

Le conséquentialisme des raisons de Parfit

Christophe Salvat

Première version : 35300 caractères

1. Introduction

Parfit a toujours entretenu avec l'utilitarisme une relation assez ambivalente. Son premier ouvrage, *Reasons and Persons* (ci-après RP) publié en 1984, a été remarqué à la fois pour ses critiques acerbes du raisonnement conséquentialiste (notamment via la conclusion répugnante) et pour sa théorie du réductionnisme ontologique, qui constitue la meilleure réponse apportée à ce jour à l'objection de la séparabilité des personnes, présentée par Rawls dans les années 70. Son second et dernier ouvrage, *On What Matters* (OWM) dont les deux premiers volumes furent publiés en 2011 et le troisième en 2017, permettent de mieux situer sa pensée. A première vue, Parfit réfute l'idée même de philosophies morales rivales, arguant que – bien comprises – les théories utilitaristes, kantienne et contractualistes peuvent être reformulées dans une seule et même théorie. De fait, Parfit développe une théorie conséquentialiste et hédoniste également influencée par l'objectivisme de Tim Scanlon et Thomas Nagel. Nous l'appellerons le conséquentialisme des raisons.

2. La nature objective des raisons

Etre rationnel, nous dit Parfit, décrit notre capacité à répondre à des raisons. Il en existe de deux types. Les raisons fondées sur les désirs (*desire-based reasons*) et les raisons fondées sur les valeurs (*value-based reasons*) (OWM, I, 45). Dans le premier cas, nos raisons sont relatives aux faits satisfaisants (ou appelés à satisfaire) nos désirs. Dans la mesure où l'importance de ces faits n'est relative qu'à nos propres désirs, et donc à notre propre personne, on peut selon Parfit également les appeler raisons relatives-au-sujet (*subject-given reasons*). Leurs forces dépendent de la force de nos désirs. Dans le second cas, les raisons sont relatives aux faits qui concernent les objets de nos désirs. On parle alors de raisons relatives-à-l'objet (*object-given reasons*). Leurs forces dépendent des faits qui nous donnent des raisons de les désirer. Ajoutons à ces deux catégories un troisième type de raisons relatives-à-l'état (*state-given reasons*), qui à la différence des raisons relatives-à-l'objet ne sont pas relatives aux faits qui concernent les objets de nos désirs mais aux états constitués par le fait d'avoir ces désirs (OWM, I, 50). Nous pouvons

avoir, par exemple, des raisons de croire en Dieu indépendamment du fait que cette croyance soit vraie ou non. Parfit considère cette dernière catégorie comme secondaire (OWM, I, 51).

Parfit rejette l'approche des raisons propriétés-des-désirs, traditionnellement utilisée pour justifier l'hédonisme, et défend une interprétation objectiviste des raisons, qui a habituellement la préférence des contractualistes. Ce parti pris, assuré avec force dès les premières pages de *On What Matters*, permet à Parfit de répondre, au moins partiellement, au dualisme de la raison pratique de Sidgwick.

Sidgwick, rappelons-le, considère que nous avons des raisons d'agir pour le mieux, pensé de façon impartiale, à moins que l'on ait également des raisons d'agir pour ce que nous considérons être notre propre bien. Dans ces cas-là, nous avons des raisons suffisantes d'agir dans un sens comme dans l'autre, mais aucune raison décisive d'agir dans un sens plutôt que dans un autre. Les deux alternatives, qui expriment respectivement le point de vue objectif (ou, selon l'expression de Sidgwick, 'le point de vue de l'Univers') et le point de vue subjectif, sont tous deux parfaitement rationnels. L'indétermination structurelle à laquelle se heurte le processus de décision individuelle chez Sidgwick est très déstabilisante pour les théoriciens de l'éthique, qui, à l'instar de Parfit, tentent de donner à l'action morale une justification rationnelle. Cela explique l'importance donnée à la résolution du dualisme de la raison pratique dans l'œuvre de ce dernier. Celle-ci s'opère en deux temps.

Parfit conteste, tout d'abord, le fait qu'il existe un fondement subjectif à la rationalité individuelle. Toutes les raisons sont, dit-il, nécessairement objectives. Il s'agit-là d'une position forte et assez nouvelle par rapport à *Reasons and Persons* (même si on la voit poindre ici et là). Pour la justifier, Parfit est amené 1° à discréditer l'idée selon laquelle la satisfaction des désirs considérés comme propriétés-du-sujet est nécessairement rationnelle et 2° à montrer que la satisfaction des désirs considérés comme propriétés-de-l'objet permet d'évaluer la rationalité d'une action. La première condition ne pose pas de problèmes particuliers¹. Dans la lignée de 'l'Indifférence à l'égard des futurs mardis' (*Future Tuesdays Indifference*), développé dans *Reasons and Persons*, Parfit développe ce qu'il nomme 'l'argument de l'agonie' (*The Agony Argument*), dans lequel il montre qu'une théorie des raisons fondée sur les désirs ne peut exclure la possibilité que des

