Aggregation and the Separateness of Persons

IWAO HIROSE
McGill University

Many critics of utilitarianism claim that we should reject interpersonal aggregation *because* aggregative principles do not take the separateness of persons seriously. In this article, I will reject this claim. I will first elucidate the theoretical structure of aggregation. I will then consider various interpretations of the notion of the separateness of persons and clarify what exactly those critics are trying to reject by appealing to the notion of the separateness of persons. I will argue that none of these interpretations can serve as the ground for rejecting aggregation.

I. INTRODUCTION

According to some moral theories, the gains and losses of different individuals can be combined and balanced in order to judge the rightness or wrongness of an act. I call this *interpersonal aggregation*, or *aggregation* for short. Aggregative moral theories hold that the disvalue of imposing heavy burdens on some limited number of people can be justified by the fact that this brings benefits to others, no matter how small these benefits may be, as long as the recipients are sufficiently numerous.[1] Critics of utilitarianism are highly sceptical of aggregation. For them, to say that aggregation should be permitted is to say that we should carve up a healthy person and transplant his organs to five needy patients because this makes the outcome strictly better. Critics find such a moral judgement to be unacceptable and therefore rule out aggregation from their proposed moral theories.

Criticism of aggregation, however, appears to go too far. Consider John Taurek's much-discussed *Rescue Case*, where we are faced with the choice between (a) rescuing one stranger and letting five different strangers die and (b) rescuing the five strangers and letting the one stranger die.[2] Suppose further that there are no morally relevant differences between these six strangers. Taurek claims that we should flip a fair coin to decide either (a) or (b). He believes that there is no impersonal perspective from which the loss of five lives is judged to be worse than the loss of one life, and that we should show equal

[1] There is another type of aggregation, i.e. *intrapersonal* aggregation. It concerns how the temporal parts of a person's life are combined within his life. I do not consider this type of aggregation here.

[2] J. Taurek, 'Should the Numbers Count?', *Philosophy and Public Affairs* 6 (1977), pp. 293–316.

© Cambridge University Press 2013
Utilitas Vol. 25, No. 2, June 2013
doi:10.1017/S0953820812000386

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

and positive respect to each stranger's life. According to Taurek, the coin-flipping best captures his belief because it gives the same chance (i.e. 50 per cent) of being saved to each stranger. But this appears counterintuitive to many people, including critics of utilitarianism. Intuitively, it is right to choose (b), i.e. to save the greater number of individuals. However, it seems that the case for saving the greater number cannot be justified without an appeal to aggregation. A problem arises for critics of utilitarianism. How can they justify the case for saving the greater number without aggregation? There have been several attempts to establish it. But none of them is satisfactory.[3]

It is not clear who first manifested criticism of aggregation explicitly and when. It is widely conceived, however, that criticism of aggregation has been an integral part of non-utilitarian moral theory at least since John Rawls published *A Theory of Justice*, in which he criticizes both classical and average utilitarianism. He puts forward two principles of justice, neither of which includes aggregation. Since then, other non-utilitarian philosophers such as Thomas Nagel, Robert Nozick and Tim Scanlon have emphasized that their proposed moral theories are non-aggregative.[4] In motivating his principles of justice, Rawls famously claims that classical utilitarianism does not take the separateness of persons seriously. Other critics of utilitarianism follow Rawls's claim. Thus, it is likely that the original argument against aggregation will be found in the idea of the separateness of persons.

In this article, I will examine the following specific question: does the notion of the separateness of persons lead us to non-aggregative moral principles? To put it another way, must we reject aggregation if we take the separateness of persons seriously? In this article, I will put forward a negative answer to this question. In section II, I will first elucidate the basic structure of aggregation before taking up the notion of the separateness of persons. In section III, I will consider with which property of aggregation the idea of the separateness of persons is concerned. I will claim that it is best understood as a criticism of aggregation. In section IV, I will raise two objections to the view that aggregative principles undermine the separateness of persons. In section V, I will consider and reject the view that the notion of the separateness of persons leads us to a version of contractualism (such as Rawls's) and hence to a non-aggregative principle. In section VI, I

[3] For the survey of this literature, see I. Hirose, 'Aggregation and Non-Utilitarian Moral Theories', *Journal of Moral Philosophy* 4 (2007), pp. 273–84.

[4] See T. Nagel, 'Equality', in his *Mortal Questions* (New York, 1979), pp. 106–27, at 125; R. Nozick, *Anarchy, State, and Utopia* (Oxford, 1974), p. 33; and T. M. Scanlon, 'Contractualism and Utilitarianism', *Utilitarianism and Beyond*, ed. A. K. Sen and B. A. O. Williams (Cambridge, 1982), pp. 103–28, at 123.

