propose Martha Nussbaum. Mais on

peut concevoir cet acte de choix de deux manières différentes, impliquant des présuppositions philosophiques différentes. L'une revient à travailler sur des fonctionnements affinés (*refined functionings*), tandis que l'autre conserve la définition des capabilités comme ensembles d'opportunités de fonctionnements.

Sen lui-même a proposé une méthode alternative pour incorporer la dimension de la liberté dans l'approche par les capabilités[413]. Le problème avec les capabilités définies comme des ensembles d'opportunités est qu'elles requièrent une quantité importante d'information et seraient donc difficile à mettre en œuvre concrètement. D'autre part, en se concentrant sur les ensembles d'opportunités, y compris les opportunités contrefactuelles, on court le risque de perdre de vue les accomplissements[414]. Les fonctionnements affinés sont des fonctionnements décrits de manière suffisamment précise pour que le fait qu'ils ont été librement choisis puisse être inclus dans cette description. Prenons l'exemple du jeûne. Une personne qui jeûne a la capabilité d'exercer le fonctionnement « manger à sa faim », mais préfère ne pas l'actualiser, pour des raisons religieuses, politiques ou thérapeutiques. La distinction entre capabilité et fonctionnement permet de saisir la différence entre le jeûne et la famine. Mais si l'on décrit le fonctionnement « manger à sa faim » avec une formule plus précise, telle que « manger librement à sa faim », il n'est plus nécessaire d'envisager le classement des combinaisons de fonctionnements alternatifs pour prendre en compte l'acte de choix dans l'évaluation de l'avantage. Dans la perspective aristotélicienne, l'exercice des fonctionnements doit être proprement « humain ». Si l'on considère que la réalisation libre de ces activités seule satisfait le critère normatif de l' « humain », alors les fonctionnements sont d'ores et déjà des fonctionnements affinés.

L'approche de Sen, avec laquelle nous travaillons, identifie les capabilités à des ensembles d'opportunités de fonctionnements. Chaque personne dispose de plusieurs combinaisons de fonctionnements possibles, et de la « puissance » nécessaire à leur réalisation (que cette puissance comprenne les dimensions émotionnelles et cognitives de l'acte de choix ou qu'elle englobe la capacité de résister à des oppositions extérieures). Cette approche est conceptuellement plus ambitieuse, puisqu'elle implique que la description des capabilités devrait inclure également les combinaisons de fonctionnements que la personne n'a pas choisi de réaliser pour évaluer le degré de liberté réelle dont elle jouit. De plus, l'évaluation des

[413]SEN, A., MUELLBAUER, J., KANBUR, R., HART, K., et WILLIAMS, B., *The Standard of Living*, Cambridge: Cambridge University Press, 1987.

[414]FLEURBAY, Marc, « Capabilities, functionings and refined functionings », *Journal of Human Development*, 7 (3), 2006, p 299-310.

fonctionnements ne dépendrait plus de leur description qualitative, mais de leur position dans le classement d'opportunités établi par la personne. Mais la richesse informationnelle n'est-elle pas ce qui importe en premier lieu pour l'approche par les capabilités ? Il serait d'ailleurs intéressant sur le plan empirique de sonder les personnes sur leur perception des options contrefactuelles afin de s'assurer qu'elles se représentaient adéquatement l'ensemble d'opportunité dont elles disposaient au moment du choix[415].

La distinction entre fonctionnements affinés et capabilités comprises comme ensembles d'opportunités implique deux compréhensions différentes de l'acte de choix.

Les fonctionnements affinés n'identifient pas nécessairement le choix à la détermination d'une option plutôt que d'une autre : la thèse de la liberté de choix n'exige pas l'existence d'options contrefactuelles pour être valide. Il suffit que l'agent ait conscience de sa participation active à l'exercice du fonctionnement pour pouvoir dire qu'elle l'a choisi librement. Avec les fonctionnements affinés, l'agent exerce un contrôle sur ses actions que l'on peut définir comme une orientation ou un accompagnement. Il suffit que l'agent se représente les raisons – au sens large du terme - d'exercer ce fonctionnement comme siennes, que l'exercice du fonctionnement soit une réponse à ses raisons ou à ses désirs[416]. L'existence d'opportunités alternatives de fonctionnements n'est pas requise.

Les ensembles capabilités sont en revanche définis par l'existence d'opportunités alternatives. Le contrôle exercé par l'agent sur ses activités est donc corrélé à sa capacité d'avoir agi autrement. Les ensembles capabilités définissent donc l'acte de choix comme la détermination d'une option plutôt que d'une autre[417][418]. Cette conception des capabilités nous semble plus prometteuse pour l'évaluation sociale et notamment pour la compréhension de la liberté en philosophie politique pour la raison suivante : les combinaisons alternatives de fonctionnements, même lorsqu'elles ne sont pas réalisées, montrent, par le fait même que

[415] Dans un document de travail sur les capabilités, sur lequel notre présentation s'appuie en partie, Erik Schokkaert mentionne l'existence de quelques travaux à ce sujet. Cf. SCHOKKAERT, Erik, "The Capability Approach", document de travail, Department of Economics, KU Leuven, ces0734, Mars 2007; ANAND, P. et VAN HEES, M., "Capabilities and Achievements: an empirical study", *Journal of Socio-Economics*, 35: 268-284.

[416] Sur les différentes conceptions de contrôle sous-tendues par la discussion sur le libre arbitre, cf. FISCHER, J. M., *The Metaphysics of Free Will*, Wiley-Blackwell, 1995.

[417] Cf. ARISTOTE, *Ethique à Nicomaque*, Livre III.

[418] Si les lecteurs s'inquiètent de voir les capabilités s'engouffrer dans les méandres du libre arbitre, on peut les rassurer en soulignant qu'il suffirait que les agents puissent se représenter les différentes combinaisons de fonctionnements possibles pour pouvoir parler d'ensembles capabilités, sans que cependant il soit nécessaire d'affirmer qu'un autre monde était réellement possible au moment où les agents en question ont effectué leur choix.

l'agent y a accès, qu'il dispose d'une maîtrise suffisante de son environnement et des circonstances tant sociales que politiques ou économiques de son existence. Cette attention portée au contexte est cruciale pour la description des facteurs de conversion des ressources en fonctionnements : lorsque Pettit rattache les capabilités à sa conception de la liberté comme non-domination, c'est en raison de ces avantages pour l'analyse. Même si une personne se trouve satisfaite de sa situation et tend à penser que celle-ci correspond à ses raisons et à ses désirs, une évaluation de son avantage ne serait pas complète si elle n'incluait pas l'information suivante : si les désirs de cette personne avaient été différents, aurait-elle pu mener sa vie autrement, ou se trouve-t-il que, par chance, ses désirs s'accordent aux circonstances ? Cette brève analyse montre que, si les ensembles capabilités échappent au problème des désirs adaptatifs, il n'en va pas évidemment de même pour les fonctionnements affinés.

Les finalités : bien-être ou liberté d'agent ?

L'analyse des fondements aristotéliciens des capabilités et des fonctionnements a montré que, ce qui fait la valeur d'une capabilité ou d'un fonctionnement donné, c'est sa contribution aux fins poursuivies par l'agent. La perspective d'Aristote était essentialiste : la fin en question est simplement la réalisation de la nature propre de l'être humain. Mais les théoriciens modernes des capabilités refusent cette conception perfectionniste de la valeur des capabilités, car elle s'accommoderait mal du « fait du pluralisme » identifié par Rawls[419]. Quelles sont alors les fins poursuivies par les agents selon l'approche par les capabilités ?

Le système des fins proposé par Sen est exposé ici dans le Tableau 2, section 3.1.1.. On voit que Sen distingue deux types de fins : le bien-être [*well-being*] et les fins d'agent. Le bien-être concerne l'avantage personnel, tandis que les fins d'agent représentent ce que l'agent souhaite voir arriver, indépendamment du souci de son bien-être. Cette distinction conduit, comme le tableau 2 le montre bien, à la définition de quatre catégories d'évaluation : les capabilités (des opportunités de bien-être [*well-being*], les fonctionnements (l'accomplissement du bien-être par des états et des activités), les libertés d'agent et les accomplissements d'agent. Toutefois, comme nous l'avons rappelé, les frontières entre capabilités et libertés d'agent sont floues. Par comparaison, l'approche de Martha Nussbaum n'adopte pas cette distinction : Nussbaum juge que la distinction entre capabilité et fonctionnement suffit pour décrire l'importance de la liberté et des valeurs de l'agent[420]. Et même Sen est parfois ambigu à ce sujet :

> *« La liberté et général et la liberté d'agent en particulier font partie du pouvoir effectif qu'une personne a, et ce serait une erreur de voir la capabilité, liée avec ces idées de liberté, seulement comme une notion de l'avantage humain : elle présente aussi un intérêt central pour comprendre nos obligations. Cette considération conduit à une opposition majeure entre bonheur [*happiness*] et capabilité en tant qu'ingrédients informationnels de base dans une théorie de la justice, puisque le bonheur ne génère pas d'obligations comme le font inévitablement les capabilités, si l'argument sur la responsabilité du pouvoir effectif est reconnu. Ceci est, à cet égard, une différence significative entre le bien-être [*well-being*] et le bonheur, d'un côté, et la liberté et la capabilité, de l'autre côté. »*[421]

[419]RAWLS, John, *Political Liberalisme*, New York, Columbia University Press, 1993; traduction française par Catherine Audard: *Libéralisme politique*, Paris, Presses Universitaires de France, « Quadrige », 1995, 2007.

[420]NUSSBAUM, Martha, *Women and human Development: the capabilities Approach*, Cambridge: Cambridge University Press, 1999. *Femmes et développement humain*, traduit de l'anglais (USA) par Camille Chaplain, Paris, Edition Des Femmes-Antoinette Fouque, 2008, p 31.

[421]SEN, Amartya K., *The Idea of Justice*, London: Allen Lane, 2009, p 271. Traduction libre..

Ce passage établit une coupure nette entre capabilité et bien-être, et « range » la capabilité du côté de la liberté.

La distinction entre bien-être et capabilité, d'une part, et fins et liberté d'agent, d'autre part, crée deux difficultés suffisamment important pour appeler à y renoncer, ainsi que le suggère Nussbaum.

La première difficulté peut être illustrée par l'exemple suivant : supposons une jeune femme, Elena, fortement impliquée dans une cause politique. Elena fait la grève de la faim pour obtenir la libération de prisonniers politiques. A première vue, cette grève de la faim nuit à son bien-être : Elena met son existence en danger. En revanche, la grève sert sa liberté d'agent, puisqu'Elena souhaite voir ces prisonniers libérés, quelles que soient les peines qu'elle doit endurer. La liberté d'agent serait donc celle de faire la grève de la faim, tandis que la capabilité serait la liberté de pouvoir s'alimenter. Ici, il semble que Nussbaum ait raison : Sen multiplie les entités sans nécessité, puisque la distinction entre les ensembles capabilités et les fonctionnements suffit pour décrire des cas tels que celui d'Elena. D'autre part, cette distinction présuppose une définition forte du bien-être, qui n'est jamais clairement explicitée. En effet, si le bien-être selon Sen inclut par exemple une durée de vie longue – il n'y aurait là rien de surprenant -, quel type de décision concernant l'euthanasie préconiserait, par exemple, l'approche par les capabilités ? Les personnes qui demandent un tel geste de la part du médecin font état de leur souhait de cesser de souffrir ; elles ne demandent pas l'euthanasie pour des raisons politiques. Autrement dit, la demande et le refus de l'euthanasie s'appuient tous deux sur une définition du bien-être, l'une comme arrêt de la douleur et l'autre comme durée de vie longue[422]. La construction d'une définition du bien-être dépend en fait pour une bonne part des valeurs individuelles des agents comme – surtout – des conceptions sociales partagées. Comme le fait remarquer Peter Singer, les sociétés contemporaines attachent une valeur à la vie qui peut conduire parfois à des prises de décision ignorant toute autre sorte de norme : Singer donne l'exemple de Peggy Stinson, une institutrice américaine qui donna naissance en 1977 à Andrew, un prématuré de vingt-quatre semaines. Bien que les parents aient dit accepter la possibilité que l'enfant ne puisse pas survivre, les médecins entreprirent de le maintenir en vie. Cela coûta très cher (l'équivalent d'un million de dollars aujourd'hui), et Andrew devait être gravement handicapé toute sa vie ; les médecins dirent aux parents

[422] Les arguments sont ici simplifiés à l'extrême. Le débat sur l'euthanasie prend en compte d'autres facteurs tout aussi importants que le bien-être, comme l'importance d'un consentement réfléchi, la dignité, la valeur de la vie en tant que telle, la légitimité de l'Etat ou du médecin pour accomplir le geste de supprimer une vie…

qu'Andrew devait souffrir horriblement à chaque respiration[423].Toutefois, pendant longtemps, les sociétés ont attaché davantage de valeur à l'honneur qu'à la vie. La conception du bien-être est donc enchevêtrée avec les valeurs de l'agent, valeurs qu'elle tient de la société dans laquelle elle vit, de ses engagements ou de ses délibérations morales. Distinguer le bien-être des fins d'agent requiert une théorie substantielle de ce qu'est le bien-être humain, théorie que l'approche par les capabilités n'est pas prête à revendiquer. D'autre part, si le bien-être est donné, l'on ne voit pas très bien dans quelle mesure la distinction entre capabilité et fonctionnement et pertinente.

La deuxième difficulté tient à la définition des fins d'agent. Lorsque l'on distingue celle-ci très nettement du bien-être, l'on ne voit pas comment les fins d'agent pourraient ne pas, à un moment donné, englober des désirs adaptatifs. Supposons une mère de famille qui a renoncé à la perspective de travailler à l'extérieur de son foyer, du fait des croyances et des valeurs propres à la société dans laquelle elle vit, croyances et valeurs qui attribuent à la femme des fonctions propres à son genre. Cette mère de famille pourrait à bon droit affirmer que ces valeurs sont les siennes ; lorsqu'un observateur ferait remarquer que sa position sociale la prive d'un revenu ou d'opportunités d'exercer des activités épanouissantes, elle répondrait que cela résulte de son engagement pour ces valeurs familiales. De même que le jeûne d'Elena la conduit à négliger son bien-être au nom de valeurs politiques, le choix de se consacrer à son foyer reviendrait à sacrifier ce même bien-être à des valeurs familiales. Dès lors, il apparaît qu'un traitement des fins d'agent séparé de celui des capabilités fait revenir par la petite porte le fameux problème des désirs adaptatifs et de l'aliénation.

Nous suivrons donc Nussbaum sur ce point, en faisant du bien-être comme des fins d'agent les finalités poursuivies par les personnes dont on évalue les capabilités et les fonctionnements.

Une typologie des fins pouvant donner de la valeur à un ensemble capabilités et être réalisées par l'exercice de fonctionnements, comprendrait donc les éléments suivants (il est important de souligner que ceux-ci ne s'excluent pas) :

(4) Le bien-être objectif [*well-being*] peut être défini à partir d'une conception plus ou moins substantielle de la « vie bonne », ou d'une vie digne d'un être humain, voire « digne d'un animal ». S'il s'agit de domaines spécifiques, ce bien-être peut être défini

[423]SINGER, Peter, *Practical Ethics,* Cambridge University Press, 1979, 2011[3], p 72.

par un consensus d'experts (par exemple, des médecins s'accordant sur la durée de vie « digne d'un être humain en 2012 ») ;

(5) Le bien-être « subjectif » [*Welfare*] correspond à des états mentaux de satisfaction, indépendamment de la condition objective de l'agent. L'évaluation des capabilités englobe cette dimension de « bonheur », à condition d'être d'une qualité cognitive suffisante pour que la recherche de ce type de satisfaction n'affecte pas le bien-être objectif de l'agent.

(6) Les fins individuelles d' « agent » sont déterminées librement par l'individu, selon sa conception privée du bien ; si ces fins sont communes, néanmoins, l'adhésion individuelle est essentielle ; les capabilités tiennent ces « fins d'agent » de la tradition libérale moderne et contemporaine ;

(7) Les fins collectives d'agent sont les fins auxquelles l'individu souscrit non pas par son jugement mais du fait même de sa « nature sociale » ou de son appartenance de naissance à une communauté donnée, qu'elle soit politique ou culturelle ; l'approche par les capabilités peut aussi inclure ce type d'analyse, par exemple en faisant du fonctionnement « participer activement à la vie de la Cité » une composante importante d'un ensemble capabilités ayant de la valeur[424]. De ce point de vue, les individus s'épanouiraient parallèlement à la collectivité dont ils sont les membres constitutifs.

(8) Une dernière finalité possible serait l'accroissement seul des capabilités : le fait même de jouir de capabilités serait à la fois le moyen et la fin de l'être humain, constitutif de son bien-être, sans qu'il soit besoin de définir une finalité extérieure à la détermination et à l'exercice des fonctionnements[425].

Le tableau 7 récapitule et schématise la conception des capabilités que nous adoptons pour la suite de notre argumentation.

Environnement :
- Environnement naturel
- Normes sociales et légales
- Relations interpersonnelles

[424] Cette possibilité théorique nuance l'individualisme éthique des capabilités.

[425] Des travaux récents tentent d'établir des parallèles entre l'approche par les capabilités et la philosophie de Spinoza. Cf. BUZAGLO, Jorge, « Capabilities : from Spinoza to Sen and Beyond », *Post-Austistic Economics Review*, 20-21, 2003.

- Formation des préférences

↘

Ressources[426]→Facteurs individuels →Ensemble **capabilité** →**Choix**[427]→**Fonctionnements**→ Accomplissement des **fins** de conversion

Tableau 7 : le processus de réalisation des capabilités.

[426] Les théories de la justice sociale basées sur les ressources ou les biens comme celles de Rawls ou de Dworkin incluent dans la liste ce que les théoriciens des capabilités classeraient parmi les facteurs de conversion : les talents, les libertés politiques, les bases sociales du respect de soi-même, etc.

[427] Le choix dépend ici strictement de la personne ; tous les facteurs susceptibles de l'influencer indépendamment du désir de la personne qui choisit sont déjà pris en compte dans la description des facteurs de conversion et de l'ensemble capabilité.

Chapitre 7 : Vers une approche préférentialiste de l'évaluation des capabilités.

Les capabilités sont une métrique de l'évaluation sociale, et l'approche par les capabilités n'offre pas de critère de distribution ou de principe d'agrégation. Le problème du poids respectif de chaque capabilité, ainsi que des capabilités de chaque personne, reste entier à cette heure.

> *« La thèse fondamentale, c'est que, pour évaluer le bien-être, les objets de valeur sont les fonctionnements et les capabilités. Elle n'implique ni que les capabilités de tout ordre sont d'égale valeur, ni que toute capabilité, même si elle est tout à fait étrangère à sa vie, doit forcément peser quelque chose dans l'évaluation du bien-être d'une personne. »*[428]

Le choix de la métrique n'implique pas une spécification des critères d'agrégation ou de distribution. Autrement dit, l'approche par les capabilités doit être complétée par une réflexion sociale sur la valeur respective de chaque capabilité et les priorités à établir en matière d'amélioration de l'avantage. C'est un problème crucial, comme le souligne Nussbaum :

> *« Je veux à présent me concentrer sur ce que je considère comme le problème le plus urgent pour une telle conception [l'approche par les capabilités], un problème qui doit être résolu, si la conception est appelée à avoir suffisamment de contenu pour être évaluée. Le problème est le suivant : quels sont les fonctionnements dont la Cité devrait s'occuper, et comment arrivons-nous à cette liste de fonctionnements ?»*[429]

L'hypothèse que nous souhaitons défendre en réponse à ce double problème philosophique est la suivante : l'approche par les capabilités a été à l'origine élaborée à partir du cadre conceptuel de l'économie normative et de la théorie du choix social, et ce serait continuer cette filiation que de traiter le problème de l'évaluation, de l'agrégation et de la liste des capabilités et des fonctionnements pertinents comme un problème de choix social. Nous nous sommes attachée à explorer les fondements philosophiques et les possibilités d'usage de ce cadre dans ce but. Un cadre de choix social assoupli peut accueillir la richesse informationnelle des capabilités. En des termes non techniques, le problème de l'évaluation respective de chaque capabilité et de la construction d'une liste, avec d'éventuelles priorités,

[428]SEN, Amartya, *Inequality Reexamined*, Oxford University Press, 1992 ; traduction française par Paul Chemla : *Repenser l'inégamité*, Paris, Editions du seuil, 2000, p 74.

[429]NUSSBAUM, Martha C., "Nature, Function, and Capability", *Oxford Studies in Ancient Philosophy*, suppl. vol. 1 (1988), pp. 145–84. Traduction libre..

est du ressort de la décision collective et de la prise en compte des intérêts bien compris de chaque membre de la société. Une théorie de la justice des capabilités, si elle doit être, serait dès lors une théorie démocratique et comparative des capabilités.

Nous exposerons dans un premier temps en des termes précis le problème de l'évaluation des capabilités et un aperçu de l'état de l'art dans ce domaine (7.1.). Notre deuxième temps présente des arguments intuitifs en faveur d'une approche de l'évaluation basée sur le concept de préférences pour les fonctionnements (7.2.).