¹ Parfit développe trois arguments à l'encontre des théories subjectives ou théories de la satisfaction des désirs. La première est qu'en fait il s'en agit pas vraiment d'une et, que les théoriciens les plus avertis ont nécessairement recours à l'objectivité des raisons. Cette critique est, je crois, tout à fait valide. Très peu de théoriciens admettent que sommes rationnels dès lors que nous satisfaisons nos désirs *quels qu'ils soient*. Certains de nos désirs sont absurdes ou peuvent nous être gravement préjudiciables. Pour se parer contre de telles éventualités, les théoriciens font alors appel à la notion de désirs informés, c'est-à-dire aux désirs que l'on aurait dans des conditions idéales. C'est le cas, notamment, des théoriciens de l'hédonisme, ce qui fait dire à Parfit, que contrairement à ce que les gens pensent l'hédonisme, et *a fortiori* l'utilitarisme, ne relève pas d'une approche subjective mais d'une approche objective des raisons (OWM, I, 67).

agents n'aient aucune raison de vouloir éviter une agonie future (OWM, I, 76). Or cette conclusion est simplement irrecevable pour Parfit, selon qui aucun être que l'on pourrait qualifier de rationnel ne peut volontairement et consciemment désirer souffrir l'agonie. En d'autres termes, la rationalité ne se réduit pas simplement à la cohérence de nos préférences mais concerne également la nature de celles-ci. Mais discréditer les fondements subjectifs de la rationalité ne suffit pas en soi à justifier une approche objective des raisons.

Déontologues et utilitaristes s'accordent généralement pour donner la priorité au point de vue impartial (*the view from nowhere*) sur le point de vue personnel. Pour Sidgwick, ces deux points de vue sont même irréconciliables. Soit j'adopte une perspective universaliste (et donc impartiale), soit j'adopte un point de vue personnel mais je ne peux pas à la fois me préférer aux autres et me considérer comme l'égal des autres. Pour Sidgwick, j'alterne les points de vue sans que l'on puisse dire exactement pourquoi j'adopte l'une plutôt que l'autre. Pour Parfit, au contraire, les deux points de vue sont concomitants. Le fait que j'adopte un point de vue personnel n'exclut pas le fait que je puisse penser de façon impartiale. Adopter un *point de vue* impartial n'implique pas que l'on soit impartial (OWM, 1, 41)². Le point de vue personnel n'est en effet pas réductible à l'égoïsme rationnel. Je suis *à la fois* capable de désirer être sauvé et de désirer que le maximum de personnes soit sauvé. Le fait que ces raisons diffèrent, voire rentrent en conflit, ne justifie pas de conclure, à l'instar de Sidgwick, à l'indétermination du choix. Le fait qu'il soit impossible de prédire avec rigueur l'issue de la procédure de décision relève davantage de la difficulté que les agents ont à comparer la force relative de ces différentes raisons plutôt que de l'impossibilité qu'ils ont à penser simultanément comme sujet et comme spectateur impartial.

La première étape de la résolution du dualisme pratique par Parfit consiste donc à rejeter la thèse de Sidgwick selon laquelle la rationalité peut se concevoir de plusieurs points de vue, deux pour être exact, et à lui substituer un point de vue unique, le point de vue objectif, mais deux façons de le traiter, avec ou sans *parti pris*³. C'est une implication directe de la thèse du 'Tous pour un' (*The Any-All thesis*) selon laquelle « Si quelqu'un a une raison personnelle de vouloir que quelque chose arrive, tout le monde a une raison impartiale, *bien que peut-être de façon plus faible*, de vouloir que cette chose arrive également. » (OWM, III, 321, je souligne). Le fait que je reconnaisse que les raisons d'autrui ne soient pas moins fondées que les miennes ne m'engage pas à les traiter comme les miennes. Elles n'ont pas pour moi la force qu'elles ont pour eux. C'est un second point de divergence avec l'utilitarisme classique mais également avec les théories contractualistes. Parfit est, en effet, l'un des rares théoriciens à ne pas faire de l'impartialité le fondement de la morale.

² Inversement, la subjectivité n'est pas incompatible, pour Parfit, à l'impartialité des raisons. Je peux reconnaître, en tant que sujet, que ma vie ne vaut pas mieux que celle qu'un autre (OWM, 1, 139).

³ L'expression de *parti pris* n'est pas de Parfit, nous l'utilisons ici pour désigner le caractère personnel des raisons objectives.