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

will consider and reject the view that the notion of the separateness of persons leads us to another version of contractualism (such as Scanlon's) and hence to a non-aggregative principle. I will conclude that the notion of the separateness of persons cannot serve as the ground for rejecting aggregation.

II. AGGREGATION DEFINED

By interpersonal aggregation, I mean, roughly, the combination of different people's goods, happiness, losses, well-being, and so on, into an objective value. By objective value, I mean the value on a numerical scale that represents the relation between sets of morally relevant features. Here is a more precise definition. Aggregation combines the values of separate morally relevant features into a single value in order to determine a comparative relation between non-singleton sets (a set containing more than two elements) of these features. Suppose that there are two sets, each of which contains n features. The relation between the two sets is as follows:

> *Relation*: $(a_1, a_2, \ldots, a_n)$ is at least as F as $(a'_1, a'_2, \ldots, a'_n)$, where F is some normative property, e.g. good, morally important, morally justifiable, morally weighty, and so on.

Being F-er is a strict relation which holds between two sets, e.g. better, more morally important, more morally justifiable, morally weightier, and so forth. For simplicity, let us assume that the F-er relation is reflexive, transitive and complete[5] and hence that the F-er relation constitutes an F-ordering. The features in each set can be good, well-being, happiness, claims, losses, pains, and so on. Aggregation combines these separate features into a single value. This is best understood by introducing a real valued function, which is a numerical representation of an ordering. The relation of the two sets is represented by an increasing function $f()$, and this defines the notion of aggregation.

> *Aggregation*: $(a_1, a_2, \ldots, a_n)$ is at least as F as $(a'_1, a'_2, \ldots, a'_n)$ if and only if $f(a_1, a_2, \ldots, a_n) \geq f(a'_1, a'_2, \ldots, a'_n)$, where $f()$ is symmetric and strictly increasing in its arguments.

When the subscripted indices above indicate persons, we have interpersonal aggregation. Thus, interpersonal aggregation is the combination of different people's morally relevant features (i.e. good,

[5] A relation F on a set X is reflexive just in case, for all x in X, x is at least as F as x. It is transitive just in case, for all x, y and z in X, if x is at least as F as y and if y is at least as F as z, then x is at least as F as z. It is complete just in case, for all x and y in X, either that x is at least as F as y or that y is at least as F as x. I assume that the set of alternatives is countable.

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

well-being, happiness, claims, loss, pain, and so on) into a real value, which represents the relation between two sets. Function $f()$ should be increasing in its arguments because, other things being equal, the increase of one feature should make the value of the combined features strictly greater.

Let me clarify what I mean by gain and loss for a person. A gain for person i occurs when $a_i > a'_i$ and a_i is brought about. A loss for person i occurs when $a_i < a'_i$ and a_i is brought about. Gain and loss are defined in terms of the difference between two states. Aggregation is concerned with the moral trade-off between the gains and losses for different groups of individuals. According to aggregative moral principles, the gains for a group of individuals can *morally* outweigh the losses for a different group of individuals. When there is no conflict between different individuals, there is no need to appeal to aggregation. If one alternative benefits some person and harms no other person, it would be agreed unanimously that this alternative should be chosen. Aggregation is used when and because there is a conflict of interests between different groups of individuals. Those opposed to aggregation hold that the combined gains for one group do not morally outweigh the losses to another group or individual. When there is a conflict of interests, non-aggregative principles resolve the conflict without combining the gains and losses of different individuals.

To understand the theoretical structure of aggregation, let me now consider the basic properties of aggregation. Interpersonal aggregation consists of four basic properties. The first property is that it requires the comparability of gains and losses in the morally relevant features of different people. I call this *interpersonal comparability*.

> *Interpersonal comparability*: For every pair of persons, it is possible to compare the morally relevant feature of one person with the morally relevant feature of the other.

It is easy to see that we cannot combine and balance the gains and losses of different individuals unless we are able to compare them. There are different types of interpersonal comparability. The type of interpersonal comparability depends on the chosen distributive principle. Take classical utilitarianism. Classical utilitarianism requires *unit-comparability* but does not necessarily require *level-comparability*. That is, for utilitarian aggregation, it suffices to compare the gains and losses for different people, and information concerning the level of different people's well-being is unnecessary. To illustrate this point, consider two comparisons. The first comparison is between $x = (10, 15)$ and $y = (11, 13)$. The second comparison is between $x' = (110, 115)$ and $y' = (111, 113)$. Classical utilitarianism judges that x is strictly better than y, and that x' is strictly better than y'. These

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

judgements can be made by the same reasoning. That is, the loss of 1 unit for person 1 is smaller than the gain of 2 units for person 2. In these two comparisons, the gains and losses for each person are the same, although their absolute level in the second comparison is higher by 100 units than in the first comparison.[6]

The second property is what I call *impartiality*.