7.1. Le problème de l'évaluation et de la liste des capabilités

Les capabilités sont une métrique multidimensionnelle de l'évaluation du bien-être et de la justice sociale. Cette métrique établie, nous avons besoin d'un schème de répartition. Une théorie de la justice basée sur les capabilités peut par exemple rechercher l'égalisation de niveau de chaque capabilité, ou encore viser la maximisation des capabilités, en partant du principe que :

> *« Avoir davantage de* chaque *capabilité pertinente ou de* chaque *fonctionnement pertinent est une amélioration claire et nette, et ce point est décidable sans qu'il soit besoin de se mettre d'accord sur les poids* relatifs *à attacher aux différents fonctionnements et capabilités. »*[430]

On peut aussi envisager des critères de répartition mélangeant égalisation et maximisation, tels que la maximin rawlsien, formulé en termes de biens premiers plutôt que de capabilités : une société juste serait celle qui élèverait en priorité le niveau de capabilité des plus défavorisés. Après avoir exposé la structure du problème de l'évaluation et de la liste (7.1.1.), nous présenterons les solutions ou les ébauches de solution de Sen (7.1.2.), de Nussbaum (7.1.3.) ainsi que des propositions plus récentes (7.1.3.).

[430]SEN, Amartya, *Inequality Reexamined*, Oxford University Press, 1992 ; traduction française par paul Chemla : *Repenser l'inégamité*, Paris, Editions du seuil, 2000, p 75.

7.1.1. Le problème de l'évaluation et le problème de la liste : dépendances et différences

Mais, quel que soit le schème de répartition que nous choisissons, nous devons d'abord assigner une valeur à chaque capabilité. Nous devons savoir quels ensembles capabilités ont la même valeur si nous sommes égalitaristes ; nous devons savoir quelle capabilité est meilleure que quelle autre si nous souhaitons maximiser les capabilités globales de la société. Or, les ensembles capabilités peuvent inclure de multiples fonctionnements : vivre longtemps, jouer au football, entretenir des relations amicales…Certaines semblent intuitivement plus importantes que d'autres. Tel est le problème de l'évaluation : dans une société avec un individu, des ressources infinies et un seul but, il n'y aurait pas un tel problème. Ce problème apparaît dans les circonstances de la justice, qui comprennent la rareté relative des ressources, l'altruisme limité (voire l'égoïsme unilatéral), et le pluralisme des valeurs[431]. L'approche par les capabilités doit donc être complétée par des décisions d'arbitrage entre différentes dimensions de l'avantage afin de spécifier ses recommandations. Comment décider de l'importance relative de chacune des dimensions de l'existence humaine ? Quelles capabilités devraient être garanties ou accrues en priorité ? Quelle procédure rationnelle permet de construire un indice de capabilités ?

Le problème de l'évaluation correspond à la question de la détermination de la procédure appropriée pour assigner une valeur à chaque capabilité, établir des priorités et prendre des décisions d'arbitrage. Le problème de la construction d'une liste concerne le contenu de l'indice des capabilités : quelles capabilités seront retenues, quelles capabilités seront exclues ?

Les liens entre le problème de l'évaluation et celui de la liste sont d'abord logiques. Le problème de l'assignation de valeurs respectives à différentes dimensions de l'existence humaine est lié au problème de la création d'une liste de capabilités pertinentes ou fondamentales pour l'évaluation sociale. En effet, dès lors que l'on assigne une valeur non nulle à une capabilité, celle-ci peut faire partie d'une liste de capabilités de valeur. Si les capabilités différentes ont des valeurs de grandeurs différentes, la liste comprendra également

[431]Cf. RAWLS, John, *A Theory of Justice*, The Belknap Press of Harvard University Press, 1971; traduction française par Catherine Audard sur un texte révisé par l'auteur en 1975: *Théorie de la Justice*, Paris, Editions du Seuil, 1987, 1997, § 22 et VAN PARIJS, Philippe, *Qu'est-ce qu'une société juste ?*, Paris, Le Seuil, 1991, chap. 10.

des priorités. Mais ils concernent également les fondements éthiques de la construction de l'indice des capabilités :

> *« Il y a deux questions ici, étroitement liées : une question de contenu et une question épistémologique. Elles sont étroitement liées parce que nous serons sceptiques, et pour cause, à l'égard d'une procédure de sélection si le contenu qu'elle génère ne correspond pas à nos intuitions et à nos jugements ; et d'un autre côté nous serons sceptiques, et pour cause encore, à l'égard d'une liste de fonctionnements si nous remarquons qu'elle ne pourrait être générée par une procédure qui nous semble être un exemple de rationalité. »*[432]

La procédure de génération de la liste définit donc la valeur relative de chaque capabilité, les modes de répartition, détermine les décisions d'arbitrage, et garantit la rationalité de l'évaluation des capabilités. La validité de la liste des capabilités dépend de la validité de la procédure de choix de ces capabilités, au même titre que la validité de l'indice premier de Rawls a pu dépendre de la validité de l'argumentation de la position originelle.

La définition du contenu de la liste spécifie la nature des capabilités et les fonctionnements devant être recherchés. Ce contenu, selon Nussbaum, ne devrait pas être trop éloigné des intuitions et des jugements du sens commun.

Ces deux versants du problème de l'indice des capabilités mettent en évidence deux modes d'appréhension différents de ce problème, qui reposent chacun sur des présuppositions philosophiques divergentes.

[432] NUSSBAUM, *op. cit.*.

7.1.2. Le silence de Sen

Si l'on refuse de traiter le problème de la définition du contenu de la liste des capabilités, on doit se contenter de fournir quelques indications sur la procédure. C'est la posture adoptée par Sen, qui a, à de multiples reprises, refusé de créer ou de valider une liste de capabilités. Dans sa première monographie sur les capabilités, Sen se contente de présenter la liste comme un problème de choix social[433] ; le classement des fonctionnements pertinents devrait rester incomplet et éviter la surdétermination, et dépend de la nature et de la culture de la société concernée[434] ; il rappelle que la métrique de la capabilité possède déjà un pouvoir discriminant, d'autant plus si on lui ajoute le principe selon lequel plus de capabilité est préférable à moins[435] ; sa position générale se résume ainsi :

> *« Le fait que, en dépit de ma défense de l'utilisation d'une « perspective des capabilités » pour comparer les avantages individuels et de ma tentative de mettre en lumière la pertinence de certaines capabilités de base en particulier (par exemple, dans « Bien-être et liberté d'agent »[436]), je n'aie pas spécifié une liste définie de « capabilités » distinctes (avec des pondérations spécifiques ou d'autres manières d'établir des priorités) a été à l'origine de reproches que j'ai reçus. Cependant, si une telle liste définie avec des priorités et des pondérations fixées était de fait l'aboutissement d'un raisonnement éthique général, pour moi, il n'est pas clair que ce puisse être compatible avec le processus démocratique d'établissement des priorités et des préséances. »*[437]

Le refus d'assigner un contenu est dû à la volonté de respecter un choix démocratique des capabilités. Autrement dit, pour Sen, le choix collectif est la source de la normativité en ce qui concerne les capabilités, qui ne fournissent que la manière de décrire le monde, ou encore le paradigme d'évaluation. D'où l'importance de penser ce processus de choix collectif de la manière la plus ouverte et la plus précise possible, et de ne pas se limiter à une interprétation réductrice du cadre conceptuel de la théorie du choix social. Dans l'*Idée de Justice*, Sen souligne la « connexion entre le raisonnement public et le choix et la pondération des capabilités dans l'évaluation sociale »[438], marquant une fois de plus sa volonté d'articuler

---

[433]SEN, Amartya ,*Commodities and Capabilities*, Amsterdam: North-Holland, 1985, p 35.

[434] SEN, Amartya, « Equality of What ? », *Choice, Welfare and Measurements*, Cambridge MA, The MIT Press, 1982; traduction française par Sophie Marnat, « Quelle égalité ? », *Ethique et Economie*, Paris, Presses Universitaires de France, Quadrige, 1993, 2008, p 212.

[435]SEN, Amartya, *Inequality Reexamined*, Oxford University Press, 1992 ; traduction française par Paul Chemla : *Repenser l'inégalité*, Paris, Editions du seuil, 2000, p 70-76.

[436]SEN, Amartya, *On Ethics and Economics*, Blackwell Publishers, Oxford, 1991; traduction française par Sophie Marnat : *Ethique et Economie*, Paris, Presses Universitaires de France, Quadrige, 1993, 2008.

[437]SEN, Amartya K., "Incompleteness and Reasoned Choice", *Synthese*, 140, 2004, n29.Traduction libre..

[438]SEN, Amartya K., *The Idea of Justice*, London: Allen Lane, 2009, p 242.

l'agrégation des préférences et des jugements de valeur à une réflexion sur les processus de délibération collective caractérisant la raison publique.

7.1.3. Une liste de capabilités nécessaires à un fonctionnement digne de l'être humain : la solution néo-aristotélicienne de Martha Nussbaum

Si l'on décide de déterminer une liste de capabilités intuitivement satisfaisante, avant de définir une procédure rationnelle susceptible de reconstituer une telle contenu, on fait le choix de se concentrer sur une théorie éthique substantielle à l'appui des capabilités avant de réfléchir sur les processus d'évaluation et de décision. C'est la voie empruntée par Martha Nussbaum. La philosophe américaine prend pour point de départ une « idée intuitive », celle d'une vie digne d'un être humain, qu'elle emprunte à la fois à Aristote et au Marx des *Manuscrits de 1844 :*

> *« L'idée essentielle est celle de l'être humain en tant qu'être libre et plein de dignité qui, homme ou femme, forge sa propre vie dans la coopération et l'échange avec les autres, plutôt que d'être passivement façonné ou malmené à la façon d'un animal vivant en « troupeau » ou en « bande »* [439]

Nussbaum s'appuie sur le fait que cette norme est présente dans la tradition philosophique, notamment chez Aristote et Marx, d'une part, et qu'elle apparaît également dans les œuvres littéraires issues de cultures diverses, auxquelles chacun est susceptible de réagir. Confiante en l'hypothèse de l'universalité de cette norme de l' « humain », Nussbaum propose de reprendre une procédure de raisonnement public conceptualisée par Rawls : le consensus par recoupement[440]. Ce consensus stipule que des personnes ayant des conceptions différentes peuvent se mettre d'accord sur ce qui est juste ; de même, Nussbaum estime que des personnes ayant des conceptions religieuses ou éthiques différentes peuvent adhérer à cette norme intuitive de l'humain, norme appelée à régir l'espace public et les décisions institutionnelles. Elle a donc mis en œuvre des « discussions transculturelles » chargées de vérifier les intuitions de ce que sont les fonctionnements et les capabilités dignes de l'être

[439]NUSSBAUM, Martha, *Women and human Development: the capabilities Approach*, Cambridge: Cambridge University Press, 1999. *Femmes et développement humain*, traduit de l'anglais (USA) par Camille Chaplain, Paris, Edition Des Femmes-Antoinette Fouque, 2008, p 111.

[440]RAWLS, John, *Political Liberalisme*, New York, Columbia University Press, 1993; traduction française par Catherine Audard: *Libéralisme politique*, Paris, Presses Universitaires de France, « Quadrige », 1995, 2007.

humain et donc de valider le résultat hypothétique et intuitif du consensus par recoupement. La liste évolue donc constamment. La version proposée en 2000 est la suivante[441] :

1. Vie : être capable d'avoir une durée de vie « normale ».
2. Santé physique : être capable de jouir d'une bonne santé et des conditions de cette santé, à savoir l'alimentation et le logement.
3. Intégrité physique : être capable de disposer de son corps afin de le protéger, de le déplacer et de jouir d'une sexualité satisfaisante et du choix de se reproduire ou non.
4. Sens, imagination, pensée : être capable d'utiliser ses sens, son imagination, sa pensée et sa raison afin d'éprouver ses expériences agréables, de produire et de profiter des œuvres, de s'exprimer, de chercher le sens de son existence.
5. Emotions : être capable d'éprouver des émotions, de jouir de relations humaines affectivement satisfaisantes.
6. Raison pratique : être capable de se forger une conception du bien et de l'examiner de façon critique.
7. Affiliation : être capable de s'engager dans des interactions sociales, d'éprouver de l'empathie, et jouir des bases sociales du respect de soi-même et de la reconnaissance mutuelle.
8. Autres espèces : être capable d'interagir avec les animaux, les plantes et la nature.
9. Jeu : être capable de rire et de jouer.
10. Contrôle de son environnement : être capable de participer à la vie politique et aux transferts économiques[442].

La liste proposée par Nussbaum a fait l'objet de vives critiques, car le raisonnement sur laquelle elle s'appuie fait figure d'argument d'autorité. Allison Jaggar offre ainsi trois axes critique de ce raisonnement[443] :

(1) Nussbaum sélectionne les intuitions morales sur lesquelles sa liste repose en raison même du fait qu'elles coïncident avec sa liste. La procédure de justification n'est donc

[441]NUSSBAUM, Martha, *Women and human Development: the capabilities Approach*, Cambridge: Cambridge University Press, 1999. *Femmes et développement humain*, traduit de l'anglais (USA) par Camille Chaplain, Paris, Edition Des Femmes-Antoinette Fouque, 2008.

[442] Cette liste, ainsi que les suivantes, fait apparaître un glissement de sens du terme « capabilité », qui décrit simultanément des ensembles d'opportunités et des fonctionnements en puissance.

[443] JAGGAR, Alison, « Reasoning about Well-Being : Martha Nussbaum's methods of Justifying the Capabilities", *The Journal of Political Philosophy*, volume 14, Issue 3, 2006, pp 301-322.

pas définie indépendamment du contenu que la philosophe souhaite justifier, et de ce fait ne justifie rien ;

(2) En recueillant et en interprétant les jugements et les intuitions des personnes des pays pauvres qui confirment sa liste, Nussbaum ne prend pas du tout en compte les biais qu'implique nécessairement un entretien entre une femme américaine éduquée et aisée avec des personnes pauvres citoyennes d'anciennes colonies. Les chances que ces femmes manifestent un désaccord sont minces. Ces problèmes sont bien connus des chercheurs en sciences sociales, et il est surprenant que Nussbaum ne se soit jamais interrogée à ce sujet ;

(3) L'attitude de Nussbaum est comprise par Jaggar comme irrespectueuse à l'égard des jugements et des voix qui pourraient exprimer un désaccord avec le contenu de la liste des capabilités fondamentales, dans la mesure où aucune procédure de discussion ou de contestation n'est prévue.

7.1.4. L'évaluation démocratique des capabilités aujourd'hui : les travaux de d'Ingrid Robeyns, de Sabina Alkire et d'Elizabeth Anderson

D'autres propositions de solutions au problème de la sélection des capabilités ont émergé récemment. Les caractéristiques communes de ces solutions sont les suivantes : elles prennent en compte l'exigence démocratique pour définir la procédure d'évaluation – sans cependant faire usage des travaux de Sen en théorie du choix social - ; elles se concentrent sur le contenu de la liste plus que sur les problèmes d'arbitrage et de pondération et ne sont donc pas des solutions complètes – elles ne prétendent pas l'être par ailleurs - .

a) Ingrid Robeyns : expertise, brainstorming et contextualisation

Ingrid Robeyns a élaboré une solution intermédiaire entre le refus de la liste et l'imposition d'une définition des capabilités : elle suggère de définir une approche procédurale adaptée à chaque contexte d'évaluation des capabilités : par exemple, l'amélioration des capabilités liées aux différences de genre requiert une liste donnée ; des inégalités d'autre ordre requerraient d'autres listes[444]…La procédure de Robeyns devrait satisfaire les critères suivants : une formulation explicite, une justification méthodologique, la sensibilité au

[444]ROBEYNS, I., "Sen's Capability Approach and Gender Inequality: Selecting Relevant Capabilities," *Feminist Economics*, 9(2/3), 2003: 61–92.

contexte, les différents niveaux de généralité (Robeyns affirme qu'il faut distinguer un contexte idéal et un contexte de second rang), l'exhaustivité plutôt que la réduction. La procédure de Robeyns comprend les étapes suivantes :

(1) Une étape de remue-méninges sans contrainte ;

(2) Une mise à l'épreuve de différents brouillons de liste avec les travaux académiques sur la question traitée (par exemple, les inégalités de genre) ;

(3) La confrontation de la liste avec d'autres listes existantes ;

(4) Une discussion de la liste avec d'autres personnes.

La méthode de Robeyns est inductive plutôt que déductive, et s'inscrit dans un contexte d'expertise plutôt que de réflexion méta-éthique ou de philosophie politique[445].

[445] Afin de se donner une idée, la liste des capabilités liées aux inégalités de genre élaborée par Robeyns comprend les éléments suivants : la vie et la santé physique ; le bien-être mental ; l'intégrité du corps et la sécurité ; les relations sociales ; l'*empowerment* politique ; l'éducation et la connaissance ; le travail domestique et la prise en charge des proches ; le travail salarié et les autres projets ; le logement et l'environnement ; la mobilité ; les activités de loisirs ; l'autonomie dans l'allocation de son temps ; le respect ; la religion.

b) Sabina Alkire : le raisonnement pratique de John Finnis et les dimensions du développement humain

Sabina Alkire a complété l'approche par les capabilités avec les travaux de John Finnis, un professeur de jurisprudence[446]. Finnis défend une théorie du raisonnement pratique qui comprend une conception de la valeur et des fins humaines, un principe d'action permettant de poursuivre ces fins et affirme la nécessité du choix libre. Il établit également des catégories centrales pour l'épanouissement humain : la vie, le savoir, le travail et le jeu, l'amitié et les relations humaines, l'auto-intégration, l'auto-direction et la transcendance. Cette liste n'est pas une liste définitive de fonctionnements et de capabilités ; sa fonction est de nourrir un débat participatif. Alkire a confronté ce schème normatif à des travaux de terrain dans des villages pakistanais[447]. La proposition d'Alkire est, dans son esprit, plus proche d'une approche démocratique plutôt que substantielle du problème de la liste. Reste à déterminer si une théorie telle que celle de Finnis est nécessaire, ou si l'analyse que nous avons proposée du cadre conceptuel de la théorie du choix social ne suffirait pas pour construire une conception du raisonnement collectif conduisant à l'évaluation et à la construction d'une liste de capabilités.

[446] Cf. FINNIS, John, *Fundamental of Ethics*, Georgetown University Press, 1983.

[447] ALKIRE, S., *Valuing Freedoms: Sen's Capability Approach and Poverty Reduction*, New York: Oxford University Press, 2002

c) *Elizabeth Anderson : les capabilités suffisantes pour l'égalité démocratique*

Elizabeth Anderson a élaboré une proposition de sélection des capabilités selon le critère d'égalité démocratique, qui implique deux principes :

(1) Les personnes ont droit à toute capabilité leur permettant d'échapper à des relations sociales d'oppression ;

(2) Les personnes ont droit à toutes les capabilités nécessaires pour fonctionner comme des citoyens égaux dans un Etat démocratique[448].

Ce critère peut sembler limité dans un premier temps. Mais Anderson précise qu' »être capable de fonctionner comme un citoyen égal » présuppose le fait d' »être capable de fonctionner comme un être humain », et l'on retrouve ici la possibilité d'inclure des capabilités relative à la santé, à l'alimentation, au logement.

L'approche d'Anderson est prometteuse, mais requiert d'être complétée en amont par une réflexion approfondie sur les fondements de ce principe d'égalité démocratique, et en aval par une spécification plus précise des critères de décision d'arbitrage entre les capabilités, une fois le niveau suffisant pour garantir l'égalité démocratique atteint.

D'autre part, la proposition d'Anderson s'inscrit dans l'opposition entre égalitarisme démocratique et égalitarisme des chances (*luck egalitarianism*). Anderson a critiqué ce dernier de manière virulente. L'égalitarisme des chances, courant influent dans les débats contemporains sur la justice distributive[449], stipule que, parmi les inégalités (que celles-ci soient mesurées en termes de ressources ou d'opportunités), seules les inégalités résultants de la « chance » (*luck*) sont injustes et doivent être annulées. Les inégalités résultant du choix de l'individu qui s'en trouve plus mal loti ne sont pas considérées comme injustes.

Selon Anderson, l'égalitarisme des chances manque ce qu'il y a de plus important pour l'égalitarisme historique: l'égal respect des personnes. Cette exigence d'égal respect interdit la formulation de jugements intrusifs sur les choix et la responsabilité des personnes. Là où l'égalitarisme des chances recherche l'atténuation d'injustices naturelles, l'égalitarisme des chances vise la disparition de l'oppression socialement créée. Là où l'égalitarisme des

[448] ANDERSON, E., "What is the Point of Equality?" *Ethics*, 109(2): 287–337, 1999.

[449] A l'origine de celui-ci, figurent notamment les travaux de Ronald Dworkin, de G. A. Cohen et de Richard Arneson. Cf. ARNESON, Richard, « Equality and equal opportunity for welfare », *Philosophical Studies*, May 1989, Volume 56, Issue 1, pp 77-93, COHEN, G.A., "On the currency of egalitarian justice", *Ethics*, Vol. 99, No. 4, 1989, pp 906-944.

chances se limite à penser l'égalité distributive, l'égalitarisme des chances recherche l'égalisation des relations.

La difficulté est que cette critique contredit d'une certaine manière l'adoption de la perspective des capabilités dans la seconde partie de l'article. En effet, en refusant de caractériser les désavantages individuels au nom de l'égal respect des citoyens, Anderson se prive en fait de la possibilité d'utiliser la métrique des capabilités, et devrait logiquement se focaliser davantage vers une approche ressourciste de la distribution[450].

D'autres propositions émergent. Certaines relèvent davantage d'un travail sur les fondements d'une éthique économique et sociale, et visent un critère du bien-être ou du Bien permettant d'établir les principales composantes d'une liste de capabilités. D'autres visent davantage la définition d'une procédure politique, idéalement démocratique, permettant de choisir les capabilités qui importent. Certaines s'inscrivent dans la démarche de construction d'une théorie de la justice sociale normative, destinée à structurer un schème institutionnel donné. D'autres visent davantage à offrir des directives pour des actions de développement et de lutte contre la pauvreté très localisées, où les acteurs individuels sont identifiés comme tels. Les contextes d'application varient également, allant d'une théorie générale du bien-être à des approches plus localisées, se focalisant exclusivement sur la santé, les inégalités de genre, l'éducation, l'environnement.