Il est alors possible de remettre en cause le principe de l'incommensurabilité des raisons et de mettre fin à l'indétermination de la raison pratique. C'est la deuxième étape du raisonnement de Parfit qui s'appuie sur la force relative des raisons qu'elles soient personnelles à l'agent ou non. Si Parfit parvient à démontrer que l'agent, en dépit de son parti pris, est capable de reconnaître la force normative des raisons qui ne lui sont pas personnelles, alors on pourra dire qu'il a réussi, non seulement à résoudre le dualisme de la raison pratique, mais également à (re)donner un fondement rationnel à la morale. On peut, je crois, lui reconnaître le fait qu'il y parvient dans certains cas, ceux pour lesquels la force normative de la raison est telle qu'elle anéantit l'effet de tout parti pris.

3. L'empreinte hédoniste des raisons

J'ai montré dans la section précédente que, bien que les raisons soient objectives, elles ne s'appliquent pas nécessairement de façon égale. Parfit rejette à la fois la position dite de la neutralité par rapport à l'agence, qui nie au sujet toute spécificité à l'égard des raisons, et la position dite de relativité par rapport à l'agence, qui considère la perspective de l'agent comme fondamentale dans l'évaluation des raisons. Il adopte une position dualiste à mi-chemin entre la théorie défendue par Nagel dans *Possibility of Altruism* (Nagel I) et celle qu'il a ensuite défendue dans *View from Nowhere* (Nagel II). Mais, à la différence de Sidgwick, ce dualisme des raisons n'est pas nécessairement synonyme d'indétermination. Pour Parfit, il existe deux types de raisons, les raisons déontiques et les raisons non-déontiques. Leur force relative dépend du degré de partialité avec lequel elles sont évaluées, qui dépend lui-même du caractère plus ou moins personnel qu'elles revêtent pour l'agent.

Les raisons déontiques peuvent être définies comme des devoirs ou des prohibitions de nature non-conséquentialiste. Il en existe deux types. Les premières sont fondées en raison. Il s'agit de principes ayant valeur universelle qui se justifient par le fait qu'ils pourraient être rationnellement choisis et/ou acceptés par tous, selon la version kantienne (*the Kantian Formula*), ou que personne ne pourrait raisonnablement rejeter, selon la version de Scanlon (*the Scanlon Formula*). Mais il peut également s'agir, non de principes, mais de devoirs ou de prohibitions *prima facie*, selon l'expression de David Ross (Ross, 1930), autrement dit de devoirs auxquels on doit se conformer à moins qu'on ait de meilleurs raisons de faire autrement. Ne pas mentir ou ne pas rompre sa promesse sont des exemples de raisons déontiques. A la différence de Kant, cependant, les prohibitions ou devoirs déontiques ne sont pas impératifs chez Parfit. Leur pouvoir décisionnel dépend de leur force relative. Il existe ainsi des cas où le mensonge, voire même le meurtre, peut être moralement justifié.

Les raisons non-déontiques sont de nature conséquentialiste, essentiellement hédonistes. A la différence des utilitaristes, les raisons non-déontiques, telles que la recherche de son bien-être personnel et de celui des autres, ne sont pas – ou plus exactement ne doivent pas être – systématiquement décisives. La force des raisons déontiques peut l'emporter sur celle des raisons non-déontiques. Selon la force que l'on donne aux unes et aux autres, cette théorie éthique prend plutôt une couleur conséquentialiste ou une couleur déontologique⁴. Les raisons non-déontiques ne sont pas nécessairement partiales dans la mesure où Parfit adopte une approche objective de l'hédonisme. Le bien-être des autres constitue ainsi une raison non-déontique mais impersonnelle. Sa force est *a priori* plus faible que les raisons non-déontiques personnelles mais rien ne nous permet de nous prononcer sur son poids par rapport aux raisons déontiques. Parfit reste en effet très évasif sur la force relative des raisons déontiques. L'essentiel de ses démonstrations porte sur la comparaison des forces entre raisons non-déontiques personnelles et impersonnelles qui sont, elles, au moins approximativement, comparables (OWM, 1, 137).

S'il est – théoriquement du moins – facile d'évaluer la force d'une raison dans le cas de raisons relatives-au-sujet (celle-ci est simplement fonction de l'intensité du désir ou de la préférence de l'agent), l'évaluation objective de la force des raisons relatives-à-l'objet pose davantage de difficultés. Elle dépend, notamment, de la qualité de la raison relative-à-l'objet, celle-ci pouvant être de nature déontique ou non-déontique. Dans le premier cas, la force de la raison relative-à-l'objet dépend de la valeur du devoir ou de la prohibition. Dans le second cas, elle dépend de la valeur de l'action elle-même. Dans les deux cas, je vais essayer de montrer, qu'elle passe néanmoins par un raisonnement de type conséquentialiste.