En dépit de la diversité des propositions de listes, deux leçons peuvent être tirées de nos analyses. La première est celle de la tension entre l'éthique libérale de l'approche par les capabilités et sa volonté de faire une proposition substantielle pour évaluer le bien-être et les inégalités. Autrement dit, l'approche par les capabilités semble libérale parce que centrée sur les libertés et perfectionniste parce que refusant de s'en tenir à une conception subjective du bien-être. La seconde est que la caractérisation des capabilités se fait mieux à petite échelle. Si, dans leurs esprits, les institutions peuvent fort bien favoriser le développement des capabilités, dans la pratique, l'évaluation des capabilités individuelles ne peut s'effectuer qu'à petite échelle et doit s'intéresser aux individus et aux relations sociales particulières, y compris dans des sphères telles que celle de la famille. A l'échelle de l'Etat-Nation, ou à

[450] Cette contradiction entre le critère d'égalité démocratique et le choix de la métrique des capabilités est mise en évidence par Thomas Pogge: cf. POGGE, Thomas, "Can the Capability Approach be Justified?", *Philosophical Topics,* 30 (2), 2002, pp 167-228.

l'échelle internationale, les capabilités seront définies de manière très générale et ne diffèreront pas tant que cela de ressources décrites de manière appropriée[451].

[451] POGGE, Thomas, "Can the Capability Approach be Justified?", *Philosophical Topics,* 30 (2), 2002, pp 167-228

7.2. Une approche préférentialiste de l'évaluation des capabilités

A présent, il est temps de proposer une solution au problème de l'évaluation et de la liste des capabilités qui s'appuie sur notre analyse fouillée du cadre conceptuel de la théorie du choix social. Sen affirme ce problème doit recevoir un traitement démocratique. Est « démocratique » ce qui relève d'un « choix collectif ». L'argument principal en faveur d'un traitement démocratique de ce problème est épistémique : la recension des préférences ou des votes des membres de la société constitue une source d'information sur les situations, les fins et les valeurs des membres à laquelle peu d'expertises peuvent prétendre accéder. D'autre part, la mise en commun de ces *data*sur les valeurs des membres de la société motive une réflexion collective afin de les rendre commensurables. Elizabeth Anderson rappelle que cette conception épistémique de la démocratie est au cœur des travaux de Sen :

> *« L'argument épistémique de Sen en faveur du rôle crucial des Etats démocratiques dans la prévention des famines reflète une compréhension de la démocratie plus large, comme l'incarnation institutionnelle du raisonnement collectif et de l'expérimentation de la manière dont nous devrions vivre ensemble. »*[452]

La démocratie selon Sen n'est pas qu'une forme institutionnelle vide, caractérisée exclusivement par sa conformité aux principes. En fait, la question de la nature des institutions est seconde : ce qui importe, c'est la réalité des processus de raisonnement, d'agrégation et d'expérimentation qui ont lieu dans le cadre institutionnel, quel qu'il soit. Cette primauté de l'exercice du choix collectif sur la forme institutionnelle reflète de fait la perspective de la théorie du choix social sur la démocratie. La démocratie est d'abord, avant tout, le lieu du choix collectif, et l'évaluation comme la sélection des capabilités en sont un exercice exemplaire.

Pourquoi choisir un processus démocratique dans des conditions non-idéales, plutôt qu'une argumentation dans des conditions idéales telle que celles préconisées par les théories de la justice hypothético-déductives (ou aspirant à l'être) telles que celles de Rawls, de Dworkin ou encore de David Gauthier ? Pourquoi privilégier l'agrégation et la comparaison des états sociaux plutôt qu'une approche contractualiste de la justice sociale (approche qui pourrait d'ailleurs être complétée par des expériences démocratiques concrètes) ? Il s'avère que le paradigme des capabilités requiert une description précise, incorporant des éléments

[452]ANDERSON, Elizabeth, "Sen, ethics and democracy", *Feminist Economics*, 0(0), 2003.

particuliers qu'un voile d'ignorance tendrait à dissimuler. C'est ici que se situe l'enjeu de la distinction entre approche transcendantale, ou idéale, de la justice sociale, et approche comparative non-idéale[453].

Dans cette dernière section, nous élaborerons une proposition préférentialiste inspirée de la conception des préférences propre à la théorie du choix social, afin d'évaluer les capabilités et de déterminer si elles peuvent intégrer une liste de capabilités nécessaires au bien-être des personnes, ou non. La comparaison d'états sociaux décrits en termes de fonctionnements et le processus démocratique au sein duquel chacun peut exprimer sa voix sont les ingrédients majeurs d'une proposition qui vise à clarifier et à faire aboutir certaines des suggestions de Sen.

Nous présenterons tout d'abord la question de l'agrégation des capabilités dans la théorie du choix social en nous appuyant sur les deux pistes de résolution du théorème d'impossibilité d'Arrow que propose Fleurbaey (7.2.1.). Ensuite, nous rappellerons la conception comparative de la préférence que nous utilisons ici (7.2.2.). Nous analyserons ensuite le concept de « préférences pour des fonctionnements », et la manière dont les fonctionnements comme objets structurent la préférence elle-même (7.2.3.). Enfin, nous examinerons le difficile problème de la relation entre préférences pour les fonctionnements et évaluations des capabilités, en soulignant l'importance de la liberté de choix pour la délibération (7.2.4.).

[453]SEN, Amartya K., *The Idea of Justice*, London: Allen Lane, 2009.

7.2.1. Capabilités et théorie du choix social : deux possibilités

L'approche par les capabilités s'est construite contre l'utilitarisme strict et exclut les évaluations seulement basées sur la conception hédoniste des préférences. Deux voies possibles subsistent alors : une première établit un indice objectif de capabilités et de fonctionnements s'appliquant à tous les êtres humains, à partir duquel on effectue des comparaisons interpersonnelles de bien-être ; une seconde est suggérée par une note de bas de page de Marc Fleurbaey dans son article consacré au choix social et aux institutions justes :

> *« L'approche selon l'équité suggère de considérer la possibilité d'avoir des indices personnalisés reflétant les préférences de chaque individu sur ses propres fonctionnements et capabilités. »*[454]

La distinction entre approche comparative et approche selon l'équité de l'évaluation sociale est basée sur une reformulation de l'axiome arrovien d'indépendance des alternatives non pertinentes en théorie du choix social (cf. section 2.2.2.). Sen a proposé d'échapper au résultat d'impossibilité de Kenneth Arrow en réintroduisant les comparaisons interpersonnelles de bien-être plutôt que de s'appuyer sur une base informationnelle de choix social constituée seulement de préférences ordinales[455]. Cette solution constitue le point de départ de ce que Fleurbaey qualifie d' »approche comparative ». Mais Fleurbaey démontre que nous pouvons échapper au résultat d'Arrow différemment. En effet, il remarque que l'axiome d'indépendance des alternatives non pertinentes comprend en fait deux propositions distinctes, qui sont deux conditions différentes de choix social (cf. aussi la section 5.2.2., § c) de notre travail) :

1) Une condition de non-comparabilité ordinale, selon laquelle, si les fonctions d'utilité individuelles changent sans altérer l'ordre de préférences, la préférence sociale devrait rester la même. La préférence sociale ne doit pas dépendre de la mesure des utilités individuelles, mais seulement de leur classement ordonné.
2) Une condition d'indépendance des utilités non pertinentes, selon laquelle une préférence sociale sur une alternative doit seulement dépendre des préférences des membres de la société sur cette même paire.

L'approche comparative stipule que nous devons abandonner la « non-comparabilité ordinale » et réintroduire la mesure cardinale des utilités ou des capabilités et la comparabilité

[454]FLEURBAEY, Marc, "Social Choice and Just Institutions", *Economics and Philosophy*, 23, 2007, p. 15-43.
[455]SEN, Amartya, *Collective Choice and Social Welfare*, San Francisco, Holden day, 1970, chapter 11.

interpersonnelle du bien-être afin de procéder à l'évaluation sociale. A l'inverse, l'approche selon l'équité défend l'abandon de l' « indépendance des utilités non pertinentes ». Aucune comparaison interpersonnelle de bien-être n'est autorisée, mais Fleurbaey propose que nous prenions en compte les préférences pour des options qui ne sont pas disponibles lorsque nous procédons à l'évaluation sociale[456].

L'indépendance des utilités non pertinentes est appropriée dans le cas d'élections de représentants : mieux vaut éviter que les votants ne choisissent un candidat décédé ou inexistant. Mais elle ne l'est pas nécessairement dans un contexte d'évaluation sociale, même démocratique : par exemple, une personne peut envier le lot de biens d'une autre personne, même si ce lot ne lui est accessible dans aucun des états sociaux possibles ; or, ce type d'information pourrait être pris en compte dans l'évaluation sociale[457].Fleurbaey souligne que son « approche selon l'équité » fait dépendre l'évaluation de la nature de l'objet comparé, c'est-à-dire de l'objet même des préférences. Sa contribution s'intéresse principalement aux paniers de biens, se situant dans une perspective rawlsienne. Mais Fleurbaey mentionne également une variante possible, les préférences pour les capabilités. La solution de Fleurbaey met en évidence un oubli important dans les efforts d'interprétation du cadre conceptuel de la théorie du choix social et de résolution du résultat d'impossibilité: la nature des objets et le paradigme utilisé pour les décrire constituent une voie à explorer. En fait, dans la caractérisation même de la rationalité des préférences et des choix, nous tendons à porter une attention excessive aux sujets qui évaluent – les agents dans la langue des théories de la décision –, à leurs motifs comme à leurs mobiles, au détriment du rôle joué par les objets et leur description dans la formation et l'énonciation des préférences.

[456] Cf. section 5.2.2.c. de ce travail.

[457] L'argument est développé à la page 20 de l'article "Social Choice and Just Institutions". Le terme « envie » fait ici référence au critère d'"absence d'envie" pour l'allocation des ressources, et non pas à un désir déraisonnable de posséder plus qu'une autre personne. Sur la signification particulière du terme "envie" en économie normative et la pertinence d'un tel critère pour la justice, cf. FLEURBAEY, Marc, *Théories économiques de la justice*, Paris, Economica, 1999, 10.6., p 222-229.

7.2.2. Les évaluations comparatives de capabilités

La tâche du théoricien du choix social sera alors d'axiomatiser l'agrégation des préférences pour les capabilités afin d'en dériver un classement collectif desdites capabilités susceptible de déterminer le contenu d'une liste et les éventuelles décisions d'arbitrage. Mais, en amont, le philosophe est amené à s'interroger sur le sens de ce qu'est une « préférence pour une capabilité », d'autant que ces deux notions, la préférence est la capabilité, sont généralement présentées comme antithétiques.

Evaluer des libertés en fonction des préférences requiert un travail de réflexion sur ces dernières. L'enjeu ici est la nature de la préférence. Nous avons vu que les évaluations comparatives représentées par des relations d'ordre en théorie du choix social peuvent recevoir différentes interprétations. Reste à déterminer l'interprétation compatible avec les capabilités.

Les analyses que nous avons menées dans les chapitres 4 et 5 de notre travail ont mis en évidence deux conceptions possibles de la préférence. L'une assimile la préférence à un désir et mesure le bien-être obtenu en fonction du degré de frustration ou de satisfaction subjective éprouvée par l'agent. L'autre identifie la préférence à une évaluation comparative fondée sur l'appréciation des propriétés respectives des états sociaux évalués. En bref, l'une se focalise sur les états d'esprits, l'autre sur les états de choses[458]. L'approche par les capabilités est caractérisée par le refus d'endosser une conception hédoniste du bien-être et ne peut donc souscrire à la première conception. L'évaluation préférentialiste des capabilités et des fonctionnements s'appuiera donc sur la conception comparative de la préférence.

Le choix d'une approche préférentialiste de l'évaluation des capabilités peut sembler éloigné de la perspective de Sen, qui n'a pas manqué de se montrer assez critique à l'égard de l'utilitarisme en général. Mais Sen lui-même rappelle que les préférences dans la théorie du choix social peuvent être comprises de manière plus large :

> *« Je dois aussi mentionner que le terme de « préférence » est utilisé [dans l'évaluation de la liberté en tant qu'opportunité] dans un sens plus général que celui qu'on lui donne parfois – par exemple quand la préférence est simplement identifiée avec le sens psychologique du verbe « préférer ». J'utilise plutôt le terme de « préférence » au sens large que lui donne la théorie du choix social, qui permet aussi quelques variations d'interprétation. […] La définition de base de la préférence donnée par Kenneth Arrow dans son livre pionnier (Arrow, 1951) comme incluant « le système complet des*

[458]GRIFFIN, J.P., *Well-Being*, Oxford: Clarendon Press, 1986.

valeurs d'une personne, y compris les valeurs au sujet des valeurs » (p. 18), est particulièrement pertinente dans ce cadre. »[459]

Nous avons vu que la théorie du choix social considère d'ordinaire que les objets de la préférence sont des états sociaux. Les objets des préférences pour les capabilités seraient donc des états sociaux décrits au travers du prisme de l'approche par les capabilités. Ce « prisme » correspond à un paradigme descriptif bien spécifique, impliquant la prise en compte non seulement des faits donnés, mais également des processus, par le biais des fonctionnements, des possibilités contrefactuelles, par le biais des ensembles capabilités, et de l'interdépendance entre le choix et l'environnement, par le biais de l'analyse des facteurs de conversion des ressources en fonctionnements. Cette perspective s'articule parfaitement avec le conséquentialisme élargi qu'a défendu Sen en philosophie morale[460] : le processus ou l'acte qui permet de produire le résultat peut être inclus dans la description de ce résultat. Par exemple, l'état de choses dans lequel je mange un poisson que j'ai pêché n'est pas le même que celui où je mange un poisson que l'on m'a donné. Le résultat est le même : je mange un poisson ; mais les processus de production de ce résultat nous apprennent beaucoup plus sur l'effort fourni, les savoirs et savoir-faire, la nature des relations sociales impliquées...Lorsque les membres de la société sont amenés à comparer et à évaluer les capabilités, ils prendront en compte ces informations grâce à un paradigme descriptif plus riche.

[459] SEN, Amartya, *Rationality and Freedom*, The Belknap Press of Harvard University Press, 2002. Trad. Fr. M.-P. d'Iribane-Jaawane, *Rationalité et liberté en économie*, Paris : Odile Jacob, 2005.

[460] SEN, Amartya, "Consequential evaluation and practical reason", *The Journal of Philosophy*, vol. 97, no 9, sept. 2000, pp 477-502.

7.2.3. Les préférences pour les fonctionnements

En quel sens une description plus riche des états sociaux permet-elle aux préférences pour les capabilités d'échapper aux difficultés inhérentes au préférentialisme ? Pour répondre, il convient de distinguer les « préférences pour les capabilités » des « préférences pour les fonctionnements ». Nous commencerons par déterminer ce que sont les préférences pour les fonctionnements avant d'examiner l'apport d'une perspective en termes de capabilités ou de libertés réelles.

a) Les préférences pour les fonctionnements : problèmes et définitions

Si l'évaluation des capabilités est basée sur les préférences, alors, en termes de choix social, cela signifie que les états sociaux évalués seront décrits en termes de fonctionnements et de capabilités plutôt que de ressources et d'expériences de satisfaction. Comment la sélection des fonctionnements peut-elle être basée sur les préférences, comprises comme des évaluations comparatives d'états sociaux appréciés en fonction de leurs propriétés ?

Parmi les risques encourus par le préférentialisme, figure celui d'une dépendance excessive des jugements de préférence par rapport au contexte. Deux types de solution s'offrent pour éviter dans une certaine mesure le classique problème des préférences adaptatives ; en substance, la première consiste à traiter le « sujet », et la seconde l' « objet ».

La première solution a été longuement développée dans la section 5.2.3.c. de notre travail. L'idée générale consiste à déterminer des conditions idéales de délibération afin de déterminer des préférences susceptibles de décrire de manière appropriée le bien-être de l'agent. La seconde solution est celle que nous proposons ici. L'idée générale est que la manière dont les états sociaux évalués sont décrits peut modifier la structure de la préférence pour ces derniers.

Dans le cas des préférences pour les fonctionnements, deux options s'offrent à nous, selon que nous adoptons une conception aristotélicienne ou neutre des fonctionnements[461].

Si les fonctionnements sont compris au sens aristotélicien de « fonctionnement propre » à l'être, cela signifie que le fonctionnement ne peut qu'être « bon » pour cet être, dans la mesure où il est adapté à sa nature. Des activités et des états préjudiciables pour l'agent ne peuvent être donc qualifiés de fonctionnements. Cette conception du fonctionnement

[461] Cette distinction est de CARTER, Ian, "The concept of freedom in the work of Amartya Sen: An alternative analysis consistent with freedom's independent value", *Notizie di Politeia* 43(4), 1996.

présuppose l'existence d'une norme objective du bien-être, norme qui serait établie en fonction de la nature même de cet être. Cette position est incompatible avec le préférentialisme, qui affirme que le bien-être dépend des préférences de l'agent. A la rigueur, cette position ne pourrait assigner aux préférences que le rôle de révélateur d'un bien-être qui existerait indépendamment de celles-ci.

Le préférentialisme doit donc s'en tenir à une conception neutre du fonctionnement. Les fonctionnements sont ainsi des activités ou des états dont la valeur dépend du jugement de l'agent. Ainsi, « jeûner » n'est pas un fonctionnement au sens aristotélicien doublé d'une conception du bien-être fondée sur la satisfaction des besoins de l'être humain compris comme être vivant (à moins que le jeûne ne joue un rôle dans la thérapie d'une pathologie). En revanche, si « jeûner » est un fonctionnement au sens neutre du terme, cela laisse la possibilité à l'agent de lui assigner une valeur positive (pour des raisons politiques ou religieuses, par exemple) aussi bien qu'une valeur négative.

b) *La structure holiste des préférences pour les fonctionnements*

Bien que les fonctionnements au sens neutre du terme n'affectent pas le contenu des préférences, ils en modifient la structure. Une distinction que Derek Parfit a élaborée afin de distinguer deux versions du préférentialisme s'avère ici utile. Il s'agit de la distinction entre « préférence globale » et « préférence sommative »[462]. Cette distinction s'applique aux préférences qui concernent l'existence d'une personne dans son ensemble, et non pas les préférences pour des événements spécifiques, tels que la dégustation d'un gâteau ou le choix de la couleur d'une chemise. Les préférences qui nous intéressent ici sont celles portent sur les choix professionnels, la santé, l'éducation, les engagements politiques ou sociaux…

Les préférences sommatives sont établies de la manière suivante : les préférences qui adviennent à des moments déterminés de l'existence de l'agent sont collectées et agrégées. L'existence que l'agent devrait préférer est l'existence dans laquelle le plus grand nombre de préférences est satisfait (ou, si l'on adopte une interprétation de la préférence en termes de désirs dont on peut mesurer l'intensité, ce sera l'existence dans laquelle la plus grande magnitude de satisfaction est obtenue). Leur structure est atomiste.

Les préférences globales n'ont pas pour objet des événements ponctuels (ou « locaux », selon la métaphore spatiale de Parfit). Elles portent sur la vie entière de l'agent considérée comme

[462]Parfit, *Reasons and persons*, Appendix I.

telle, ou encore sur des portions de son existence que l'on peut considérer comme un tout. Leur structure est holiste.

Les préférences pour les fonctionnements ne peuvent être sommatives car les fonctionnements ne sont justement pas des événements ponctuels ou des occurrences limitées. Ce sont des processus. D'autre part, il est difficile d'envisager l'agrégation de fonctionnements. On peut les combiner pour constituer une existence, mais on ne peut les rendre commensurables, c'est-à-dire les mesurer sur une échelle commune. Un fonctionnement doit être considéré comme un tout.

Une préférence pour une combinaison de fonctionnements est une préférence pour une existence décrite sous la forme d'un vecteur[463] de fonctionnements. La valeur d'une existence ainsi décrite ne peut être réduite à la valeur agrégée de ses composantes. Par exemple, si je combine « être adéquatement nourri », « avoir un abri », « être en bonne santé », et « prendre part à la vie de la communauté », j'obtiens une existence que je peux valoriser. Mais si je soustrais « être adéquatement nourri » de cette combinaison, cela n'implique pas que cette existence devienne juste « moins valable ». En fait, elle cesse d'avoir quelque valeur que ce soit, puisque je constate que je ne peux réaliser les autres fonctionnements sans celui-ci. Les fonctionnements ne sont donc pas simplement superposés, ils s'enchevêtrent. La valeur d'une combinaison de fonctionnements ne se réduit pas à la valeur agrégée de ses composantes, et l'ajout d'un fonctionnement à une combinaison existante peut créer une valeur bien supérieure à celle de la somme agrégée de l'ancienne combinaison et du nouveau fonctionnement.

L'exemple proposé ci-dessus nous permet également de comprendre en quel sens les fonctionnements ne peuvent pas être commensurables. Comment aligner « être adéquatement nourri » et « avoir un abri » sur une échelle commune ? Une approche basée sur la valeur marchande des ressources peut offrir une solution : comparer le prix de chacun de ces fonctionnements sur le marché. Une approche utilitariste comparera le degré de satisfaction procuré par chacun de ces fonctionnements[464]. Mais l'approche par les capabilités se veut pluraliste et ne promeut pas une telle réduction. Comme le soutient Sen, « aucune

[463]Un vecteur est un ensemble dont les éléments sont ordonnés.

[464] L'exemple que nous avons donné peut laisser penser que ces fonctionnements sont incommensurables parce que nécessaires à la survie : ils satisfont des besoins de base. Mais ce n'est pas le nœud du problème. Outre le fait qu'une approche basée sur les besoins ne permet que d'évaluer des situations d'extrême pauvreté et peut s'avérer insuffisamment exigeante dans les contextes où lesdits besoins sont satisfaits mais des droits fondamentaux sont violés, par exemple, une telle approche ne laisse pas de place à l'exercice de la qualité d'agent et à l'évaluation des fonctionnements par l'agent elle-même. C'est en ce sens que l'approche par les capabilités ambitionne de se détacher de l'approche en termes de besoins.

comparaison fonctionnelle-intervaluationnelle »[465] n'est possible. Cela ne fait pas sens de comparer le degré d'accomplissement d'un fonctionnement donné ou d'une capabilité. Soit une personne accomplit une capabilité, soit elle ne le fait pas. On ne peut être « adéquatement nourri » à moitié. Par conséquent, l'exercice d'évaluation des fonctionnements doit considérer chaque fonctionnement comme un tout, dont la valeur dépend de son adéquation aux valeurs, aux projets et aux finalités de l'agent[466].

[465]SEN, Amartya, *Commodities and Capabilities*, Amsterdam, North Holland, 1985.

[466] Dans le cas où l'agent s'avère incapable de procéder à l'évaluation elle-même (le cas d'un très jeune enfant, par exemple), l'évaluation positionnelle que nous avons développée dans la section 5.1. à partir des travaux de Sen s'appliquerait ici.

c) La structure narrative des préférences pour les fonctionnements

Il est intéressant d'associer l'évaluation des capabilités à une approche narrative des préférences. Plutôt que d'envisager l'objet des préférences comme une agrégation d'éléments séparés, il peut-être plus judicieux de concevoir des préférences pour des objets globaux[467]. Dans un article consacré à l'autonomie narrative et au bien-être, Johan Brännmark signale la structure narrative des existences humaines : celles-ci pourraient être décrites comme un tout, un processus continu, dans une perspective holiste, plutôt que comme la somme de parties[468]. Cette structure narrative permet de discriminer les préférences selon leur objet : certaines préférences portent sur des caractéristiques temporaires de nos existence, tandis que d'autres concernent des états et des activités durables. Les préférences qui importent pour évaluer notre bien-être bien compris sont donc, selon Brännmark, les préférences pour des états et des activités durables, c'est-à-dire les préférences dont l'objet se caractérise par sa structure narrative. Or, il s'avère que les fonctionnements décrivent des processus et des activités durables, et non pas des états de satisfaction ou de possession de ressources[469]. Par conséquent, si l'évaluation sociale repose sur les préférences pour des capabilités de fonctionnements, les préférences qui importeront seront celles qui ont pour objet des caractéristiques durables des états sociaux. Cela permet, comme le suggère Brännmark, d'exclure les préférences pour des états à court terme. Il y aurait ici une justification possible pour restreindre le domaine des préférences admissibles dans le contexte d'une formalisation de l'évaluation des capabilités, restriction qui permettrait d'échapper au résultat d'impossibilité d'Arrow.

Si les personnes ayant des préférences de structure narratives sont appelées à se considérer elles-mêmes comme les auteurs et les protagonistes de leur existence, justifiant ainsi le recours à la notion d' « autonomie narrative », il n'en reste pas moins, comme le souligne Brännmark, que cette structure narrative met en évidence le fait que nos existence sont ancrées dans un environnement social. Les préférences pour des états temporaires tendent à ignorer cet ancrage, car leur pauvreté descriptive les amène à ignorer les processus sociaux qui sous-tendent les existences individuelles. Par ailleurs, l'évaluation des capabilités et des

[467] Cf. PARFIT, Derek, *Reasons and Persons*, Oxford :Clarendon Press, 1984, Appendix I.

[468] BRÄNNMARK, Johan, "Leading a Life of One's Own: On Well-Being and Narrative Autonomy", in OLSARETTI, Serena (ed.), *Preferences and Well-Being*, Cambridge University Press, 2006

[469] Brännmark n'établit pas de lien entre sa proposition et l'évaluation des fonctionnements. Mais l'analogie de structure justifie le rapprochement.

fonctionnements partage certaines caractéristiques communes avec la conception rawlsienne des « projets de vie »[470] :

> *« Tout d'abord, les désirs et les besoins humains ont certaines caractéristiques d'ensemble, une urgence relative et des cycles de récurrence, des phases de développement affectées par des facteurs physiologiques et autres. En second lieu, les projets doivent s'accorder avec les exigences des capacités et des compétences humaines, leur maturation et leur développement, leur éducation et leur formation en vue de tel ou tel but. En troisième lieu, je postulerai un principe essentiel de motivation, le principe aristotélicien. Enfin, il faut tenir compte des faits généraux et de l'interdépendance sociale. »*

Certes, la définition des projets de vie par Rawls vise à établir un indice uniforme de biens premiers, alors que nous cherchons les conditions d'élaboration d'indices personnalisés de capabilités. Néanmoins, la structure du problème est la même, dans la mesure où les objets évalués sont des objets évolutifs, dynamiques, et interagissant avec le contexte social et environnemental. Les « préférences » énoncées dans la position originelle pour certains projets de vie s'opposent donc, au même titre que les préférences pour les capabilités établies dans des conditions non-idéales, aux préférences pour des états de choses brefs, fixes et détachés de leur contexte. Lorsque nous réfléchissons aux activités que nous souhaitons entreprendre dans cette perspective dynamique et contextualisée, nous devons envisager les relations avec les autres personnes : vont-elles contribuer à la réalisation de nos fins, ou au contraire s'y opposer ? Nos finalités sont-elles raisonnables du point de vue d'autrui ? Peut-on recourir à la solidarité ? L'évaluation de fonctionnements exercés sur le long terme se produit quasi fatalement dans un contexte social, et rend d'autant plus probable la prise en compte de l'autre dans le processus de réflexion sur les capabilités.

[470]RAWLS, John, *A Theory of Justice*, The Belknap Press of Harvard University Press, 1971; traduction française par Catherine Audard sur un texte révisé par l'auteur en 1975: *Théorie de la Justice*, Paris, Editions du Seuil, 1987, 1997, § 65.

d) Préférences globales pour les fonctionnements et degré d'adaptation

Revenons à présent à la question des préférences adaptatives. Celles-ci se caractérisent comme étant des réponses passives à la restriction, réelle ou perçue, de l'ensemble des options faisables (cf. section 5.2.2. de notre travail). Les préférences adaptatives sont des préférences qui sont formées en fonction de l'environnement de choix, c'est-à-dire de la nature et de l'étendue des options offertes à l'agent (ou de celles qu'elle perçoit comme telles).

En soi, que les préférences s'adaptent à l'environnement de choix n'a rien de préoccupant. C'est même une question de santé mentale que de ne pas toujours désirer ou vouloir l'impossible. Les préférences adaptatives sont problématiques lorsque qu'elles dépendent exclusivement de l'environnement de choix, au prix même de ce que l'agent pourrait avoir de bonnes raisons d'évaluer.

Notre hypothèse est que la structure des préférences globales que sont les préférences pour les fonctionnements les protège davantage d'une dépendance excessive au contexte que la structure des préférences sommatives. Une préférence sommative est une somme de préférences pour des états locaux du monde, et non pour le monde dans sa globalité. Par conséquent, une préférence « locale », pour reprendre la terminologie de Parfit, s'énonce à partir d'un ensemble d'opportunités défini en un temps et en un lieu donné. Autrement dit, les options offertes au choix sont appréhendées en fonction de ce temps et de ce lieu donné, et plus généralement d'un contexte donné. Cette caractéristique des préférences locales favorise la dépendance au contexte, dans la mesure où l'agent n'a pas l'opportunité de comparer ce contexte précis avec d'autres contextes de choix. Nous avons déjà mis en évidence dans la section 5.2. de notre travail l'intérêt de définir des options extérieures à l'environnement de choix pour lutter contre les préférences adaptatives au point d'être une base d'information problématique. En comparant différents contextes de choix, l'agent est plus à même d'adopter une attitude critique à l'égard des options qui lui sont offertes. Mais les préférences locales ne permettent pas cela.

En revanche, les préférences globales sont des préférences pour des constituants de l'existence humaine que l'on peut considérer comme des touts, voire même pour une existence dans son ensemble. Or, une période entière de la vie, ou une vie entière, n'est pas localisée dans un contexte unique. Une vie raisonnablement longue voir le monde qui l'environne changer. La conjoncture sociale et économique, la culture et les coutumes, le

climat, l'entourage, l'aspect physique et les capacités intellectuelles, évoluent au cours d'une vie. Une préférence globale est une préférence pour des états de choses qui peuvent advenir au sein de différents contextes. C'est, en un sens, une préférence « transcontextuelle » ou « transmondes »[471]. Une telle préférence met à profit l'intérêt critique qu'offrent les comparaisons de différents contextes.

L'exemple récurrent de préférence adaptative qui inquiète tous ceux qui ont fait leur l'approche des capabilités est celui des préférences exprimées par des femmes dans des sociétés conservatrices pour des conditions de vie déplorables[472]. Une femme battue peut choisir de demeurer dans le foyer où elle endure vexations, violences et privations et sembler préférer cela. Est-ce une préférence pour un fonctionnement? Est-ce une préférence locale ou globale? Lorsque l'on interroge son comportement, il apparaît que celui-ci est une réponse à sa situation particulière, en un temps et en un lieu donnés. Elle pourrait n'avoir ni revenue ni lieu où se réfugier. Elle pourrait croire que la violence domestique est le lot inévitable de toute femme mariée: sa mère, ses soeurs, ses amies pourraient endure un sort similaire dans les milieux les plus sexistes. Peut-être juge-t-elle qu'elle mérite les coups reçus pour un plat trop salé ? Mais si, réfléchissant sur l'ensemble de son existence, si, appelée à comparer les différentes vies possibles qui s'offrent à elle, sa situation présente et ses expériences passées comme ses opportunités futures, est-ce si sûr qu'elle soutiendra *mordicus* que son sort est celui qu'elle a toujours voulu et qu'elle souhaitera jusqu'à sa mort ? C'est peu probable.

Bina Agarwal a remis en cause l'idée selon laquelle les femmes vivant dans des sociétés pauvres et conservatrices adhéreraient complètement aux normes qui les oppressent[473]. Les études empiriques ont montré qu'une adhésion apparente dissimule souvent en réalité des stratégies de résistance qui montrent bien que la personne perçoit sa situation comme défavorable à son bien-être. Elle cite des observations de terrain au Pakistan : les femmes dissimulent du riz ou des chèvres pour les revendre et avoir un revenu en dépit de l'interdiction de travailler, dissimulent la nourriture en principe réservée aux hommes. Ces stratégies de résistance peuvent être rapprochées de celles des travailleurs à la chaîne qui

---

[471]H. E. Baber utilise la logique des mondes possible pour caractériser sa proposition préférentialiste d'évaluation des capabilités (il ne reserve pas de traitement spécifique aux fonctionnements, contrairement à ce que nous proposons ici, et conçoit la valeur des capabilités du point de vue de leur aptitude à satisfaire les preferences actuelles ou probables de l'agent. Les préférences probables sont les préférences susceptible d'advenir dans des mondes possible dont les caractéristiques sont proches du nôtre). Le terme "transmonde" ("*transworld"*) appliqué aux préférences apparaît dans son article à ce sujet. Cf. BABER, H. E., « Worlds, Capabilities and Well-Being", *Ethical Theory and Moral Practice,* Vol 13 (4), 2010, p 377-392.

[472].NUSSBAUM, Martha, « Adaptive preferences and women's options », *Economics and Philosophy*, vol. 17, Issue 01, 2001, pp 67-88

[473] AGARWAL, Bina, « 'Bargaining' and Gender Relations Within and Beyond the Household », *Feminist Economics* 3, no. 1, 1997.

trouvaient le moyen de s'octroyer des pauses en dépit de l'interdiction de leurs supérieurs. L'obéissance n'implique pas l'acceptation.

Plus encore, Alison Jaggar a questionné l'asymétrie entre le jugement de l' « expert » (le philosophe, le théoricien des capabilités, l'économiste…) et celui des « pauvres » dans le compte-rendu critique qu'elle a fait de l'usage du raisonnement pratique dans la construction de la liste des capabilités de Martha Nussbaum[474]. Il faut donc être assez prudent lorsque l'on taxe une préférence d' « adaptative ». Prudent, parce que les biais épistémiques sont plus que jamais possibles : comme l'a montré Agarwal, une observation trop hâtive ne parviendra pas vraiment à rendre compte de l'appréciation authentique de leur situation par les personnes. Prudent aussi, parce que, dans son esprit même, l'approche par les capabilités s'est voulue respectueuse des valeurs de chacun et à l'écoute des voix des pauvres ; écarter du revers de la main des préférences parce qu' « adaptatives » pourrait en fait relever d'une approche somme toute paternaliste, voire condescendante, voire même quelque peu néocolonialiste lorsque – comme cela arrive encore souvent – l' « expert » appartient à l'ancienne puissance coloniale et le « pauvre » à l'ancienne colonie…

Le degré de dépendance au contexte nous semble un critère suffisamment neutre pour éviter ces écueils et permettre de distinguer les préférences adaptatives pertinentes pour l'évaluation de la pauvreté et des inégalités de celles qui ne le sont pas, sans pour autant imposer une norme extérieure du bien-être que les personnes appelées à évaluer leur propre situation n'endosseraient pas nécessairement.

[474]JAGGAR, Alison, « Reasoning about Well-Being : Martha Nussbaum's methods of Justifying the Capabilities", *The Journal of Political Philosophy,* volume 14, Issue 3, p 301-322, sept. 2006.

7.2.4. Des préférences pour les fonctionnements à l'évaluation des ensembles capabilité

Jusqu'ici, nous nous en sommes tenus à l'évaluation des fonctionnements. Qu'en est-il des capabilités ? Quelle est leur valeur ajoutée ?

a) Evaluation des capabilités et préférences pour les fonctionnements

Nous travaillons ici avec les capabilités définies comme des ensembles d'opportunités de fonctionnements, dans la mesure où cette conception permet d'englober les considérations relatives à la liberté. Par conséquent, attribuer une valeur à une capabilité parmi d'autres capabilités revient à classer des capabilités, c'est-à-dire des ensembles d'opportunités. Mais nous devons encore déterminer si le classement et la pondération des ensembles capabilités est fonction de leur contenu ou de leur étendue (ou les deux à la fois). Notre problème général est donc le suivant : en quel sens et dans quelle mesure la valeur des capabilités est-elle basée sur les préférences pour les fonctionnements ?

Une première méthode d'évaluation des ensembles d'opportunités – de fonctionnements dans le cas des capabilités - est l'évaluation élémentaire. L'évaluation élémentaire, selon la terminologie de Sen, stipule que la valeur d'un ensemble capabilité est égale à la valeur de son meilleur élément[475]. Par conséquent, la valeur d'un ensemble capabilité serait déterminée par la préférence pour la combinaison de fonctionnements qui figure parmi les éléments de cet ensemble capabilité, et ce quels que soient les autres éléments. Mais cette évaluation élémentaire fait face à deux objections. La première concerne la nature du « meilleur » élément : est-ce le meilleur en termes de bien-être objectif, ou selon les préférences et les valeurs de l'agent ? La seconde est la suivante : pourquoi devrions-nous évaluer les capabilités au lieu des fonctionnements si, après tout, ce qui fait la valeur d'un ensemble capabilité est son meilleur fonctionnement ? Ce qui est en jeu dans l'évaluation des capabilités, c'est que l'agent a accès à un ensemble de différentes options, et pas seulement à la meilleure d'entre elles. En d'autres termes, les capabilités ont de la valeur parce qu'elles impliquent des scénarios contrefactuels de choix.

Une seconde méthode serait l'évaluation cardinale de la liberté. Elle consiste à évaluer les ensembles d'opportunités en fonction de leur étendue. Des travaux formels ont été proposés

[475]SEN, Amartya ,*Commodities and Capabilities,* Amsterdam: North-Holland, 1985, p 39.

par Pattanaik et Xu dans cette direction[476]. Leur proposition consiste à définir le degré de liberté en fonction du nombre d'options disponibles pour l'agent. Plus un ensemble d'opportunités inclut d'options, plus il a de la valeur en termes de liberté. Ici, l'intuition de départ serait que le plus grand nombre de préférences possibles pour des fonctionnements devrait pouvoir être satisfait. Cette approche a été discutée parce qu'elle ne rend pas compte de la qualité des options. Elle implique que, si nous ajoutons des options telles que « être tué » ou « être torturé » à l'ensemble d'opportunités, celui-ci verra sa valeur s'accroître du simple fait qu'il contient plus d'options. Même si, *a priori,* une personne peut avoir toutes les préférences possibles, *a posteriori,* il s'avère que nos préférences sont dans une bonne mesure assez similaires. Nous souhaitons à peu près tous boire de l'eau potable et être traités avec respect[477]. Dans cette perspective, l'ensemble *a priori* le plus étendu possible de fonctionnements ne fait guère sens.

Une troisième réponse pourrait être formulée en termes de « préférences pour les capabilités ». Les préférences ne détermineraient pas seulement quels fonctionnements ont de la valeur pour l'agent, mais également dans quelle mesure l'agent valorise sa liberté de choix. Par exemple, un agent qui croit que la liberté de choix n'a aucune valeur préférera avoir un ensemble capabilité qui inclut seulement les fonctionnements qu'elle préfère. Une personne qui prend plaisir au choix ou qui craint les changements de conjoncture pourrait inclure dans son ensemble capabilité le plus de combinaisons de fonctionnements possible, y compris des fonctionnements pour lesquels elle n'a aucune préférence.

Cette dernière solution laisse le choix même de se situer dans l'espace des capabilités ou dans celui des fonctionnements à l'arbitre de l'agent. L'approche par les capabilités n'aurait même plus de raison philosophique de défendre la liberté de choix, puisque les agents pourraient préférer ne pas avoir cette liberté.

La valeur d'un ensemble capabilité devrait donc être influencée par les préférences pour les fonctionnements sans cependant être déterminée par celles-ci. Les préférences changent, et les personnes apprécient d'avoir la liberté de choisir et de comparer les différentes options qui s'offrent à elles. Une piste possible est d'inclure dans les ensembles capabilités des combinaisons de fonctionnements pouvant être préférées, étant données les caractéristiques de

[476]PATTANAIK, P. K., and XU, Y., "On Ranking Opportunity Sets in Terms of Freedom of Choice", *Recherches Economiques de Louvain*, 56: 383-390, 1990.

[477] Je remercie Robert Heeger de m'avoir aidée à souligner ce point.

la position occupée par l'agent et la similitude des préférences que nous avons signalée ci-dessus[478].

[478]Baber propose une solution de cet ordre en utilisant le critère de "fécondité" des états de chose. Un état de chose valable est un état de choses "fécond" au sens où il est susceptible de satisfaire non seulement les préférences actuelles de la personne, mais également les préférences qu'elle pourrait avoir dans des mondes possible similaires au monde actuel. Cf. Baber, « Worlds, Capabilities and Well-Being ».

b) *Remarques sur la délibération comme comparaison d'opportunités*

Le fait de devoir sélectionner les fonctionnements que l'on préfère dans un ensemble capabilité nous oblige également à considérer avec davantage d'attention les autres fonctionnements. En effet, lorsque les personnes comparent les éléments de leur ensemble capabilité, cette comparaison implique leur perception des opportunités. Les agents ne pensent pas qu'à ce qu'elles veulent accomplir, mais également à ce qu'elles auraient pu faire d'autre. Elles considèrent des états contrefactuels qu'elles n'auraient peut-être pas choisis à première vue. Si unepersonne raisonne sur les opportunités, elle est amenée à s'interroger sur la faisabilité de résultats qu'elle n'aurait pas spontanément choisis. L'activité qui consiste à réfléchir sur nos opportunités implique également une réflexion sur la perception, les émotions et la capacité de prendre des décisions. Tout cela constitue nos opportunités réelles. Nous avons vu dans le paragraphe c) de la section 5.2.2. que les préférences adaptatives peuvent être efficacement contrées par l'exercice qui consiste à envisager des possibilités extérieures, ou perçues comme extérieures à l'environnement de choix. Le second exercice serait ici d'envisager les possibilités contrefactuelles après avoir effectué le choix, ou au moment de l'effectuer. Le fait de considérer non seulement ce qui va arriver, mais également ce qui pourrait arriver, ou encore ce qui aurait pu être, ainsi que les processus, actions et causes ayant conduit à ces états de fait, peut modifier la perception de faisabilité. D'autre part, un agent peut éprouver de la satisfaction du seul fait qu'elle dispose d'une vaste étendue d'options. Comme l'écrit François Hudon, nous souhaitons généralement avoir davantage d'argent (sans envisager bien entendu le cas pathologique de l'avarice), même si nous ne comptons pas en dépenser la totalité : nous préférons simplement disposer d'une liberté de choix qui se traduit par l'existence d'un ensemble d'options raisonnablement étendu et diversifié[479]. Raisonner sur les opportunités pourrait ainsi contribuer à l'émancipation des personnes, grâce aux activités de conception et d'imagination des possibles et des contrefactuels. Cette thèse est encore à l'état de conjecture, et elle gagnerait sans doute à être étayée par des travaux empiriques.