Commençons par le cas le plus simple, celui des raisons non-déontiques dont la force est déterminée par le degré de 'bonté' (*goodness*) de l'action. Parfit distingue, dès les premières pages de *Reasons and Persons*, trois grands types de théories de la valeur : les théories hédonistes, les théories du succès et les théories des listes objectives (RP, 4). Cette catégorisation est aujourd'hui très largement adoptée aujourd'hui. Mais, de façon assez ironique, il développe de fait dans *On What Matters* une quatrième approche, à mi-chemin entre hédonisme et liste objective (seule les théories du succès ou de la satisfaction des préférences ne trouvent aucun mérite à ses yeux).

Bien qu'il ne soit pas hédoniste au sens strict du terme, dans la mesure où il refuse de faire de la recherche du plaisir le motif exclusif de l'action humaine et qu'il lui donne un fondement objectif plutôt que subjectif, Parfit accorde une place très importante au bien-être des individus. Plus exactement, il reconnaît la diminution de la souffrance humaine comme l'une des premières causes légitimes de l'action. Il se distingue en cela de Kant et de Ross, notamment, qu'il accuse de ne pas avoir réalisé que la souffrance,

⁴ Parfit contesterait sans doute cette assertion arguant que sa philosophie dépasse ces distinctions.

bien qu'elle ne soit pas 'mauvaise' (*bad*) moralement parlant, n'en est pas moins 'mauvaise' dans une perspective des raisons (*reason-implying sense*) (OWM, I, 371-72). Il va même plus loin, et reprend la thèse de Nagel selon laquelle, la souffrance est mauvaise, non seulement d'un point de vue personnel mais également d'un point de vue impersonnel (*impersonally bad*) (OWM, I, 372). L'argument de l'Agonie (OWM, I, 76) montre qu'il est irrationnel, quelques soient nos préférences subjectives, de préférer une situation de plus grande souffrance à une situation de moindre souffrance, y compris, – et c'est un point important – lorsque la plus grande souffrance est plus éloignée dans le temps que la moindre souffrance. L'objectivité et l'universalité de la souffrance humaine justifient, il me semble, l'adoption par Parfit du principe de neutralité par rapport au temps⁵.

Plus une souffrance est importante, plus on a de raisons de vouloir l'éviter. Deux facteurs sont décisifs pour déterminer la force relative de ces raisons. Le premier facteur est, tout d'abord, celui de l'agence. La force de la raison n'est pas la même si la souffrance est mienne ou non si elle est celle d'un proche ou d'un parfait étranger. L'objectivité des souffrances ne se traduit pas nécessairement par un principe de neutralité par rapport à l'agence. Le second facteur est celui de la quantité et/ou de l'intensité de la souffrance ou du plaisir. Contrairement aux utilitaristes classiques, Parfit rejette la possibilité d'une quantification exacte des plaisirs et des peines mais reconnaît, néanmoins, l'existence d'une relation de causalité entre ce qu'on appellera la *gravité* des souffrances (ou *l'ampleur* des plaisirs) et la force des raisons : « Si nous devons choisir entre une douleur brève mais intense et une autre douleur beaucoup plus longue et moins intense, on pourrait dire qu'une de ces expériences est pire que l'autre, dans le sens où on aurait plus de raisons de choisir l'une plutôt que l'autre. Mais il ne pourrait pas, je suggère, y avoir de vérité précise sur la force relative de ces raisons. » (OWM, I, 132). Imaginons le cas d'un bateau échoué dans lequel un seul homme (sur deux) puisse être sauvé. Le fait que l'autre passager doive vivre légèrement plus longtemps, et légèrement plus heureux, que vous, ne constitue pas en soi une raison pour vous de vous sacrifier pour lui (OWM, III, 340), pas plus pourrait-on ajouter, que cela ne constitue une raison décisive pour un sauveteur de le préférer à vous. De la même manière, le nombre de personnes affectées par la souffrance (ou la mort) a une influence importante sur la force de nos raisons mais celle-ci n'est pas directement fonction de celui-ci. La question de la force des raisons ne se pose naturellement pas de la même façon si je (ou une personne extérieure) met(s) en balance ma propre souffrance (ou mort) et celle de deux, cinq, vingt ou plusieurs centaines de milliers de personnes. Cela explique pourquoi, qu'on peut avoir des raisons suffisantes mais non décisives de sacrifier pour sauver la vie de deux ou trois autres personnes, mais qu'on a des raisons décisives de se sacrifier pour sauver la vie de plusieurs milliers de personnes (ainsi que l'exemple des malheureux liquidateurs de Tchernobyl l'a montré). Les raisons non déontiques, fondées sur la

⁵ On comprend mal, sinon, le statut de ce principe dans la pensée d'un auteur ayant adopté une approche réductionniste de l'identité personnelle et développé une théorie de la rationalité fondée sur les désirs *présents* (CP).

gravité des souffrances ou l'ampleur des plaisirs, ce que nous avons appelé l'empreinte hédoniste, qu'elles soient personnelles et/ou impersonnelles, sont toujours comparables, mais elles ne le sont qu'*approximativement*.