Les deux principaux arguments proposés par Sen en faveur d'une perspective de l'évaluation fondée sur les libertés sous la forme de capabilités sont les suivants :

[479] HUDON, François, « La liberté dépend-elle des préférences individuelles ? », *Raisons politiques*, 3, n° 43, 2011, pp 35-58.

(3) La liberté aurait une valeur instrumentale, car plus la personne a de choix et plus elle a de chances de réaliser son bien-être et ce qu'elle valorise.

(4) La liberté aurait une valeur intrinsèque, dans la mesure où l'acte de choix aurait de la valeur lui-même.

Notre brève analyse de la délibération sous la forme de la comparaison de combinaisons de fonctionnements dans un ensemble capabilité met en évidence une valeur peut-être inattendue de la perspective des capabilités, que l'on peut qualifier de « valeur épistémique ». Le fait de disposer de la liberté de choix permet à la personne de comparer ce qu'elle valorise à des options qu'elle n'aurait pas envisagé autrement, et ainsi de renforcer sa préférence grâce à l'examen critique qui procède de toute comparaison. La personne peut également être amenée à remettre en cause sa préférence première. Il est à notre que même dans un monde déterminé, où la liberté de choix ne serait qu'illusoire, la connaissance des alternatives apparentes permettrait à la personne de mieux appréhender son environnement et ce qui la détermine à suivre une voie plutôt qu'une autre.

Nous avons ainsi énoncé une proposition permettant d'articuler le cadre conceptuel préférentialiste et démocratique de la théorie du choix social avec les capabilités, paradigme d'évaluation soucieux à la fois de la liberté de choix et du bien-être des personnes. L'approche par les capabilités requière une base d'information plus riche, et cela passe par la relaxation d'une des deux propositions qui constituent la condition arrovienne d'indépendance des alternatives non pertinentes. D'autre part, la théorie du choix social nous permet de préciser que la préférence peut et même doit être comprise comme une évaluation comparative davantage que comme une expérience psychologique. Cela étant, il nous a été possible de proposer une méthode préférentialiste d'évaluation des capabilités, passant par les préférences pour les fonctionnements. Les ensembles capabilités incluent des combinaisons de fonctionnements susceptibles d'être préférées (même si le fait d'être « susceptible d'être préféré » ne peut être déterminé qu'*a posteriori*). Cela signifie que l'acte de choix consiste à comparer les différentes combinaisons de fonctionnements dans un ensemble capabilité. Une telle comparaison peut s'avérer profitable pour lutter contre une excessive dépendance des préférences à l'égard de l'environnement immédiat, ainsi que pour mieux connaître les différentes déterminations possibles de notre existence.

Conclusion

Les notions de choix collectif et de préférence ne sont pas techniques par elles-mêmes, ni étrangères au sens commun. Le choix de traiter le choix collectif comme agrégation des préférences est donc un pari risqué, dans la mesure où nous ne sommes jamais loin du malentendu lorsque vient le moment de confronter cette approche à nos habitudes de langage. Non que les définitions que nous avons proposées soient totalement étrangères à ces habitudes ; mais, en visant à les préciser et à en éliminer les contradictions, le risque de ne pas se faire comprendre subsiste. On pourrait alors reprendre à notre compte ce que Spinoza dit à propos du vocabulaire scolastique qu'il utilise :

> *« Je sais que ces mots ont dans l'usage ordinaire un autre sens, mais mon dessin est d'expliquer la nature des choses et non le sens des mots, et de désigner les choses par des vocables dont le sens usuel ne s'éloigne pas entièrement de celui où je les emploie, cela soit observé une fois pour toutes. »*[480]

Le pari était d'autant plus risqué que les notions que nous avons traitées appartiennent aujourd'hui en premier lieu à la discipline de l'économie, alors qu'il s'agissait ici d'une thèse de philosophie, ayant pour objet des concepts et des questionnements fondamentaux d'épistémologie, de philosophie morale et de philosophie politique. Le dialogue entrepris avec les économistes durant ce travail de recherche a permis d'esquisser les divisions du travail possibles, sans toutefois rechercher à tout prix le cloisonnement disciplinaire.

[480] SPINOZA, Baruch, *Ethique*, III. Traduction Charles Appuhn, Paris : Vrin[2], 1977.

1. Un bref bilan

Les résultats de notre travail surprendront ceux qui voient dans l'agrégation des préférences un dispositif nécessairement individualiste, orienté par la maximisation de l'intérêt personnel à l'exclusion de toute considération morale ou altruiste, et voué à l'échec de ce fait. Notre thèse est que le cadre conceptuel de la théorie du choix social et de l'économie normative n'implique pas une telle interprétation (même s'il ne l'exclut pas), et qu'il peut de ce fait constituer une base pour une théorie démocratique de la justice ayant pour métrique les capabilités, dans l'esprit des idées développées par Amartya Sen dans l'*Idée de Justice.* L'étude des textes de Bentham, de Pareto, d'Arrow et de Sen, pour ne citer qu'eux, ainsi que l'analyse des concepts de d'agrégation, de préférence et de capabilité, a permis de proposer des arguments en faveur de cette thèse, arguments que nous espérons convaincants.

Notre première partie a fait la genèse du rôle joué par l'individu dans les contributions correspondant à des étapes importantes de l'histoire de l'agrégation des préférences. Cette analyse historique et critique du présupposé individualiste du cadre conceptuel de la théorie du choix social a mis en évidence que ce présupposé ne relevait pas d'une nécessité logique, mais plutôt d'un choix interprétatif arbitraire. Le contexte même d'application d'une théorie du choix social et de l'évaluation collective suppose un individu inséré dans des rapports sociaux. L'individu n'est pas une donnée naturelle ; ses finalités sont reconstruites différemment selon le mode d'expression de ses attentes choisi par le théoricien. Il est donc peut-être moins arbitraire de parler de positions plutôt que d'individus pour désigner la source des évaluations substantielles que la théorie du choix social et l'économie du bien-être se chargent par la suite d'agréger.

Cependant, l'agrégation, par elle-même, suppose une hétérogénéité des préférences ou des valeurs. Mais cette hétérogénéité peut traduire des divergences entre des groupes, des positions, des situations, et pas nécessairement entre des individus. Réfuter le caractère nécessairement individualiste de la théorie du choix social ne remet donc pas en question l'existence d'un problème du choix social et la pertinence du résultat d'Arrow pour exprimer certaines difficultés inhérentes à l'existence collective.

Enfin, nous avons mis en évidence la possibilité de concevoir le processus même d'agrégation des préférences comme un constituant de l'individualité. A travers la définition de ses attentes et le mode de formulation de ces dernières, l'individu se construirait peu à peu, tel un

regroupement de qualités, au travers du triple prisme de l'introspection, du regard de l'autre et du point de vue de la collectivité elle-même.

L'analyse conceptuelle de la préférence que nous avons proposée dans la IIe partie a dissipé certaines ambiguïtés, élargi le champ des interprétations possibles du cadre conceptuel de la théorie du choix social, et permis des choix interprétatifs cohérents et pertinents compte tenu de la nature de notre projet, qui était d'associer conceptuellement théorie du choix social au sens large et approche par les capabilités.

Concernant la nature de la préférence, nous avons défendu une conception l'assimilant à un jugement de valeur comparatif disposant éventuellement au choix et à l'action, mais sans en être cependant une cause inévitable. Eu égard à la position de l'évaluateur et au choix du critère du « préférable », l'enjeu déterminant est la légitimité de l'attribution d'une préférence à une personne – individuelle ou collective - . Une attribution à long terme de la préférence autorise une énonciation de celle-ci par un évaluateur externe, bien informé et bienveillant, à condition que l'agent puisse se reconnaître à long terme dans la préférence et l'identifier comme sienne. La construction d'une préférence apparaît donc comme un processus délibératif long et pouvant impliquer d'autres personnes jouant un rôle de conseiller. Nous avons également découvert qu'en mettant de côté les conceptions substantielles du bien-être et du bien, la connaissance et le mode de description des états sociaux deviennent cruciaux pour la détermination de la préférence. En fait, la rationalité de la préférence dépend étroitement de l'information dont dispose l'agent sur les états sociaux, leurs caractéristiques, leur faisabilité, leurs processus de production, leur devenir, la situation de l'agent par rapport à ces états ainsi que celle de ceux dont il se soucie…Le bon usage de cette information dépend également de la capacité de l'agent à imaginer des alternatives improbables ou encore à se représenter sous des couleurs plus vives les traits les plus pertinents pour la satisfaction de ses goûts comme pour la réalisation de ses valeurs.

Notre IIIe partie a jeté les bases d'une approche « démocratique » de l'évaluation des capabilités. Les problèmes relatifs à la construction d'une liste de capabilités et aux décisions d'arbitrage les concernant y sont abordés comme des problèmes sociaux et politiques de décision collective, et non comme des problèmes moraux de détermination de la norme de la vie bonne. La décision collective est toutefois, et c'est indispensable, nourrie par la description empirique des états sociaux, et leur comparaison. A la base des évaluations

collectives des capabilités, se trouveraient donc des préférences dont la valeur cognitive est déterminée par quatre facteurs que nous avons mis en évidence au cours de notre travail :

(1) Le rattachement de la préférence à une position pouvant être décrite objectivement, plutôt qu'à une source individuelle opaque et mystérieuse ;

(2) La thèse selon laquelle le bien-être comme satisfaction des préférences requiert une parfaite connaissance des options ou des états sociaux évalués pour que l'on puisse à proprement parler attribuer la préférence à la personne qui va subir les effets de ces options ou de ces états sociaux ;

(3) Le filtre des fonctionnements comme objets de préférence conduit à prendre en compte des préférences globales plutôt que locales, c'est-à-dire des préférences pour des existences, des activités et des états au long court. Les préférences pour les fonctionnements se caractérisent donc par la structure narrative de leur objet, ce qui constitue un filtre supplémentaire.

(4) Enfin, l'activité de choix d'une combinaison de fonctionnements constituant une existence humaine prend la forme d'une comparaison de fonctionnements alternatifs dans un ensemble capabilité. Cette activité possède en soi une valeur épistémique : nous apprenons beaucoup de la valeur respective de chaque fonctionnement en les comparant les uns aux autres. La perspective des capabilités enrichit ainsi non seulement celle des fonctionnements en protégeant la personne des risques de changement de préférence comme de conjoncture, ou en garantissant son autonomie et sa liberté de choix, mais également en lui offrant l'opportunité, si ce n'est de choisir, du moins celle de juger et d'évaluer une multiplicité d'existences possibles.

Ce travail de recherche a mis peu à peu en évidence l'importance de la connaissance pour la conception du choix collectif comme agrégation des préférences. L'histoire de cette conception, que nous avons retracée dans notre introduction, a vu réapparaître régulièrement, chez Raymond Lulle comme chez Bentham, chez Condorcet comme chez Amartya Sen, le souci de définir une conception de l'agrégation des préférences qui soit satisfaisante pour effectuer le « bon choix », ce bon choix étant aussi celui qui est le plus près de la vérité. Cette « vérité » du choix collectif n'est pas une entité subsistant indépendamment de la société ou du groupe concernés : il s'agit de la vérité des attentes et des revendications des personnes concernées, qui doit être représentée au mieux, dans toute sa diversité, par le choix collectif. Ici, le souci d'éviter l'arbitraire logique rejoint les inquiétudes causées par l'arbitraire moral

ou politique. Il faut éviter la manipulation des votes, une description inappropriée des attentes, une échelle de comparaison arbitraire, la dictature au sens formel du terme, ou encore une base informationnelle trop restreinte, parce que, justement, le vrai choix collectif est celui qui est représentatif des préférences des membres de la collectivité, et c'est dans ce caractère représentatif que réside son acceptabilité, ce que Condorcet appelait la « justice » et la « raison ».

2. De nouvelles pistes de recherche

Ce travail n'est en fait qu'une ébauche, que des recherches futures permettront de préciser, de compléter ou d'orienter vers d'autres problématiques philosophiques.

Notre troisième partie a mis en évidence la possibilité d'utiliser une analyse conceptuelle précise des conditions de validité de la préférence en vue de jeter les bases d'une théorie démocratique de la justice. Le concept de préférence joue en fait un rôle crucial dans une bonne partie des théories contemporaines de la justice : comme le souligne Jon Elster dans sa contribution sur les préférences adaptatives,

> *« En fait, le problème est pertinent pour toutes les théories éthiques et de la justice qui prennent en considération le vouloir. La théorie de John Rawls semble échapper à cette difficulté, parce qu'elle repose sur les biens premiers plutôt que sur les utilités ou les préférences. Mais en fait même sa théorie nécessite la préférence afin de comparer des paniers non-dominés de biens premiers [...] »*[481]

Les seules théories morales et politiques normatives pouvant échapper à la préférence sont celles qui imposent de l'extérieur une conception du Bien indépendante des attentes, aspirations, valeurs et désirs des membres de la société (et même, dans ce cas, la préférence peut jouer le rôle de révélateur, comme nous l'avons vu dans la section 5.2.3. a) de notre travail).

Aborder la question de l'agrégation des préférences pour les capabilités par l'intermédiaire de la théorie du choix social peut sembler être une voie tortueuse : pourquoi ne pas simplement interpréter l'approche par les capabilités comme un utilitarisme amélioré, au moins dans sa forme ? Mais la théorie du choix social apporte deux dimensions qui ne sont pas intrinsèques à l'utilitarisme. D'une part, en se concentrant sur les ordres de préférences plutôt que sur la mesure de la satisfaction, elle favorise une interprétation de la préférence en termes de jugements plutôt que de désirs. Nous avons vu avec Jonathan Riley que les normes

[481] ELSTER, Jon, "Sour grapes – utilitarianism and the genesis of wants", *in* Sen, Amartya Kumar and Williams, Bernard (eds), *Utilitarianism and beyond*, Cambridge: University Press, 1982.

procédurales utilitaristes ne sont équivalentes aux normes démocratiques que lorsque la base d'information est définie en termes de préférences et non d'utilités[482]. D'autre part, rappelons que l'agrégation des préférences dans la théorie du choix social relève d'une approche à profils multiples : le théoricien du choix social ne cherche pas à définir un choix collectif pour un profil particulier de préférences, mais souhaite définir une règle d'agrégation valable pour tous les profils de préférences possibles. En d'autres termes, ce n'est pas l'évaluation contingente qui est recherchée, mais la garantie constitutionnelle : garantie selon laquelle, quelles que soient les préférences des membres de la société, le système de choix collectif en dérivera toujours une évaluation collective.

La poursuite de ce travail exigera une analyse conceptuelle de l'agrégation des préférences pour les capabilités. En travaillant sur une préférence reconsidérée comme jugement informé, ainsi que sur les problèmes de commensurabilité propres à la procédure d'agrégation, il sera alors possible d'envisager des modes d'introduction de la délibération au sein même de l'opposition classique entre les deux[483].

Le concept même de préférence pour les capabilités peut faire l'objet d'une investigation approfondie sur l'influence de la perception des opportunités sur la délibération et le choix en général. Cette investigation pourrait être mise en relation avec des travaux en sociologie, en sciences politiques et en économie.

D'autre part, la poursuite d'une réflexion sur le choix collectif consiste en une réévaluation des formes historiques et contemporaines de la démocratie, au sens le plus large du terme. L'approche par les capabilités manque encore à ce jour d'une réflexion sur la nature et la forme des institutions les plus à même de contribuer à son avancée ; tout ce que nous savons est qu'elles doivent être « démocratiques ».. Il conviendra de déterminer à la fois les procédures de décision appropriées à des contextes aussi divers que la décision internationale, locale, en entreprise, dans la famille, dans les associations, les syndicats, les universités...

Si la dernière partie de notre travail oriente le cadre conceptuel de la théorie du choix social vers le terrain des théories de la justice sociale, d'autres usages des analyses des deux

[482]RILEY, Jonathan, « Utilitarian Ethics and Democratic Government », *Ethics*, Vol. 100, No. 2, Janv. 1990, pp 335-348.

[483] John Dryzek et Christian List ont montré par ailleurs que la tradition du choix social peut être réconciliée avec celle de la démocratie délibérative : en réévaluant les implications des théorèmes d'Arrow et de Gibbard-Satterwaithe, il apparait que la délibération favorise les conditions dans lesquelles une prise de décision démocratique est possible. Cf. DRYZEK, John S., LIST, Christian, « Social Choice Theory and Deliberative Democracy : A Reconciliation », *British Journal of Political Science*, 2003.

premières parties sont possibles. On pourrait par exemple s'interroger de manière critique sur les raisons pour lesquelles une interprétation très particulière et réductrice du cadre conceptuel de la théorie du choix social a prévalu. Notre analyse conceptuelle motiverait ainsi un travail de généalogie critique du discours interprétatif du cadre formel welfariste. Ces analyses pourraient, si elles sont validées par la suite, contribuer à l'épistémologie des branches de l'économie qui partagent avec l'économie du bien-être certaines de ses notions, notamment celle d'utilité.

Bibliographie

Monographies

ALKIRE, S., *Valuing Freedoms: Sen's Capability Approach and Poverty Reduction,* New York: Oxford University Press, 2002.

ANDERSON, Elizabeth, *Value in Ethics and Economics,* Cambridge and London: Harvard University Press, 1993.

ARISTOTE, *Ethique à Nicomaque.* Trad. J. TRICOT, Paris : Vrin, 1994.

ARISTOTE, *Physique.* Trad. P. PELLEGRIN, Paris: Garnier-Flammarion, 1999.

ARISTOTE, *Politiques.* Trad. P. PELLEGRIN, Paris: Garnier-Flammarion, 1990.

ARISTOTE, *Topiques.* Trad. J. TRICOT, Paris: Vrin, 2004.

ARROW, Kenneth, *Social Choice and Individual Values*, New Haven, Yale university Press, 1951 (2ème édition révisée, 1961).Traduction française par l'Association de Traduction Economique de l'Université de Montpellier, Paris, Calmann-Levy, 1974.

ARROW, Kenneth, *Collected Papers,* Cambridge, MA: The Belknap Press of Harvard University Press, 1983.

BENMAKHLOUF, Ali, *Frege. Le nécessaire et le superflu,* Paris : Vrin, 2002.

BENMAKHLOUF, Ali (éd.), *Pauvreté et richesse : perspectives économiques et philosophiques,* Casablanca : Le Fennec, 2010.

BENMAKHLOUF, Ali, *L'identité : une fable philosophique,* Paris : Presses Universitaires de France, 2011.

BENTHAM, Jeremy, *An Introduction to the principles of Morals and Legislation,* Londres, 1780, 1823. Réédition Oxford, Clarendon Press, 1907.

BLACK, Duncan, *The Theory of Committees and Elections,* Cambridge: Cambridge University Press, 1958.

BORDA, J.-C., "Mémoire sur les élections au scrutin", *Mémoires de l'Académie Royale des Sciences,* 1781, p 657-665.

BOURDIEU, Pierre, *La distinction, critique sociale du jugement*, Paris : Editions de Minuit, 1979.

BRIGHOUSE, H. et ROBEYNS, I. (ed.), *Measuring Justice: Primary Goods and Capabilities,* Cambridge University Press, 2010.

BRANDT, Richard, *A Theory of the Good and the Right,* Oxford : Oxford University Press, 1979.

BRENNAN, Jason, *The Ethics of Voting,* Princeton University Press, 2011.

BROOME, John, *Weighing Goods*, Oxford : Blackwell, 1991.

BUCHANAN and TULLOCK, G., *The Calculus of Consent: Logical Foundations of Constitutional Democracy*, University of Michigan Press, 1962.

COHEN, G.A., *Karl Marx's Theory of History. A Defence,* Princeton: Princeton University Press, 1978.

CONDORCET, Marie-Jean-Antoine de Caritat, marquis de, *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Paris, 1785.

COX, GW, *Making Votes Count : Strategic Interaction in the World's Electoral System,* New York: Cambridge University Press, 1997.

CROCKER, D. A., *Ethics of Global Development: Agency, Capability and Deliberative Democracy*, Cambridge: Cambridge University Press, 2008.

DARWALL, Stephen, *Welfare and Rational Care,* Princeton University Press, 2004.

DAVIDSON, Donald, *Essays on Actions and Events*, Oxford: Clarendon Press, 2001.

DAVIS, John, *Individuals and Identity in Economics*, Cambridge University Press, 2010.

DE CUES, Nicolas, *De concordantia catholica* , 1434.

DEWEY, John, *The Future of Liberalism,* in BOYDSTON, Jo Ann (ed.), *The Later Works, 1882-1898*, vol. 11, Carbondale (Illinois): Southern Illinois University Press, 1981-1990.

DIDEROT, *Le Rêve de d'Alembert, Entretien entre d'Alembert et Diderot* et *Suite de l'entretien*, 1830. Edition critique d'après le manuscrit autographe de Diderot, thèse complémentaire pour le doctorat ès lettres présentée à la Faculté des lettres de l'Université de Paris, par Paul VERNIERE, Paris, Didier, 1951.

DODGSON, C. L., *A Method on taking votes on more than two issues,* in P. Black, 1958.

DWORKIN, Ronald, *Sovereign Virtue,* Cambridge MA: Harvard University Press, 2000.

EDGEWORTH, F. Y., *Mathematical Psychics,* London: 1881.

ELSTER, Jon, *Rational Choice*, Oxford: Blackwell, 1986.

ELSTER, Jon (ed.), *The Multiple Self,* Cambridge University Press, 1986.

ELSTER, Jon and HYLLAND, Aanund (eds), *Foundations of social choice theory*, Cambridge University Press, 1986.

ELSTER, Jon, et ROEMER, John (éditeurs), *Interpersonal Comparisons of Utility*, 1991.

ENGEL Pascal, *Introduction à la philosophie de l'esprit*, Paris, La Découverte, 1994.

FARQUHARSON, R., *Theory of voting*, New Haven : Yale University Press, 1969.

FISCHBACH, Franck, *Manifeste pour une philosophie sociale,* Paris : La Découverte, 2009.

FISCHER, J. M., *The Metaphysics of Free Will,* Wiley-Blackwell, 1995.

FLEURBAEY, Marc, *Théories économiques de la justice*, Paris, Economica, 1996.

FLEURBAY, Marc, *Fairness, Responsibility and Welfare,* Oxford University Press, 2008.

FLEURBAEY, Marc et MANIQUET, François, *A Theory of Fairness and Social Welfare*, Cambridge University Press, 2011.

FOUCAULT, Michel, *Surveiller et punir,* Paris : Gallimard, 1975.

FREGE, Gottlob, *Begriffsschrift*, 1879. Trad. Fr. BESSON, Corine, *Idéographie,* Paris : Vrin, 1999.

FREGE, Gottlob, *Recherches logiques,* trad. Claude Imbert, *in* FREGE, *Ecrits logiques et philosophiques*, Paris, Seuil, 1971.

GAERTNER, Wulf, *A Primer in Social Choice Theory*, Oxford : Oxford University Press, 2006.

GAUTHIER, David, *Morals by Agreement*, Oxford, Clarendon Press, 1986.*Morale et contrat, Recherche sur les fondements de la morale*, Traduction de l'anglais (USA) et introduction par Serge Champeau, Mardaga, Liège, 2000.

GOSSERIES, Axel, *Penser le justice entre les générations : de l'affaire Perruche à la réforme des retraites*, Paris : Aubier, 2004.

GOSSERIES, Axel and MEYER, Lukas, *Intergenerational Justice,* Oxford University Press, 2009, pp 301-322.

GOTOH, Reiko et DUMOUCHEL, Paul (eds.), *Against Injustice : The New Economics of Amartya Sen*, Cambridge University Press, 2010.

GRANGER, Gilles-Gaston, *La mathématique sociale du marquis de Condorcet,* Paris, PUF, 1956.

GRIFFIN, J.P., *Well-Being*, Oxford: Clarendon Press, 1986.

GRÜNE-YANOFF, Till, et HANSSON, Sven Ove, *Preference Change. Approaches from philosophy, economics and psychology,* Springer, 2009.

HABERMAS, Jürgen, *Strukturwandel der Öffentlichkeit*, 1963. *L'espace public: archéologie de la publicité comme dimension constitutive de la société bourgeoise;* traduit de l'allemand par M.B. de Launay. -Paris: Payot, 1978 - rééd.1986.

HABERMAS, Jürgen, *Faktizität und Geltung : Beiträge zur Diskurstheorie des Rechts und des Demokkratischen Rechtsstaats,* Frankfurt : Suhrkamp Verlag, 1992. Trad. Française ROCHLITZ, rainer et BOUCHINDHOMME, Christian, *Droit et démocratie : Entre faits et normes,* Paris : Gallimard, 1997.

HALEVY, Elie, *La Formation du radicalisme philosophique,* Paris, Félix Alcan, 1901-1904, réédité avec une postface de P. Mongin, Paris, PUF, 1996.

HARAWAY, Donna, *Manifeste cyborg et autres essais. Sciences - Fictions - Féminismes*, Anthologie établie par Laurence Allard, Delphine Gardey et Nathalie Magnan. Éditions Exils, 2007.

HARE, R. M., *Essays on Philosophical Method,* London: McMillan, 1971.

HARE, R. M., *Moral Thinking: its levels, methods and point*, Oxford: Oxford University Press, 1982.

HARSANYI, John, *Essays in Ethics, Social Behaviour and Scientific Explanation*, Dordreicht, Reidel, 1976.

HAUSMAN, Daniel, *Preference, Value, Choice and Welfare,* Cambridge: Cambridge University Press, 2012.

HELVETIUS, *De l'Esprit,* Paris [Liège] : Durand [Bassompierre], 1759.

HIRSCHMAN, Albert O., *The Passions and the Interests,* Princeton University Press, 1997.

HOBBES, Thomas, *Leviathan.* Trad. G. MAIRET, Paris : Gallimard, 2000.

HUME, David, *A Treatise of Human Nature*, 1739. Rééd par L.A. Selby-Bigge, Oxford : Oxford University Press, 2000.

HUTCHESON, Francis, *An Inquiry into the Original of Our Ideas of Beauty and Virtue,* London , 1725.

JOUAN, Marlène et LAUGIER, Sandra (dir.), *Comment penser l'autonomie ? Entre compétences et dépendances,* Paris : Presses Universitaires de France, 2009.

KANT, Immanuel, *Logique.* Trad. L. Guillermit, Paris: Vrin, 1989.

KANT, Immanuel, *Métaphysique des Moeurs.* Trad. A. RENAUT, Paris : Garnier-Flammarion, 1994.

KOLM, Serge-Christophe, *Modern Theories of Justice,* Cambridge (Massachussetts), The M.I.T. Press, 1995.

KYMLICKA, *Contemporary political philosophy: an Introduction,* Oxford University Press, 1990, 1999 ; traduction française Marc Saint-Upéry, *Les théories de la justice: une introduction,* Paris: La Découverte, 1999.

LAVAL, Christian, *Jeremy Bentham : le Pouvoir des Fictions,* Paris : Presses Universitaires de France, collection « Philosophies », 1994.

LAVAL, Christian, *L'homme économique. Essai sur les racines du libéralisme,* Paris : Gallimard, 2007.

LEVI, Isaac, *Hard Choices. Decision Making under unresolved conflict.* Cambridge: Cambridge University Press, 1986.

LIST, Christian et PETTIT, Philip, *Group Agency : the possibility, design and status of corporate agents,* Oxford University Press, 2011.

LUDWIG, P. et PRADEU, T. (éds.), *L'individu,* Paris : Vrin, 2008.

LULLE, Raymond, *Blanquerna* (c. 1283). Trad. Fr P. GIFREU, Monaco: éditions du Rocher, 2007.

LULLE, Raymond, *De Arte Eleccionis* (1299).

McLEAN Ian and URKEN, Arnold B.(eds), *Classics of Social Choice,* University of Michigan Press, 1995.

MARX, Karl, *Manuscrits économico-philosophiques de 1844*, présentation, traduction et notes par F. FISCHBACH, Paris, Vrin, collection « Textes & Commentaires », 2007.

MARX, Karl, *Contribution à la critique de l'économie politique,* Editions sociales, 1976.

MARX, Karl, *Le Capital. Livre I,* Trad. Sous la direction de J.-P. LEFEBVRE, Paris : PUF « Quadrige », 1993.

MARX, Karl, *Le Capital. Livres II et III,* trad M. RUBEL, Paris : Gallimard 1968.

MARX, Karl et ENGELS, F., *L'Idéologie Allemande,* Trad. De l'allemand par CARTELLE, Renée et BADIA, Gilbert, Paris : Editions Sociales, 1971.

MILL, J. S., *On Liberty,* London, 1859. Traduction française L. Lenglet et Dupont White, *De la Liberté*, Paris: Gallimard, 1990.

MILL, J.S., *Utilitarianism*, London, 1861.S. Traduction C. AUDARD et P. THIERRY, *L'Utilitarisme,* Paris: Presses Universitaires de France, 1998.

MILL, J.S., *Considerations on Representative Government,* London, 1861. Trad. P.SAVIDAN, *Considérations sur le gouvernement représentatif,* Paris: Gallimard, 2009.

MILL, J.S., *The Subjection of Women,* London, 1869.

MIROWSKI, Philip, *Machine Dreams: Economics become a Cyborg Science,* Cambridge: Cambridge University Press, 2002.

MORISHIMA Mishio, *Why has Japan "Succeeded"? Western Technology and Japanese Ethos,* Cambridge University Press, 1982.

MOULIN, Hervé, *Axioms of Cooperative Decision-Making,* Cambridge, C.U.P., 1988.

NOZICK, Robert, *Anarchy, State and Utopia*, New York, Basic Books, 1977; traduction française par Evelyne d'Auzac de Lamartine, *Anarchie, Etat et Utopie*, Paris, Presse Universitaires de France, 1988, 2003, 2008.

NUSSBAUM, Martha, *Women and human Development: the capabilities Approach*, Cambridge: Cambridge University Press, 1999. *Femmes et développement humain,* traduit de l'anglais (USA) par Camille Chaplain, Paris, Edition Des Femmes-Antoinette Fouque, 2008.

NUSSBAUM, Martha, *Frontiers of justice : Disabilities, Nationality, Species membership,* The Belknap Press of Harvard Universitu Press, 2006.

NUSSBAUM, Martha, *Not for profit: Why Democracy needs the humanities*, Princeton University Press, 2010. Traduction française Solange Chavel, *Les émotions démocratiques,* Paris: Flammarion, 2011.

NUSSBAUM, Martha, *Creating Capabilities: the Human Development Approach,* Harvard University Press, 2011.

NUSSBAUM, Martha, et SEN, Amartya, *The Quality of Life*, New York, Oxford University Press, 1993.

OLSARETTI, Serena (ed.), *Preferences and Well-Being,* Cambridge University Press, 2006.

PARETO, Vilfredo, *Manuel d'Economie politique*, 1907, réédité in *Œuvres complètes*, VII, Genève, Droz, 1966.

PARFIT, Derek, *Reasons and Persons*, Oxford : Clarendon Press, 1984.

PASCAL, Blaise, *Pensées*, 1669. Rééd. in *Oeuvres completes,* Paris: Seuil, 1963.

PETTIT, Philip (ed.), *Consequentialism,* Aldershot: Darthmouth, 1993.

PETTIT, Philip, *Republicanism: A Theory of Freedom and Government,* Oxford: Clarendon Press, 1997. Trad. Fr. Patrick Savidan et J.-F. Spitz, *Républicanisme*, Paris: Gallimard, 2004.

PICAVET, Emmanuel, *Choix rationnel et vie publique*, Paris, P.U.F., 1998.

PIGOU, Alfred, *The Economics of Welfare*, Londres, MacMillan, 1920.

PLINE LE JEUNE, *Lettres,* Trad. DE SACY et PIERROT, Paris : Garnier, 1920.

POGGE, Thomas, *Realizing Rawls,* Ithaca: Cornell University Press, 1989.

PUFENDORF, *De Jure Naturae et Gentium* (1672). Trad. BARBEYRAC, Jean, *Droit de la Nature et des Gens,* Amsterdam : Henri Schelte, 1771.

PUTNAM, Hilary, *Reason, Truth and History,* Cambridge: Cambridge University Press, 1981. Traduction GERSCHENFELD, Abel, *Raison, Vérité et Histoire,* Paris: Les Editions de Minuit, 1984.

PUTNAM, Hilary, *The Collapse of the Fact/Value dichotomy and other essays,* Harvard University Press, 2002. Trad. M. CAVERIBERE et M. COMETTI, *Fait/Valeur: la fin d'un dogme et autres essais,* Paris : Editions de l'Eclat, 2004.

RAWLS, John, *A Theory of Justice*, The Belknap Press of Harvard University Press, 1971; traduction française par Catherine Audard sur un texte révisé par l'auteur en 1975: *Théorie de la Justice*, Paris, Editions du Seuil, 1987, 1997. Préface de la traduction française par John Rawls.

RAWLS, John, *Political Liberalism*, New York, Columbia University Press, 1993; traduction française par Catherine Audard: *Libéralisme politique*, Paris, Presses Universitaires de France, « Quadrige », 1995, 2007.

RAWLS, John, *Justice as Fairness : A Restatement,* Cambridge (MA: The Belknap University Press, 2001. Trad. B. GUILLARME, *La Justice comme Equité,* Paris: La Découverte, 2003.

RILEY, Jonathan, *Liberal Utilitarianism : Social Choice Theory and J. S. Mill's Philosophy,* Cambridge University Press, 1988.

ROEMER, John, *Theories of Distributive Justice*, Cambridge (Massachussets): Harvard University Press, 1995.

ROUSSEAU, Jean-Jacques, *Du Contrat Social* (1765). Paris : Garnier-Flammarion, 2006.

SAINT-SERNIN, Bernard, *Les Mathématiques de la Décision,* Paris, PUF, 1973.

SAMUELSON, P.A., *Foundations of Economic Analysis,* Cambridge, MA: Harvard University Press, 1947.

SCHEFFLER, Samuel, *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions,* Oxford University Press, 1982.

SEN, Amartya, *Collective Choice and Social Welfare,* San Francisco, Holden day, 1970.

SEN, Amartya, *Poverty and Famines*, Oxford University Press, 1981.

SEN, Amartya, *Choice, Welfare and Measurements,* Cambridge MA, The MIT Press, 1982.

SEN, Amartya, *Resources, Values and Development,* Cambridge, MA: Harvard University Press, 1984.

SEN, Amartya , *Commodities and Capabilities,* Amsterdam: North-Holland, 1985.

SEN, Amartya, *On Ethics and Economics*, Blackwell Publishers, Oxford, 1991; traduction française par Sophie Marnat : *Ethique et Economie*, Paris, Presses Universitaires de France, Quadrige, 1993, 2008.

SEN, Amartya, *Inequality Reexamined*, Oxford University Press, 1992 ; traduction française par Paul Chemla : *Repenser l'inégalité*, Paris, Editions du seuil, 2000.

SEN, Amartya, *Development as Freedom*, New York, Alfred A. Knopf Inc., 1999; traduction française par Michel Bessières: *Un nouveau modèle économique. Développement, justice, liberté.*, Paris, Editions Odile Jacob, 2000, 2003.

SEN, Amartya, *Rationality and Freedom*, The Belknap Press of Harvard University Press, 2002; traduction française par Marie-Pascale d'Iribane-Jaawane : *Rationalité et liberté en économie*, Paris, Editions Odile Jacob, 2005.

SEN, Amartya, *Identity and Violence : The Illusion of Destiny,* New York, London: WW Norton & Company, 2006. Traduction fr. KLEIMAN-LAFON, Sylvie, *Identité et Violence,* Paris: Odile Jacob, 2007.

SEN, Amartya K., *The Idea of Justice*, London: Allen Lane, 2009. Traduction française Paul Chemla et Eloi Laurent, *L'Idée de Justice,* Paris : Flammarion, 2009.

SEN, A., MUELLBAUER, J., KANBUR, R., HART, K., et WILLIAMS, B., *The Standard of Living,* Cambridge: Cambridge University Press, 1987.

SEN, Amartya, et WILLIAMS, Bernard, *Utilitarianism and Beyond*, Cambridge University Press, 1962.

SHAFTESBURY, Anthony Ashley Cooper, 3rd Earl of Shaftesbury, *Characteristics of Men, Manners, Opinions, Times*, édition par Lawrence E. Klein, Cambridge: Cambridge University Press, 1999.

SIGDWICK, Henry, *Methods of Ethics,* London: McMillan & Co, 1874, 1907[7].

SINGER, Peter, *Practical Ethics,* Cambridge University Press, 1979, 2011[3].

SMITH, Adam, *The Theory of Moral Sentiments,* 1759, rééd. Oxford : Clarendon Press, 1976. Trad. Fr. BIZIOU, M., GAUTIER, C., PRADEAU, J.-F., *Théorie des sentiments moraux*, Paris : P.U.F., 2011.

SPINOZA, Baruch, *Ethique*. Traduction Charles Appuhn, Paris : Vrin, 1977[2].

STEWART, Frances, *Basic Needs in Developing Countries*, Baltimore, Md: John Hopkins University Press, 1985.

STREETEN Paul, avec BURKI S. J., UL HAQ M., HICKS N., STEWART F., *First Things First*, Oxford: Oxford University Press for the World Bank, 1981.

THIESSE, Anne-Marie, *La création des identités nationales,* Paris : Seuil, 1999.

TIERCELIN, Claudine, *Hilary Putnam, l'héritage pragmatiste*, Paris: Presses Universitaires de France, "Philosophies", 2002

TOCQUEVILLE, Alexis de, *De la démocratie en Amérique* (1830)*,* Paris : Garnier-Flammarion, 1981.

VALIER, Jacques, *Brève histoire de la pensée économique*, Paris : Flammarion, 2009.

VAN PARIJS, Philippe, *Qu'est-ce qu'une société juste ?*, Paris, Le Seuil, 1991.

VAN PARIJS, Philippe, *Real Freedom for All. What (if anything) can justify capitalism*?, Oxford University Press, 1995.

WALSH, Vivian C., *Scarcity and Evil,* Prentice-Hall, 1961.

WEBER, Max, *Die Protestantische Ethik und Der Geïst Des Kapitalismus*, 1920[2], trad. Fr. J. Chavy, Paris, Plon, 1964.

WILLIAMS, Bernard, *Moral Luck,* Cambridge University Press, 1981. Traduction française *La Fortune morale: moralité et autres essais,* Paris : PUF, 1994.

WITTGENSTEIN, Ludwig, *Philosophische Untersuchungen*, Blackwell Publishers, 1953. Traduction française par F. Dastur, M. Elie, J.-L. Gautero, D. Janicaud, E. Rigal, *Recherches philosophiques,* Paris : Gallimard, 2004.

Ouvrages de référence

AUDARD, Catherine, *Anthologie historique et critique de l'utilitarisme,* Paris : PUF, 1999.

GOODIN, Robert E., PETTIT, Philip and POGGE, Thomas (eds), *A Companion to Contemporary Political Philosophy (Second Edition),* Oxford: Wiley-Blackwell, 2012.

MESURE, Sylvie, et SAVIDAN, Patrick (directeurs), *Dictionnaire des Sciences Humaines,* Paris, PUF, 2006.

The Cambridge Companion to Rawls, Cambridge University Press, 2003.

Articles de revues

AGARWAL, Bina, « 'Bargaining' and Gender Relations : Within and Beyond the Household », *Feminist Economics* 3, no. 1, 1997.

ALEXANDER, J. M., "Capability Egalitarianism and Moral Selfhood", *Ethical Perspectives,* 10, 1, 2003, p 3.

ANAND, P. et VAN HEES, M., "Capabilities and Achievements: an empirical study", *Journal of Socio-Economics,* 35, pp 268-284.

ANDERSON, Elizabeth, "What is the Point of Equality?" *Ethics*, 109(2), 1999, pp 287-337.

ANDERSON, Elizabeth, "Unstrapping the straitjacket of 'preference': a comment on Amartya Sen's contributions to philosophy and economics." *Economics and Philosophy*, 17, 2001, pp 21-38.

ANDERSON, Elizabeth, "Sen, Ethics and Democracy", *Feminist Economics,* 0(0), 2003.

ARNESON, Richard, « Equality and equal opportunity for welfare », *Philosophical Studies*, Volume 56, Issue 1, 1989, pp 77-93.

ARNESON, Richard, « Liberalism, Distributive Subjectivism, and Equal Opportunity for Welfare », *Philosophy and Public Affairs*, Vol. 19, No. 2, 1990, pp 158-194.

BABER, H. E., « Worlds, Capabilities and Well-Being", *Ethical Theory and Moral Practice,* Vol 13 (4), 2010, p 377-392.

BARRETT, Richard, PETRON-BRUNEL, Anne, SALLES, Maurice, « A new approach to rights in social choice theory which incorpores utilitarianism », *Social Choice and Welfare* 22, 2004, p 17-28.

BAUJARD Antoinette, « Conceptions of freedom and ranking opportunity sets. A typology. », *Homo Oeconomicus* 24, 2007, p 1-24.

BAUJARD, Antoinette, "A return to Bentham's felicific calculus: from moral welfarism to technical non-welfarism", *European Journal of the History of Economic Thought,* vol 16, no 3, 2009, pp 431-453.

BAUJARD, Antoinette, "Collective interest versus individual interest in Bentham's felicific calculus. Questioning welfarism and fairness."*European Journal of the History of Economic Thought,* vol 17, no 4, 2010, pp 607-634.

BERGSON, Abram, "A Reformulation of certain aspects of Welfare Economics", *Quarterly Journal of Economics*, 52, 1938, pp 310-334.

BERGSON, Abram, "On the Concept of social Welfare", T*he Quarterly Journal of Economics*, Vol. 68, No. 2, 1954, pp. 233-252.

BLACK, Duncan, "On The Rationale of Group Decision Making", *The Journal of Political Economy,* 56, 1948, pp 23-34.

BOSSERT, Walter, SUZUMURA, Kotaro, « Multi-profile Intergenerational Social Choice », *Social Choice and Welfare,* vol. 37, 3, sept. 2011, pp 493-509.

BOVENS, Luc, « Sour Grapes and Character Planning », *The Journal of Philosophy,* Vol. 89, No. 2, feb. 1992, pp 57-78.

BROGAN, Albert, « The fondamental value universal », *Journal of Philosophy, Psychology and Scientific Method*, 16, 1919, pp 96-104.

BRUCKNER, Donald W., "In defense of adaptive preferences", *Philos Stud,* 142, 2007, pp 307-324.

BUZAGLO, Jorge, « Capabilities : from Spinoza to Sen and Beyond », *Post-Austistic Economics Review*, 20-21, 2003.

CARTER, Ian, "The concept of freedom in the work of Amartya Sen: An alternative analysis consistent with freedom's independent value", *Notizie di Politeia* 43(4), 1996.

CLAASSEN, Ruther, DÜWELL, Marcus, "The Foundations of Capability Theory: Comparing Nussbaum and Gewirth", *Ethical Theory and Moral Practice,* 2012.

COHEN, G.A., « Karl Marx and the withering away of social sciences » *Philosophy and Public Affairs,* 1, no 2, hiver 1972.