Notons que a grande majorité des exemples présentés par Parfit évoque la souffrance ou la mort potentielle. C'est le cas notamment (dans OWM) de *Lift*, de *Tunnel*, de *Sidetrack* et de *Bridge* (qui sont tous des reformulations du célèbre cas du *Trolley*), de celui la *Nuclear Policy* (inspiré des bombardements de Hiroshima et Nagasaki), ou encore de ces exemples fictifs aux noms évocateurs : *Child Drowning*, *Shipwreck*, *Lifeboat*, *Fire*, *Earthquakes*, *Desperate Plight*, *Whimsical Despot*, *Murderous Theft*, *Gangster*, *Lesser Evil*, *Depression* etc. Rares sont les exemples qui justifient la force d'une raison non-déontique par l'intensité et/ou la quantité de plaisir qu'elle pourrait procurer. Doit-on en conclure que Parfit défend, au moins implicitement, une forme de conséquentialisme négatif ? On peut raisonnablement le penser⁶. Parfit semble être d'avis – et je le rejoins sur ce point – qu'aucun plaisir, aussi intense soit-il et/ou aussi répandu soit-il, ne peut justifier la souffrance, serait-elle modérée, d'autrui. Alors que, dit-il, la qualité de la vie a une valeur (positive) nécessairement limitée, il n'existe pas de limites au mal (*badness*) produit par les souffrances humaines (RP, 407).

Parfit fait face, il me semble, à une importante difficulté. Si la force des raisons impersonnelles est fonctionnellement liée au bien auxquelles elles sont liées, alors il devrait toujours être possible de se déterminer, une fois la force de nos raisons personnelles établies. Il est possible que la force de la raison personnelle 'sauver ma vie' soit au moins trois fois supérieure à celle de la raison impersonnelle 'sauver une autre vie', même si l'on est dans l'incapacité de déterminer un ratio entre les forces relatives de ces deux raisons. Mais je dois être capable d'identifier un seuil, même approximatif, à partir duquel la force de la raison 'sauver x vies' est clairement supérieure à la force de la raison 'sauver ma vie'. Si tel n'est pas le cas, on sera contraint de conclure, avec Sidgwick, que ces raisons ne sont pas comparables. Or, cela est impossible pour deux raisons. La première est qu'il semble difficile, voire impossible, de déterminer un standard d'équivalence entre raisons personnelles et raisons impersonnelles. Si chacun admet, par exemple, qu'il serait immoral de préférer s'épargner une égratignure que de sauver la vie d'un tiers, il est extrêmement difficile d'établir - même approximativement - à partir de quel degré de gravité, il est tolérable de sacrifier la vie d'un autre pour s'éviter une blessure. Le second élément, mis en avant par Parfit, est que l'empreinte hédoniste des raisons (partiales ou impartiales) n'est pas seule à devoir être prises en compte. Parfit croit ainsi que « nous avons des raisons personnelles et impartiales qui ne nous sont pas fournies, ou pas totalement fournies, par les faits relatifs à notre bien-être ou celui des autres. Nous avons des raisons d'essayer de faire différentes choses qui ont

⁶ Dans *Reasons and Persons*, Parfit affirme néanmoins que « le soulagement de la souffrance devrait avoir un poids plus important que celui qui lui ai ordinairement donné par les Utilitaristes Positifs. Mais il ne semble pas plausible qu'il doive avoir priorité absolue, ou même un poids beaucoup plus important. » (RP, 344-45).

de valeur en elles-mêmes, telles que découvrir la vérité, créer ou préserver la beauté, réaliser plusieurs autres buts bons ou valables. Nous avons aussi des raisons morales à la fois neutres et relatives à la personne, comme les raisons de prendre soin de nos propres enfants, et des raisons de faire en sorte que les avantages et les inconvénients soient plus également distribués entre les différentes personnes » (OWM, III, 321). Parfit conclut que « bien que l'hédonisme soit une part importante de la vérité concernant ce qui est bon ou mauvais, nous refusons d'admettre qu'il s'agisse là de toute la vérité » (OWM, III, 353). Le poids des raisons déontiques ne serait donc pas à minimiser. Mais comment l'évaluer ? Comment comparer une raison non déontique, telle que la recherche du confort personnel, à une raison déontique, telle que la préservation de la nature ? Selon que l'on raisonne en conséquentialiste ou en déontologue, on aura tendance à donner systématiquement plus de poids aux raisons déontiques ou, au contraire, aux raisons non déontiques, rendant ainsi factice le dépassement de l'indétermination sidgwickienne. Une possibilité pourrait, naturellement, consister à évaluer le poids des raisons déontiques en fonction de la valeur de la règle qui les porte, autrement dit, à donner une raison non déontique aux raisons déontiques. C'est ce que fait le conséquentialisme de la règle, auquel Parfit semble adhérer. Néanmoins, ainsi que je le suggère ci-dessous, le conséquentialisme des raisons n'est pas réductible à un conséquentialisme de la règle.

4. Conséquentialisme des raisons vs conséquentialisme de la règle

J'ai essayé de montrer ci-dessus que la théorie des raisons qu'il développe dans *On What Matters* est dotée d'une forte dimension hédoniste et conséquentialiste, même si les raisons déontiques y prennent une place plus ou moins importante selon les cas. J'ai appelé cela le conséquentialisme des raisons. Pour conclure cet article, je voudrais montrer que le conséquentialisme des raisons diffère fondamentalement du conséquentialisme de la règle, auquel Parfit apporte pourtant son soutien.

Parfit se montre critique vis-à-vis de l'utilitarisme de l'acte et semble défendre – par défaut – un utilitarisme de la règle. Trois éléments théoriques peuvent motiver ce choix. La perspective kantienne adoptée dès le premier volume de *On What Matters* en constitue le premier. Le conséquentialisme de la règle, qui implique que « nous devrions agir selon les principes ou les règles qui sont optimifiques, qui sont universellement souhaitables et qui ne peuvent pas être raisonnablement rejetés » (principe F, OWM, III, 434), est, en effet, le produit logique de deux principes universellement acceptés (*UA principles*), l'argument kantien selon lequel « tout le monde devrait suivre les principes dont l'acceptation universelle est rationnellement voulue ou choisie par tous » (principe A) et l'argument conséquentialiste selon lequel « il existe des principes optimifiques universellement acceptés que tout le monde a des raisons impartiales de choisir » (principes C et D, OWM, I, 378). Parfit appelle cela l'*Argument Kantien pour le Conséquentialisme de la Règle* (OWM, I, 379).

Le deuxième élément est introduit dans le premier volume de OWM et est repris et développé dans le troisième. Il s'agit de la légitimité de la préférence pour ses proches, pratique remise en cause par le principe d'impartialité, commun aux approches déontologiques et utilitaristes, principe aussi appelé principe de neutralité par rapport à l'agence. Notez, j'y reviendrais plus bas, que Parfit ne rejette pas totalement le principe d'impartialité à l'égard des personnes, mais qu'il accorde néanmoins un statut particulier aux personnes qu'on aime, et en particulier à ses enfants. La neutralité par rapport à l'agence, explique Parfit, est contraire à l'intuition morale et est rationnellement injustifiable. Son adoption devrait logiquement être contraire à toute philosophie morale conséquentialiste et hédoniste : « Si nous prenions également soin du bien-être de tout le monde, nous serions dans l'incapacité de vraiment aimer une personne particulière, et nous perdriions ainsi beaucoup de nos objectifs, centres d'intérêts et ambitions desquels dépend notre bonheur » (OWM, III, 414). Le conséquentialisme de l'acte, selon Parfit, ne permet pas de différencier l'utilité de ceux qu'on aime de celle des autres⁷. Le conséquentialisme de la règle, au contraire, permettrait d'établir une préférence, universellement acceptée, entre les personnes qu'on aime et celles qui nous sont étrangères. Les principes optimifiques « nous permettraient, dans de nombreux cas, et nous obligeraient dans beaucoup d'autres, à donner une forte priorité au bien-être de nos propres enfants. » (OWM, I, 385).

Le troisième et dernier élément en faveur du conséquentialisme de la règle⁸ est assez classique et a trait à l'imperceptibilité des effets de certaines actions individuelles, imperceptibilité qui pourrait justifier une absence d'action dans le cadre d'un conséquentialisme de l'acte. Contrairement à ce dernier, le conséquentialisme de la règle permet de justifier, voire d'exiger, de suivre les règles, qui, si elles sont universellement suivies, produisent le plus grand bien, même si l'effet d'une seule action est imperceptible pour l'agent. C'est le cas, notamment, du vote démocratique dans de grands pays ou de la lutte contre le réchauffement climatique.

Malgré tout cela, et au regard du troisième volume de OWM, on est amené à douter de l'appartenance de la théorie de Parfit au conséquentialisme de la règle. Pour dire les choses autrement, un certain nombre de points nous amènent à conclure que le conséquentialisme des raisons qu'il développe, notamment dans le troisième volume de OWM, n'est pas compatible avec le conséquentialisme de la règle qu'il semble vouloir adopter.

⁷ C'est, en fait, inexact. L'utilitarisme des préférences, notamment, permet de surpondérer l'utilité de ceux qu'on aime. Parfit raisonne presque toujours implicitement dans un cadre hédoniste ainsi que le montre l'exemple du canot de sauvetage qu'il utilise dans ce cas précis.