COHEN, G.A., "On the currency of egalitarian justice", *Ethics,* Vol. 99, No. 4, 1989, pp 906-944.

CROCKER, D.A., « Functioning and Capability : the Foundations of Sen's and Nussbaum's development Ethics", *Political Theory,* Vol. 20, No. 4, Nov. 1992.

DRYZEK, John S., LIST, Christian, « Social Choice Theory and Deliberative Democracy : A Reconciliation », *British Journal of Political Science,* 2003.

EDGEWORTH, F.Y., « The Pure Theory of Taxation », in *Papers Relating to Political Economy,* London, 1925.

FANSTEN, Michel, « L'échec électoral est une science exacte », *Mathématiques et sciences humaines*, 185, 2009.

FEREJOHN, John and PAGE, Talbot, « On the foundations of intertemporal choice », *American Journal of Agricultural Economics,* 60 (2), 1978, pp 269-275.

FLEURBAEY, Marc, « Choix social : une difficulté et de multiples possibilités », *Revue économique,* Vol. 51, No. 5, sept. 2000, pp 1215-1232.

FLEURBAEY, Marc, « On the informational basis of social choice », *Social Choice and Welfare*, 21, 2003, pp 347-384.

FLEURBAEY, Marc, « Capabilities, functionings and refined functionings », *Journal of Human Development,* 7 (3), 2006, p 299-310.

FLEURBAEY, Marc, "Social Choice and Just Institutions", *Economics and Philosophy,* 23, 2007, pp 15-43.

FLEURBAEY, Marc, "The Importance of what People Care About", *Politics, Philosophy and Economics,* 11 (4), 2012, pp 415-447.

FRANKFURT, Harry, "Alternate Possibilities and Moral responsibility", *The Journal of Philosophy,* Vol. 66, No. 23, 1969, pp 829-839.

FRANKFURT, Harry, "Freedom of the Will and the Concept of a Person", *The Journal of Philosophy,* Vol. 68, No. 1, 1971, pp 5-20.

FRIED, Greg, "Teaching Arrow's Theorem", *Teaching Philosophy* 33(2), 2010, pp 173-186.

GAERTNER W., PATTAINAK P. et SUZUMURA K., "Individual Rights Revisited", *Economica,* 59, 1992, pp 161-177.

GIBBARD, Allan, "Manipulation of voting schemes: a general result", *Econometrica* 41 (4), 1973, pp 587–601.

GIBBARD, A. « A Pareto-Consistant Libertarian Claim », *Journal of Economic Theory* 7, 1974, 388-410.

GIBBARD, A., « Preference Strength and Two Kinds of Ordinalism », *Philosophia,* 7 (2), 1978, pp 255-264.

GILARDONE, Muriel, "Amartya Sen sans prisme", *Cahiers d'Economie Politique*, 1, 2010, n° 58.

GOLDFINGER, Johnny, « The Value of Social Choice for Normative Political Theory", *The Good Society,* 11, 2, 2002, pp 33-37

GRIFFIN, J.P., « Are there incommensurable values ? », *Philosophy and Public Affairs 7,* No. 1, 1977.

HARSANYI, John, "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility", *Journal of Political Economy* 63, 1955, pp 309-321.

HAUSMAN, Daniel, « Rational Choice and Social Theory: a Comment", *The Journal of Philosophy*, Vol. 92, No. 2, 1995, pp 96-102

HEATHWOOD, « The Problem of Defective Desires », *Australasian Journal of Philosophy,* 83, 2005, pp 487-504.

HUDON, François, « La liberté dépend-elle des préférences individuelles ? », *Raisons politiques,* 3, n° 43, 2011, pp 35-58.

IGERSHEIM, Herrade, « Du paradoxe libéral-parétien à un concept de métaclassement des préférences », *Recherches économiques de Louvain* vol. 73, 2/2007, p. 173-192.

JAGGAR, Alison, « Reasoning about Well-Being : Martha Nussbaum's methods of Justifying the Capabilities", *The Journal of Political Philosophy,* volume 14, Issue 3, 2006, pp 301-322.

LIST, Christian et PETTIT, Philip, « Aggregating sets of judgments : two impossibility results compared », *Synthese*, 140, 2004, pp 207-235

LITTLE, L.M.D., « Social Choice and Individual Values », *Journal of Political Economy* 60, 1952, p 422-432

McLEAN, Ian and LONDON, J., "The Borda and Condorcet Principles. Three Medieval Applications", *Social Choice and Welfare,* 7, 1990, pp 99-108.

MARTIN, Mathieu et MERLIN, Vincent, « Les apports de la théorie du choix social pour l'analyse de la démocratie », *Cahiers d'Economie Politique* n° 47, 2004/2.

MONGIN, Philippe, « Les préférences révélées et la formation de la théorie du consommateur ». In: *Revue économique*. Volume 51, n°5, 2000. pp. 1125-1152

MONGIN, Philippe, « Value judgments and Value neutrality in Economics », *Economica*, 73, 2006, pp 257-286

MONGIN, Philippe, "Une source méconnue de la théorie de l'agrégation des jugements", *Revue économique,* vol. 63, 2012/4, pp 645-657.

MONGIN, Philippe et DIETRICH, Franz, « Un bilan interprétatif de la théorie de l'agrégation logique », *Revue d'économie politique*, 120 (6), 2010, pp 929-972.

NÄSSTROM, Sofia, "The Challenge of the All-Affected Principle", *Political Studies,* vol. 59, 2011, pp 116-134.

NUSSBAUM, Martha C., "Nature, Function, and Capability", *Oxford Studies in Ancient Philosophy*, suppl. vol. 1, 1988, pp. 145–84

NUSSBAUM, Martha, « Adaptive preferences and women's options », *Economics and Philosophy*, vol. 17, Issue 01, 2001, pp 67-88.

OLSARETTI, Serena, "Endorsement and Freedom in Amartya Sen's Capability Approach", *Economics and Philosophy,* no 21, issue 01, 2005, pp 89-108.

PATTANAIK, P. K., and XU, Y., "On Ranking Opportunity Sets in Terms of Freedom of Choice", *Recherches Economiques de Louvain*, 56, 1990, pp 383-390.

PETER, Fabienne, "Gender and the Foundations of Social Choice: The Role of Situated Agency", *Feminist Economics* 9 (2-3), 2003, pp 13-32.

PETTIT, Philip, "Capability and Freedom: A Defence of Sen", *Economics and Philosophy* 17, 2001, pp 1-20.

PILDES, Richard et ANDERSON, Elizabeth, "Slinging Arrows at Democracy: Social Choice Theory, Value Pluralism and Democratic Politics",*Columbia Law Review*, Vol. 90, No. 8, 1990, pp. 2121-2214

POGGE, Thomas, "Can the Capability Approach be Justified?", *Philosophical Topics,* 30 (2), 2002, pp 167-228.

QIZILBASH, Mozaffar, "Capability, Happiness and Adaptation in A. Sen and J. S. Mill", *Utilitas,* Vol. 18, Issue 1, 2006.

QIZILBASH, Mozaffar, "Social choice and individual capabilities", *Politics Philosophy Economics* 6, 2007, pp 169-192.

READER, Soran, "Does a Basic Needs Approach need Capabilities", *The Journal of Political Philosophy,* Vol. 14, number 3, 2006, pp 337-350.

RILEY, Jonathan, « Utilitarian Ethics and Democratic Government », *Ethics,* Vol. 100, No. 2, 1990, pp 335-348.

ROBEYNS, I., "Sen's Capability Approach and Gender Inequality: Selecting Relevant Capabilities," *Feminist Economics*, 9(2/3), 2003, pp 61–92.

ROBEYNS, Ingrid, "The Capability Approach in Practice," *Journal of Political Philosophy*, 14(3), 2006, pp 351–76.

ROBEYNS, Ingrid, "The Capability Approach: A Theoretical Survey," *Journal of Human Development*, 6(1), 2005, pp 93–114.

SAMUELSON, Paul A., « A Note on the Pure Theory of Consumer's Behaviour: An Addendum", *Economica,* New Series, Vol. 5, No 19, 1938, pp 353-354.

SATTERTHWAITE, Mark, "Strategy-proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions", *Journal of Economic Theory* 10, 1975, pp 187–217.

SATZ, D. et FEREJOHN, J., « Rational Choice and Social Theory", *The Journal of Philosophy*, Vol. 91, No. 2, 1991, pp 71-87

SCANLON, Thomas, « Preference and Urgency », *Journal of Philosophy* 72 (19), 1975, pp 655-669.

SCHOKKAERT, Erik, "Capabilities and Satisfaction with Life", *Journal of Human Development and Capabilities,* vol 8(3), 2007, pp 415-430.

SEN, Amartya, "The Nature and Classes of Prescriptive Judgments", *Philosophical Quarterly,* 17/66, 1967, pp 46-62.

SEN, Amartya, "The Impossibility of a Paretian Liberal", *Journal of Political Economy,* vol. 78, no 1, 1970, pp 152-157.

SEN, Amartya, "Choice Functions and Revealed Preference", *Econometrica*, 41, 1973.

SEN, Amartya, "Utilitarianism and Welfarism", *The Journal of Philosophy,* Vol. 76, no 9, 1979, pp 463-489.

SEN, Amartya, "Description as a choice", *Oxford Economic Papers,* New Series, Vol. 32, No 3, 1980, pp 353-369.

SEN, Amartya, "Plural Utility", *Proceedings of the Aristotelian Society*, New Series, Vol. 81, 1980-1981, pp 193-215.

SEN, Amartya, « Social Choice Theory : A Reexamination", *Econometrica*, 45, Janvier 1977, pp 53-89. Rééd. *In Choice, Welfare and Measurement*, Oxford: Basic Blackwell Publishers, 1982.

SEN, Amartya, "Equality of What?", in S. McMURRIN (éd.), *Tanner Lectures on Human Values,* Cambridge: Cambridge University Press, 1980.

SEN, Amartya, « Positional Objectivity », *Philosophy & Public Affairs*, Vol. 22, No. 2, 1993, pp. 126-145.

SEN, Amartya, "Rationality and Social Choice", *American Economic Review,* vol. 85(1), 1995, pp 1-24.

SEN, Amartya, "How to Judge Voting Schemes", *Journal of Economic Perspectives,* vol. 9(1), 1995, pp 91-98.

SEN, Amartya, « The Possibility of Social Choice », *The American Economic Review,* Vol. 89, No 3, 1999, pp 349-378.

SEN, Amartya, "Democracy as a Universal Value", *Journal of Democracy* 10.3., 1999, pp 3-17.

SEN, Amartya, "Consequential evaluation and practical reason", *The Journal of Philosophy,* vol. 97, no 9, 2000, pp 477-502

SEN, Amartya, "Sraffa, Wittgenstein and Gramsci", *Journal of Economic Literature,* vol. 41(4), 2003, pp 1240-1255.

SEN, Amartya K., "Incompleteness and Reasoned Choice", *Synthese,* 140, 2004.

SEN, Amartya, "Capabilities, Lists and Public Reason: Continuing the Conversation", *Feminist Economics,* vol. 10(3), 2004, pp 77-80.

SEN, Amartya, "Why Exactly is Commitment important for Rationality?", *Economics and Philosophy* 21(1), 2005, pp 5-14.

SEN, Amartya, "What do we want from a Theory of Justice?", *Journal of Philosophy* 103(5), 2006, pp 215-238.

SEN, Amartya, "Reason, Freedom and Well-Being", *Utilitas,* vol. 18 (01), 2006, pp 80-96.

SEN, Amartya, "The Reach of Social Choice Theory", *Social Choice and Welfare,* 39, 2012, pp 259-272.

SUPPES, Patrick, "The Pre-History of Kennneth Arrow's social choice and individual values", *Social Choice and Welfare,* vol. 25, issue 2-3, 2005, pp 319-326.

TEMKIN, Larry, "Intransitivity and the Mere Addition Paradox", *Philosophy and Public Affairs,* Vol. 7, No 2, 2007, pp 138-187.

VALENTINI, Laura, "A Paradigm Shift in Theorizing About Justice? A Critique of Sen.", *Economics and Philosophy,* Vol. 27, Issue 03, 2011, pp 297-315.

VALLENTYNE, Peter, "Capabilities VS Opportunities for Well-Being", *Journal of Political Philosophy*, Volume 13, Issue 3, 2005, pages 359–371.

WALBY, Sylvia, « Sen and the Measurement of Justice and Capabilities: A Problem in Theory and Practice", *Theory, Culture and Society,* 29, no 1, 2012, pp 99-118.

WALSH, Vivian, "Sen after Putnam", *Review of Political Economy,* 15, 2003, p 331.

YAARI et BAR-HILLEL, "On Dividing Justly", *Social Choice and Welfare,* 1, 1984, pp 1-24.

Chapitres de livres

ANDERSON, E., "Justifying the capabilities approach to justice", in Brighouse and Robeyns (eds.), 2010, pp 81-100.

ARNESON, Richard, "Desire Formation and Human Good", in OLSARETTI, Serena (ed.), *Preferences and Well-Being,* Cambridge University Press, 2006.

ASPREMONT, Claude, "Economie du bien-être et utilitarisme », in L.A. Gérard-Varet et J.C. Passeron (éds), *Le modèle et l'enquête. Les usages du principe de rationalité dans les sciences sociales,* Paris, Editions de la Maison des Sciences de l'Homme, 1995.

BRÄNNMARK, Johan, "Leading a Life of One's Own: On Well-Being and Narrative Autonomy", in OLSARETTI, Serena (ed.), *Preferences and Well-Being,* Cambridge University Press, 2006.

BROOME, John, "Reasoning with preferences?", in OLSARETTI (ed.), *Preferences and Well-Being,* Cambridge Univesrity Press, 2006.

BYKVIST, Krister, « Preference formation and intergenerational justice », in GOSSERIES, Axel and MEYER, Lukas, *Intergenerational Justice,* Oxford University Press, 2009, pp 301-322.

COHEN, G.A, "Equality of What? On Welfare, Goods and Capabilities", *in* Nussbaum, Martha and Sen, Amartya (eds), *The Quality of Life,* Oxford: Oxford University Press, 1993.

DAVIDSON, Donald, "Psychology as Philosophy", *in* BROWN S. C. (ed.), *Philosophy of Psychology,* The Macmillan Press and Barnes, Noble, Inc, 1974.

ELSTER, Jon, "Sour grapes – utilitarianism and the genesis of wants", *in* Sen, Amartya Kumar and Williams, Bernard (eds), *Utilitarianism and beyond,* Cambridge: University Press, 1982.

FRIEDMAN, Milton, « The Methodology of Positive Economics", in *Essays in Positive Economics*, Chicago University Press, 1953.

GIBBARD, A., « Interpersonal Comparisons: Preferences, Goods and the Intrinsic Reward of a Life", in ELSTER, Jon and HYLLAND, Aanund (eds), *Foundations of social choice theory*, Cambridge University Press, 1986.

GOODIN, Robert E. « Laundering Preferences », in ELSTER, Jon and HYLLAND, Aanund (ed.), *Foundations of Social Choice Theory,* Cambridge University Press, 1986.

HARSANYI, John, "Morality and the Theory of Rational Behaviour", in SEN and WILLIAMS (ed.), *Utilitarianism and Beyond,* Cambridge University Press, 1982.

NUSSBAUM, Martha, "Non-relative Virtues: an Aristotelian Approach", *in* Nussbaum, Martha and Sen, Amartya (eds), *The Quality of Life,* Oxford: Oxford University Press, 1993.

PETTIT, Philip, « Decision Theory and Folk Psychology", *in* BACHARACH, M. et HURLEY S. (eds), *Foundations of Decision Theory: Issues and Advances*, Blackwells, Oxford, 1991, pp. 147-75.

PETTIT, Philip, "Preferences, Deliberation and Satisfaction", in OLSARETTI (ed.), *Preferences and Well-Being,* Cambridge University Press, 2006.

ROBERTS, Kevin, "Social Choice Theory and the Informational Basis Approach", in MORRIS, Christopher (ed.), *Amartya Sen*, Cambridge University Press, 2009, pp 115-137.

ROBEYNS Ingrid,"Justice as fairness and the capability approach", in: Kaushik Basu and RAVI KANBUR(eds.), *Arguments for a Better World. Essays for Amartya Sen's 75th Birthday*, Oxford University Press, 2006, pp. 397–413.

SEN, Amartya, "Choice, Orderings and Morality", in KÖRNER, Stephan (éd.), *Practical Reason,* Oxford: Basil Blackwell, 1974.

VECA, Salvatore, "La Possession", in Collectif, *Quelle philosophie pour le XXIe siècle? L'Organon du nouveau siècle,* Paris : Gallimard/Editions du Centre Pompidou, 2001.

WHELAN, F. G.'Prologue: Democratic Theory and the Boundary Problem', in R. J. Pennock and J.W. Chapman (eds), *Liberal Democracy.* NewYork: NewYork University Press, 1983, pp. 13–47.

WIGGINS, David, "Claims of Need", in *Needs, Values, Truth*, Oxford: Blackwell, 1987, pp 1-57.

Articles de dictionnaire ou d'encyclopédie

ANGNER, Erik, "Subjective Measures of Well-Being: Philosophical perspectives," in KINCAID, Harold and DON ROSS (Eds.) *The Oxford Handbook of Philosophy of Economics*, Oxford: Oxford University Press, 2009, pp. 560-579

ARROW, Kenneth, « Formal Theories of Social Welfare », in WIENER, Philip P. (ed), *Dictionary of the History of Ideas*, Charles Scribner's Sons, 1973, vol. 4, pp 276-284. Reed. ARROW, Kenneth, *Collected Papers,* Cambridge, MA: The Belknap Press of Harvard University Press, 1983.

FLEURBAEY, Marc, et MONGIN, Philippe, "Théorie du choix social et économie normative", in CANTO-SPERBER, Monique (directrice), *Dictionnaire de philosophie morale*, Paris, P. U.F., 1996, p. 243-251.

SALLES, Maurice, articles « agrégation des préférences » et « bien-être et choix social », in BEITONE, Alain, CAZORLA, Antoine, DOLLO, Christine, DRAI, Anne-Mary, *Dictionnaire des Sciences Economiques* , Paris : Armand Colin, 2006.

Thèses

BYKVIST, Krister, *Changing Preferences: A Study in Preferentialism*, Thèse. Université d'Uppsala, 1998.

GILARDONE, Muriel, *Contexte, sens et portée de l'approche par les capabilités de Amartya Kumar Sen*, Thèse de doctorat, Université Lumière Lyon 2, Faculté de sciences économiques et de gestion, 2007.

IGERSHEIM, Herrade, *Liberté et Choix social - Contribution à l'analyse de la liberté en économie normative.* Thèse de doctorat, Université Louis Pasteur, Strasbourg, 2004.

Documents de travail

GIBBARD, A., *Intransitive social indifference and the Arrow dilemma*, mimeo, Département de Philosophie, Université de Michigan, 1969.

GOTOH, Reiko, "A Note on Capability Comparison and Social Evaluation", Paper for the workshop on Multidimensional Poverty, Oxford Poverty and Human development initiative (OPHI), Dept of International Development, University of Oxford, 29 May-1 June 2007.

SALLES, Maurice, « La théorie du choix social : de l'importance des mathématiques », Center for Research in Economics and Management (CREM), University of Rennes 1, University of Caen and CNRS, Economics Working Paper Archive (University of Rennes 1 & University of Caen)no 200617

SCHOKKAERT, Erik, "The Capability Approach", document de travail, Department of Economics, KU Leuven, ces0734, Mars 2007.

Communications de colloques, cours, conférences

ELSTER, Jon, « Les décisions collectives », cours donné au Collège de France, Paris, 15 janvier-7 mai 2009. Retranscription audio : http://www.college-de-france.fr/site/jon-elster/#|q=/site/jon-elster//_audiovideos.jsp| ; Consulté le 20 décembre 2012.

GILARDONE, Muriel et BAUJARD, Antoinette, « Individual Preferences VS Positional Judgments », *Communication dans le cadre du Colloque « Philosophie Economique »,* Sciences Po Lille, 21-22 juin 2012.

HAWI, Rima, « Rawls et l'économie : les leçons des archives de Harvard », *Communication dans le cadre du Colloque « Philosophie Economique »,* Sciences Po Lille, 21-22 juin 2012.

KOLM, Serge-Christophe, "What sense Social Choice?" C.G.P.C., 1993, and International Economic Association Conference on Social Choice, Hernstein, 1994.

Ressources en ligne

CRISP, Roger, « Well-Being », *Stanford Encyclopedia of Philosophy,* dec. 2008, http://plato.stanford.edu/entries/well-being/ ; Consulté le 6 décembre 2012.

DRIVER, Julia, "The History of Utilitarianism", *The Stanford Encyclopedia of Philosophy*, 2009, http://plato.stanford.edu/archives/sum2009/entries/utilitarianism-history/ ; Consulté le 24 août 2012.

FLEURBAEY Marc, "Economics and Economic Justice", *The Stanford Encyclopedia of Philosophy,* 2012, http://plato.stanford.edu/archives/sum2012/entries/economic-justice/ ; Consulté le 2 décembre 2012.

FLEURBAEY, Marc, et TUNGODDEN, Bertil, « The tyranny of aggregation VS the tyranny of non-aggregation in social choices : a real dilemma », *LES Philosophy Papers,* November 2007, http://www2.lse.ac.uk/CPNSS/pdf/DP-withcover_LSE-PhilosophyPapers/PP-02-07/02-07-online.pdf ; Consulté le 3 novembre 2012.