⁸ A la fin du troisième volume de OWM, Parfit suggère un quatrième élément qu'il ne développe pas suffisamment. Le conséquentialisme de la règle se renforcerait si l'on pouvait montrer qu'il est conforme à 'la meilleure version' de la Moralité du Sens Commun (tout comme cette dernière serait également renforcée si l'on pouvait montrer qu'elle est conforme au conséquentialisme de la règle) (OWM, III, 433).

On peut s'interroger, tout d'abord, sur la compatibilité entre l'externalisme assumé de Parfit et la théorie des motivations sous-jacente au conséquentialisme de la règle. S'il existe des raisons objectives, et que celles-ci sont suffisantes pour déterminer l'action de l'agent, le conséquentialisme de l'acte devrait, en toute logique, se suffire à lui-même. Prenons le problème du réchauffement climatique auquel Parfit attachait une importance toute particulière. Un conséquentialiste de l'acte considèrera qu'il a le devoir d'agir de manière à éviter autant que peut se faire la destruction des grands équilibres naturels qui constituent un danger pour les générations à venir ainsi que pour toutes les espèces vivantes. Toutes les actions qui satisfont ce critère sont bonnes et, selon la théorie externaliste, cette raison devrait suffire à motiver l'agent à agir dans ce sens. Le comportement des autres, ainsi que leurs croyances, ne rentre pas en ligne de compte. Selon le conséquentialisme de la règle, en revanche, les agents n'ont intérêt à respecter l'environnement que si les autres le respectent également. C'est la raison pour laquelle, selon ces théoriciens, la participation au vote est aussi faible dans les grands pays démocratiques. Ils proposent donc de substituer à la raison individuelle 'il est bon de voter' la raison collective 'il est bon que tout le monde vote'. C'est donc l'obéissance à la règle qui est ici bonne en soi (plus exactement l'obéissance à une règle que l'on estime être obéie par les autres), et non l'action de voter (c'est-à-dire l'action qui consiste à respecter la règle du vote indépendamment du faire savoir, ou de croire, si elle est respectée par les autres). Est-ce une position compatible avec l'externalisme pour lequel nos raisons nous sont fournies par les propriétés objectives du monde indépendamment des désirs et des croyances individuelles ? Je ne le pense pas.

Le conséquentialisme de la règle juge, ensuite, de la valeur de la règle uniquement en fonction de sa contribution à la maximisation du bien total. Il est bon de dire la vérité car on peut raisonnablement penser que dire la vérité est, en général, une règle optimifque. Supposons un instant, cependant, que nous vivions dans une société qui valorise les 'fake news' et que l'on attache plus d'importance à ce que l'on veut croire qu'à ce qui est vrai. On aurait alors tort, dans ce contexte, de dire la vérité. Cette position n'est, cependant, pas conforme à la pensée de Parfit, pour lequel des raisons déontiques peuvent toujours contre-balancer des raisons non-déontiques. Ainsi, explique Parfit, « on peut raisonnablement penser que, dans de nombreux cas, il serait mal en soi de traiter les gens d'une certaine façon, telle que de les tromper ou de les contraindre, ou en violant les promesses qu'on leur a faites. Le mal intrinsèque de telles actions surpasserait souvent le bien de leurs conséquences» (OWM, III, 421). Il ne peut donc exister pour Parfit de règles optimifques fondamentalement contraires aux raisons déontiques (même si des conflits occasionnels sont possibles). Néanmoins, rien, en théorie, n'interdit au conséquentialisme de la règle de valoriser une règle 'immorale' si celle-ci est optimisatrice.

Parfit, enfin, loue le conséquentialisme de la règle pour la latitude qu'il donne aux individus de préférer leurs proches, et notamment, leurs enfants aux étrangers lorsque le conséquentialisme de l'acte nous obligerait à respecter le principe de neutralité par

rapport à l'agence. L'argument est, ainsi que je le souligne dans une note, contestable. L'utilitarisme des préférences peut donner plus ou moins de poids à l'utilité des personnes selon leur degré de proximité avec l'agent. Dworkin a d'ailleurs vu dans cette possibilité une entorse inacceptable au principe d'égalité (Dworkin 1977). Dans le conséquentialisme de l'acte, il ne s'agit que d'une possibilité. Même si les agents ont des raisons 'objectives' de préférer leurs enfants à ceux qui leur sont étrangers, ils n'ont pas l'obligation de systématiquement le faire. Ils ont, en effet, également des raisons objectives d'aider de parfaits inconnus. La force respective de ces raisons dépend du bien produit par celles-ci. Il est parfois préférable de donner à un pauvre inconnu qu'à son enfant qui ne manque de rien. Or, ce qui est contingent dans le conséquentialisme de l'acte devient absolu dans le conséquentialisme de la règle. Si la règle 'préférer son enfant aux autres' est optimifque, elle doit alors être systématiquement respectée. Parfit accepterait-il cette conclusion ? L'ensemble de son œuvre, ainsi que son engagement pour l'altruisme effectif, prouve que non.

Pour terminer, on peut donc conclure que le conséquentialisme des raisons que présente Parfit dans le troisième volume de OWM n'est pas réductible au conséquentialisme de la règle qu'il défend dans le premier volume et qu'il réaffirme à la fin du troisième volume. On peut le regretter. Le conséquentialisme des raisons est beaucoup plus riche de possibilités et théoriquement beaucoup plus innovant que le conséquentialisme de la règle, largement décrédibilisé après les critiques de David Lyons.

5. Conclusion

Nous avons montré ci-dessus que Parfit défendait un conséquentialisme des raisons que l'on aurait tort de confondre avec un conséquentialisme de la règle. Cette conclusion met à mal projet d'unification des philosophies éthiques défendu dans OWM. Selon la Triple Théorie de Parfit, la déontologie kantienne, le contractualisme scanlonien et le conséquentialisme de la règle ne sont, en réalité, que les trois approches différentes d'une même réalité morale, à l'instar d'une montagne que l'on peut grimper sur différents versants. Son interprétation de Kant et de ses lois universelles, qu'il compare à un conséquentialisme de la règle, a été beaucoup critiquée (Otsuka 2009). Thomas Scanlon a lui-même pris certaines distances vis-à-vis de l'interprétation de son œuvre par son ami (OWM, II). Dans les deux cas, on lui a reproché d'avoir exagéré les similitudes existant entre ces deux courants de pensée et le conséquentialisme de la règle. Le lien entre conséquentialisme et conséquentialisme de la règle avait jusqu'ici fait l'objet de peu de discussion. Peut-être a-t-on pensé qu'il n'y avait pas là matière à débat. Pourtant, ainsi que j'ai essayé de le montrer, son interprétation du conséquentialisme diffère fondamentalement du conséquentialisme de la règle qu'il dit vouloir adopter. Ce résultat vient donc conforter l'idée selon laquelle la Triple Théorie est erronée. Cela doit-il nous conduire à douter de l'intérêt de la philosophie de Parfit ? Absolument pas. Ses

analyses restent d'une très grande richesse et contribueront certainement à renouveler le débat éthique. Le fait qu'elles convergent ou non vers une théorie unique n'est après tout pas ce qui importe. Ce qui importe vraiment, *what matters*, est qu'il développe une nouvelle forme de conséquentialisme, beaucoup plus flexible et moins contraignante que le conséquentialisme de la règle, et que celle-ci a un potentiel théorique et pratique⁹ beaucoup plus important que celui du conséquentialisme de la règle.

Daniels, Norman. 1979. "Moral Theory and the Plasticity of Persons." *The Monist* 62 (3):265-287.

Dworkin, R. 1977. *Taking Rights Seriously*. Cambridge (Ma): Harvard University Press.

Korsgaard, C.M. 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press.

Lazari-Radek, Katarzyna de, and Peter Singer. 2017. "Parfit on Objectivity and 'The Profoundest Problem of Ethics'." In *Does Anything Really Matter? Essays on Parfit on Objectivity*, edited by Peter Singer, 279-296. Oxford: Oxford University Press.

Nagel, T. 1989. *The View From Nowhere*: Oxford University Press, USA.

Otsuka, M. 2009. "The Kantian Argument for Consequentialism". *Ratio*, 22 (1). pp. 41-58.

Parfit, D. 2016. "Conflicting Reasons." *Etica & Politica / Ethics & Politics* XVIII (1):169-186.

Parfit, Derek. 1982. "Personal Identity and Rationality." *Synthese* 53.

Scanlon, T.M. 2014. *Being Realistic about Reasons*. Oxford: Oxford University Press.

Scanlon, T.M. . 1998. *What We Owe to Each Other*. paperback edition 2000 ed. Cambridge, Massachusetts and London, England: The Belknap Press of Harvard University Press.

Scheffler, Samuel. 1982. "Ethics, Personal Identity, and Ideals of the Person." *Canadian Journal of Philosophy* 12 (2):229-246.

Schultz, Bart. 1986. "Persons, Selves, and Utilitarianism." *Ethics* 96 (4):721-745. doi: 10.2307/2381096.

Sidgwick, Henry. 1981. *The Methods of Ethics*. Seventh Edition ed. Indianapolis/Cambridge: Hackett Publishing Company.

Williams, Bernard. 1981. "Internal and External Reasons." In *Moral Luck: Philosophical Papers*, 20-39. Cambridge: Cambridge University Press.

⁹ Pour une tentative d'application de cette théorie voir C. Salvat, « L'éthique face à la GPA », à paraître.