ROBEYNS, Ingrid, "The Capability Approach", *The Stanford Encyclopedia of Philosophy (Summer 2011 Edition)*, Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/sum2011/entries/capability-approach/ ; Consulté le 2 décembre 2012.

ZUPPINGER, Thibaud, "La survenance, entre axiologie et anthropologie philosophique », *Raison Publique,* 27 avril 2010, http://www.raison-publique.fr/article293.html; Consulté le 31 août 2012.

Rapports officiels

BAUJARD, Antoinette et IGERSHEIM, Herrade, "Expérimentation du vote par note et du vote par approbation lors de l'élection présidentielle du 22 avril 2007", *Rapport pour le Centre d'Analyse stratégique,* 2007.

Annexes

Annexe 1 : notation standard en logique mathématique

$\exists$: quantificateur existentiel (« il existe au moins un... »)
$\forall$: quantificateur universel (« pour tout... »)
$\rightarrow$: implication (« si, alors »)
$\leftrightarrow$: équivalence (« si et seulement si »)
$\sim$: négation
$\vee$: alternation (« ou » inclusif)
& : conjonction (« et »)
= : identité (« le même que »)
$\in$: appartenance d'un élément à un ensemble (« appartient à »)
$\subset$: sous-ensemble d'un ensemble (« est contenu dans »)
$\cap$: intersection d'ensembles (« éléments appartenant à chacun de ces ensembles distincts »)
$\cup$: union d'ensembles (« éléments appartenant tous ces ensembles joints »)

Annexe 2 : les relations de préférence et les ensembles de choix[484]

Soit *x R y* représentant une relation binaire entre *x* et *y.* Il s'agit d'une relation de préférence faible. *x R y* signifie « *x* est au moins aussi bonne que *y* », ou encore « *x* est préférée ou indifférente à *y* ». Cette relation binaire est définie sur un ensemble d'éléments *S.*

Cette relation peut satisfaire les propriétés suivantes :

(1) Réflexivité : $\forall x \in S : x\ R\ y$.
(2) Complétude : $\forall x, y \in S : (x \neq y) \rightarrow (x\ R\ y \vee y\ R\ x)$
(3) Transitivité : $\forall x, y, z \in S : (x\ R\ y\ \&\ y\ R\ z) \rightarrow x\ R\ z$.
(4) Anti-symétrie : $\forall x, y \in S : (x\ R\ y\ \&y\ R\ x) \rightarrow x = y$
(5) Asymétrie : $\forall x, y \in S : x\ R\ y \rightarrow \sim(y\ R\ x)$
(6) Symétrie : $\forall x, y \in S : x\ R\ y \rightarrow y\ R\ x$

Les relations satisfaisant certaines propriétés et utilisées plus fréquemment sont nommées selon la terminologie suivante, qui est celle de Sen[485] :

Propriétés de la relation	**Dénomination de Sen**	**Dénominations alternatives**
1 Réflexivité et transitivité	Quasi-ordre	Pré-ordre
2 Réflexivité, transitivité et complétude	Ordre	Pré-ordre complet ; quasi-ordre complet ; ordre faible
3 Réflexivité, transitivité et anti-symétrie	Ordre partiel	Ordre
4 Réflexivité, transitivité, complétude et anti-symétrie	Chaîne	Ordre linéaire ; ordre complet ; ordre simple
5 Transitivité et asymétrie	Ordre partiel strict	
6 Transitivité, asymétrie et complétude	Ordre fort	Ordre ; ordre strict ; ordre complet strict

[484] ARROW, Kenneth, *Social Choice and Individual Values*, New Haven, Yale university Press, 1951 (2ème édition révisée, 1961). Traduction française par l'Association de Traduction Economique de l'Université de Montpellier, Paris, Calmann-Levy, 1974; SEN, Amartya, *Collective Choice and Social Welfare,* San Francisco, Holden day, 1970

[485] Il s'agit de la terminologie de Sen.

On peut définir les relations de « préférence stricte » *P* et d' « indifférence » *I* en fonction de la relation de préférence faible *R*.

PREFERENCE STRICTE : $x\ P\ y \leftrightarrow [x\ R\ y\ \&\sim (y\ R\ x)]$

INDIFFERENCE : $x\ I\ y \leftrightarrow [x\ R\ y\ \&y\ R\ x]$

Un élément peut être caractérisé comme le « meilleur » élément (ou l'élément préféré) de l'ensemble *S* s'il est au moins aussi bon que (préféré à) n'importe quel autre élément de *S*.

MEILLEUR ELEMENT : Un élément *x* dans *S* est un meilleur élément de *S* eu égard à la relation de préférence *R* si et seulement si :

$$\forall y : (y \in S \rightarrow x\ R\ y)$$

CHOIX : L'ensemble des meilleurs éléments dans *S* est appelé l'ensemble de choix, et est noté ainsi :

C (S, R).

Annexe 3 : le théorème d'impossibilité d'Arrow (démonstration de Sen)[486]

Soit X un ensemble d'états sociaux. La relation de préférence de l'individu i est R_i, et il y a n individus : $i = 1, ..., n$.

Soit R la relation de préférence sociale. On suppose que chaque individu a un ordre, c'est-à-dire que, pour chaque i, R_i est réflexive, transitive et complète. De même, la relation R est un ordre[487].

DEFINITION 1 : une règle de choix social est une relation fonctionnelle f telle que, pour n'importe quel ensemble de n ordres individuels $R_1, ..., R_n$ (un ordre est attribué à chaque individu), une et seulement une relation de préférence sociale R est déterminée : $R = f(R_1, ..., R_n)$.

DEFINITION 2 : une fonction de bien-être social est une règle de choix social f . La gamme de ces règles correspond à l'ensemble de relations d'ordre sur X. Cette restriction est la condition O sur f.

La règle de choix social f doit remplir certaines conditions jugées « raisonnables »[488] :

CONDITION U (domaine non-restreint) : la règle f doit être valable pour n'importe quelle combinaison d'ordres individuels logiquement possible.

CONDITION P (Pareto-faible) : pour toute paire d'états sociaux x, y dans X, $[\forall i : x P_i y] \rightarrow x P y$.

CONDITION I (indépendance des alternatives non pertinentes) : soit R et R' les relations de préférence sociale binaires déterminées par f et correspondant respectivement à deux ensembles de préférences individuelles, $(R_1, ..., R_n)$ et $(R'_1, ..., R'_n)$. Si, pour toutes les paires

[486] ARROW, Kenneth, *Social Choice and Individual Values*, New Haven, Yale university Press, 1951 (2ème édition révisée, 1961).Traduction française par l'Association de Traduction Economique de l'Université de Montpellier, Paris, Calmann-Levy, 1974; SEN, Amartya, *Collective Choice and Social Welfare,* San Francisco, Holden day, 1970.Nous présentons ici la démonstration de Sen, simple et élégante.

[487] Sen a souligné que c'est une exigence sans doute trop forte : « Pour la relation de préférence sociale R, une telle supposition n'est pas facile à faire, et de fait c'est une partie de notre exercice que de voir si R aura ces caractéristiques ou non. Nous n'exigeons donc pas à ce niveau, par conséquent, que R soit un ordre. ». En revanche, la préférence sociale arrovienne est bel et bien un ordre.

[488] Les conditions correspondent à la version de 1963 du théorème d'Arrow. Dans cette version, la condition U n'a pas de nom et est la condition 1' ; la condition I de Sen la condition 3 d'Arrow ; la condition D de Sen est la condition 5 d'Arrow.

d'alternatives *x, y* dans un sous-ensemble *S* de *X*, $x R_i y \leftrightarrow x R'_i y$, pour tout *i*, alors *C(S, R)* et *C(S, R')* sont les mêmes.

CONDITION D (non-dictature) : il n'existe aucun individu *i* tel que, pour chaque élément dans le domaine de la règle *f*, $\forall x, y \in X : x P_i y \rightarrow x P y$.

La « société » est composée d'au moins deux « individus »[489], et ceux-ci se voient offrir au moins trois options distinctes.

THEOREME D'IMPOSSIBILITE D'ARROW : Il n'existe pas de règle de choix social *f* satisfaisant à la fois les conditions U, P, I et D.

La preuve de Sen introduit deux définitions supplémentaires et un lemme.

DEFINITION 3 : Un ensemble d'individus V est <u>à peu près décisif</u>[490] concernant la paire d'option *x, y* si *x P y* à chaque fois que $x P_i y$ pour chaque *i* dans V, et que $y P_i x$ pour chaque *i* hors de V.

DEFINITION 4 : Un ensemble d'individus V est <u>décisif</u> pour la paire *x, y* si *x P y* lorsque $x P_i y$ pour chaque *i* dans V.

Soit un certain individu *J*. *D(x, y)* signifie que *J* est « à peu près décisif » (cf. définition 3) pour *x, y*. *Ď(x, y)* signifie que *J* est « décisif » pour *x, y*.

LEMME 1 : S'il existe un individu *J* qui est à peu près décisif pour n'importe quel ordre d'alternatives, alors une règle de choix social satisfaisant les conditions U, P et I implique que *J* doit être un dictateur.

<u>Preuve :</u> Supposons que *J* est à peu près décisif pour une alternative donnée *x, y* : $\exists x, y \in X : D(x, y)$. Soit *z* une autre alternative, et *i* tous les individus autres que *J*. Supposons que *J* préfère strictement *x* à *y* et *y* à *z* : $x P_J y$ et $y P_J z$. Et supposons que dans le même temps tous les individus autres que *J* ont des ordres de préférence tels que : $y P_i x$ et $y P_i z$.

Si $[D(x, y) \& x P_J y \text{ et } y P_i x] \rightarrow x P y$.

---

[489] Evidemment, il n'y a pas vraiment de « choix collectif » pour des sociétés à un individu. Remarquons bien que les « individus » en question peuvent être des groupes partageant une même préférence (les Etats d'une fédération, les communautés religieuses dans un Etat, un groupe de travailleurs représenté par un syndicat et partageant une préférence pour l'augmentation de son salaire...), ou encore qu'un même « individu » peut avoir plusieurs préférences et « vivre » un problème analogue à celui du choix social (par exemple, lorsqu'il fait face à un dilemme, ou encore lorsqu'il fait l'épreuve de la faiblesse de sa volonté et préfère à la fois rédiger sa thèse et visiter des sites Web).

[490] Sen explique qu'un individu est à peu près décisif s'il gagne lorsque les individus non décisifs ont des préférences opposées aux siennes, et qu'il est décisif quelles que soient les préférences des individus non décisifs. Cela signifie qu'on ne peut savoir si un individu est à peu près décisif qu'une fois que l'on connaît le profil des préférences de chacune des personnes appelées à se prononcer, tandis que l'on peut spécifier *a priori* si un individu est décisif ou pas. Un individu pourrait être décisif alors même que tous les individus non décisifs peuvent être d'accord avec lui. Cette distinction est intéressante dans la perspective de l'analyse du pouvoir et de la liberté.

De plus, la condition P fait que : [$y\ P_J\ z$ &$y\ P_i\ z$] $\rightarrow y\ P\ z$.

Mais, [$x\ P\ y$ &$y\ Pz$] $\rightarrow x\ P\ z$ puisque P est supposée être transitive.

On arrive donc à $x\ P\ z$ sans rien savoir des préférences des individus autres que J concernant cette paire x, z. Or, du fait de la condition I, les préférences des i sur les autres paires x, y et y, z ne devraient pas avoir d'incidence sur la préférence sur x, z. Par conséquent, $x\ P\ z$ doit être la conséquence de la préférence de J pour x plutôt que z. Mais cela implique que J est décisif sur x, z.

(Proposition 1) $D\ (x, y) \rightarrow \check{D}(x, z)$.

Supposons à présent que $z\ P_J\ x$ et $x\ P_J\ y$, tandis que $z\ P_i\ x$ et $y\ P_i\ x$. La condition faible de Pareto requiert donc que la société préfère également z à x : $z\ P\ x$. Sachant que $D\ (x, y)$, que $x\ P_J\ y$ et que $y\ P_i\ x$, il s'en suit que $x\ P\ y$. La préférence sociale étant transitive, $z\ P\ y$. Or,

$z\ P_J\ y$, tandis que nous ne connaissons pas les préférence de i pour cette alternative. De ce fait, J se trouve être décisif pour le classement de z et de y. Donc :

(Proposition 2) $D\ (x, y) \rightarrow \check{D}(z, y)$.

Le même raisonnement peut être effectué en intervertissant y et z. Donc :

(Proposition 3) $D\ (x, z) \rightarrow \check{D}(y, z)$.

Le raisonnement conduisant à la Proposition 1 avec x à la place de z, z à la place de y et y à la place de x fait que l'on obtient :

(Proposition 4) $D\ (y, z) \rightarrow \check{D}(y, x)$.

Récapitulons : $D\ (x, y)$ implique cinq propositions :

$D\ (x, y) \rightarrow \check{D}\ (x, z)$, d'après la proposition 1.

$\rightarrow D\ (x, y)$, d'après les définitions 3 et 4.

$\rightarrow \check{D}\ (y, z)$, d'après la proposition 3.

$\rightarrow D\ (y, z)$.

$\rightarrow \check{D}\ (y, x)$, d'après la proposition 4.

Par conséquent,

(Proposition 5) $D\ (x, y) \rightarrow \check{D}\ (y, x)$.

En intervertissant x et y dans les propositions 1, 2 et 5, nous obtenons la proposition suivante :

(Proposition 6) $D\ (y, x) \rightarrow$ [$\check{D}\ (x, y)$ &$\check{D}\ (z, x)$ & $\check{D}\ (x, y)$].

Puisque être décisif implique être à peu près décisif :

D (x, y) $\rightarrow$ *Ď (y, x)*, selon la proposition 5.

$\rightarrow$*D (y, x).*

Il s'en suit que, à partir de la proposition 6 :

(Proposition 7) *D (x, y)* $\rightarrow$ [*Ď (x, y)* &*Ď (*z, *x)* & *Ď (x, y)*].

Si l'on combine les propositions 1, 2, 5 et 7, nous constatons que, étant données les conditions U, P et I, *J* est décisif pour n'importe quelle alternative de l'ensemble *(x, y, z).* Par conséquent, il suffit qu'un ensemble de trois options comprenne *x* et *y* pour que *J* soit un dictateur.

Supposons à présent qu'il y ait plus de trois options. De cet ensemble, nous prenons *u* et *v*.

(5) Soit *u* et *v* sont identiques à *x* et *y,* auquel cas la démonstration de *Ď (x, y)* ci-dessus vaut également pour *Ď (u, v).*

(6) Soit seulement *u* ou *v* est identique à *x* ou *y* : l'ensemble des options consistera par exemple de *x = u,* de *y* et de *v.* Sachant que *D (x, y),* il s'en suit que *Ď (u, v)* et que *Ď (v, u).*

(7) Soit *u* et *v* sont différents de *x* comme de *y*.

   a. Avec *(x, y, u),* on obtient *Ď (x, u),* et *Ď (x, u)* $\rightarrow$*D (x, u).*

   b. Avec *(x, u, v),* et sachant que *D (x, u),* il s'en suit que *Ď (u, v)* et *Ď (v, u)* d'après la démonstration ci-dessus.

Par conséquent, pour *x* et *y* donnés, si *D (x, y),* alors *Ď (u, v)* pour toutes les paires ordonnées possibles *(u, v).* Par conséquent, *J* est un dictateur.

- Le lemme 1 est démontré.

Ce lemme est utilisé pour démontrer le théorème d'impossibilité d'Arrow.

Preuve : le raisonnement est le suivant. Pour démontrer l'impossibilité de l'existence d'une règle de choix social satisfaisant simultanément U, P, I et D, il faut prouver qu'il y a nécessairement un individu qui est à peu près décisif pour la préférence prononcée sur une paire d'options. Pour ce faire, la démonstration consiste à montrer que l'affirmation contraire aboutit à une contradiction (il s'agit d'un raisonnement par l'absurde).

La condition P stipule que, pour chaque paire d'options, l'ensemble de tous les individus est décisif. Il existe donc au moins un ensemble d'individus décisifs. Par conséquent, il existe donc au moins un ensemble d'individus à peu près décisifs. Prenons l'ensemble d'individus à

peu près décisifs le plus petit sur une paire d'options quelconque. Cet ensemble sera désigné par la lettre *V*, et il sera décisif sur *(x, y).*

Si *V* contient plus d'un individu, on distingue deux sous-ensembles de *V*. V_1comprend un unique individu, et V_2 comprend le reste des individus appartenant à *V*.L'ensemble *W* est l'ensemble de tous les individus qui n'appartiennent pas à *V*(dans le cas où tous les individus sont décisifs, l'ensemble *W* est tout simplement vide).

La condition U stipule que toutes les préférences logiquement possibles doivent être prises en compte. Prenons donc les préférences suivantes :

$\forall i$ dans V_1, $x\ P_i\ y\ \&y\ P_i\ z$.

$\forall j$ dans V_2, $z\ P_j\ x\ \&x\ P_j\ y$.

$\forall k$ dans *W*, $y\ P_k\ z\ \&z\ P_k\ x$.

Puisque *V* est à peu près décisif sur *(x, y),* et que chaque individu dans *V* préfère *x* à *y,* et chaque individu dans *W* préfère *y* à *x,* il s'en suit que *x P y.*

V_2 ne peut pas être décisif sur *(y, z),* car V_1 est aussi à peu orès décisif et a une préférence opposée. Par conséquent, *~(z P y).*

Rappelons que la relation de préférence sociale doit être complète et transitive. Il faut donc que *y R z.* Si $x\ P\ y\ \&y\ R\ z \rightarrow x\ P\ z$. Le problème est que V_2, et *W,* préfèrent strictement *z* à *x.* Il se trouve donc que l'individu *i* dans V_1 est à peu près décisif pour une paire d'options. Or, d'après le Lemme 1, si un individu est à peu près décisif sur une paire, il doit être un dictateur. Or, cela viole la condition D.

- Le théorème d'impossibilité de l'existence d'une règle de choix social satisfaisant simultanément U, P, I et D, dit « théorème d'Arrow », est donc démontré.

Annexe 4 : l'impossibilité du libéral parétien[491]

Cette démonstration reprend les définitions de la préférence et de l'indifférence, la Définition 1 et la Définition 2, ainsi que les conditions U et P du théorème d'Arrow. Sen a proposé de remplacer les fonctions de bien-être social utilisées par Arrow par des fonctions de décision sociale. Ces fonctions ne requièrent pas une préférence sociale qui ait la forme d'un ordre. Il suffit qu'elle puisse générer une fonction de choix. Cela permet d'étendre la gamme des règles de choix social susceptibles de produire un choix social.

DEFINITION 5 : Une fonction de décision sociale est une règle de choix social f . La gamme de ces règles inclut toute relation de préférence faible R susceptible de générer une fonction de choix $C(S, R)$ sur l'ensemble des options X. Cette restriction est la restriction O* sur f.

Si X est fini, une fonction de bien-être social est toujours une fonction de décision sociale, mais pas inversement[492].

Une condition supplémentaire dite de « libéralisme » est introduite. Cette condition représente la garantie de la protection d'une « sphère privée », une paire d'options sur laquelle l'individu concerné est décisif.

CONDITION L (libéralisme) : Pour tout individu i, il doit exister au moins une alternative (x, y) tel que l'individu est décisif pour cette alternative. Cela signifie que ses préférences sont représentées dans la préférence sociale, et ce quelques soient les préférences des autres membres de la société. Par conséquent :

$x P_i y \rightarrow x P y.$

$y P_i x \rightarrow y P x.$

Afin de rendre la preuve encore plus forte, on peut restreindre la condition de libéralisme en faisant en sorte qu'elle ne tienne que pour deux individus dans la société (si elle ne tenait que pour un, ce serait une dictature, pas une société libérale).

CONDITION L* (libéralisme minimal) : il existe au moins deux individus j et k, chacun étant décisif sur une paire qui lui est attribuée, soit (x, y) pour k et (z, w) pour j.

---

[491]SEN, Amartya, "The Impossibility of a Paretian Liberal", *Journal of Political Economy,* vol. 78, no 1, jan-mar 1970, pp 152-157.

[492] Si X est infini, une fonction de bien-être social n'est pas forcément une fonction de décision sociale (un classement de préférences sur un nombre infini d'états sociaux ne permet pas systématiquement de définir un ensemble de choix).

THEOREME D'IMPOSSIBILITE DU LIBERAL PARETIEN : il n'existe pas de fonction de décision sociale satisfaisant U, P et L*.

Preuve :

(x, y) et *(z, w)* doivent être distinctes pour que la condition L* puisse tenir.

(8) Si elles ont un élément en commun, par exemple, si *x* est identique à *z* :

Dans le cas (autorisé par U) où $x\ P_k\ y$, $w\ P_j\ x$ et où pour tout *i*, $y\ P_i\ w$, L* requiert que *x P y* et *w P x*, tandis que P requiert que *y P w*. Cela crée un cycle : il n'y a pas de meilleur élément et donc pas de choix possible.

(9) Supposons alors que *(x, y)* et *(z, w)* n'ont aucun élément en commun :

Dans le cas (autorisé par U) où $x\ P_k\ y$, $z\ P_j\ w$ et où pour tout *i*, ($w\ P_i\ x$ & $y\ P_i\ z$), L* requiert que *x P y* &*z P w*. La condition P requiert que *w P x* &*y P z*. Cela viole également l'acyclicité requise pour spécifier un ensemble de choix. Par conséquent, il n'existe pas de fonction de décision sociale qui puisse satisfaire simultanément L*, P et U.

- Le théorème d'impossibilité du libéral parétien est démontré.

COROLLAIRE : Il n'existe pas de fonction de décision sociale qui puisse satisfaire simultanément U, P et L.