> *Impartiality*: Other things being constant, two alternatives are morally indifferent if they differ only with regard to the identities of people.

The basic idea of impartiality is that, given other things being equal, the permutations of personal identities do not affect our moral judgements. To see the gist of this property, consider a two-person case with two alternatives x = (losing an arm, losing no arm) and y = (losing no arm, losing an arm), where the brackets show what happens to persons 1 and 2. The only difference between x and y is that the personal identities are permutated. The other features remain the same. Impartiality holds that x is morally indifferent to y. For instance, it would be reasonable for person 1 to reject the principle that always chooses x because he is not shown equal respect.[7] In the aforementioned definition of aggregation, I assume a real-valued function that represents the relation between two sets of morally relevant features. The value of the set of morally relevant features is not altered even if people's positions are swapped. That is, $f(a_1, a_2, a_3, \ldots, a_n) = f(a_2, a_1, a_3, \ldots, a_n) = f(a_1, a_3, a_2, \ldots, a_n) = \ldots = f(a_n, a_{n\text{-}1}, \ldots, a_2, a_1)$. The notion of impartiality captures this basic feature of aggregation.

The third property is *Pareto*.[8]

[6] There is general scepticism about interpersonal comparability in both economics and philosophy. In economics, it is claimed that there is no scientific or objective basis for interpersonal comparisons of utility as utility represents a person's mental state. In ethics, some people are sceptical about interpersonal comparison because of incommensurability of values. However, I believe that those who advocate the notion of the separateness of persons are not concerned with interpersonal incomparability or incommensurability.

[7] It is not my contention that this is the most plausible notion of impartiality. There might be other ways to define the notion of impartiality. All I claim is that aggregation entails the condition of impartiality I stated above. Needless to say, if there is a morally relevant difference between the two persons (e.g. one person is evil and the other is virtuous), we must consider additional elements such as desert, responsibility, and so on. In such a case, the notion of impartiality must be revised. However, the notion of impartiality holds that two alternatives are morally indifferent if they differ only with regard to the identities of people. Throughout this article, I assume that there are no morally relevant differences amongst people.

[8] Notice that this version of Pareto is different from the Pareto condition in economics. In economics, the Pareto condition is defined in terms of preference: if x is strictly preferred to y by some individuals, and if y is not strictly preferred to x by any individual, x is collectively preferred to y. Notice also that the Pareto condition is logically independent of the person-affecting view in ethics. The person-affecting view holds that x is better

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

Pareto: Other things being equal, if one alternative is strictly *F*-er for some person than another alternative, and if it is equally as *F* for other people, then it is strictly *F*-er than the other.

Simply put, Pareto holds that the gain for a person counts as strictly positive, and that moral judgements should be sensitive to the gain for this person as far as it is morally relevant. If a moral principle does not acknowledge the improvement of a person's condition, then there would be a legitimate ground for this person to object to this principle. In the general definition of aggregation, I take the function $f()$ to be strictly increasing in its arguments. This means that, given other people's morally relevant features remaining constant, an increase in one person's morally relevant feature results in an increase in the value of the set of that morally relevant feature. More precisely, $f(a_1, \ldots, a^*_i, \ldots, a_n) > f(a_1, \ldots, a_i, \ldots, a_n)$ if $a^*_i > a_i$. Pareto represents this feature of aggregation.

The fourth property is *continuity*. Continuity allows us to represent the *F*-ordering by a real-valued function $f()$.[9]

Continuity: For any set of morally relevant feature x, define $A(x) = \{y \mid y \text{ is at least as } F \text{ as } x\}$ and $B(x) = \{y \mid x \text{ is at least as } F \text{ as } y\}$. Then, $A(x)$ and $B(x)$ are closed (in other words, $A(a_i)$ and $B(a_i)$ contain their own boundaries, for any a_i in the set of morally relevant features).

When an ordering is continuous, it can be represented by a real-valued function. Continuity is like the Archimedean property of the real numbers. For any positive numbers x and y, there is a natural number n such that nx is greater than y. In ethics, it implies, intuitively, that for any amount of A there is some amount of B which is *F*-er.

Given the theoretical structure of aggregation, we can see why some distributive principles are non-aggregative. Take Maximin. Maximin ranks states of affairs, and the relative goodness is represented in such a way that a state of affairs $x = (w_1, w_2, \ldots, w_n)$ is at least as good as another state of affairs $y = (w'_1, w'_2, \ldots, w'_n)$ if and only if min $(w_1,$

than y only if x is better for some person than y. The Pareto condition holds that x is better than y if x is better for some person and no worse for any person than y.

[9] For the general discussion of continuity, see G. Arrhenius, 'Superiority in Value', *Philosophical Studies* 123 (2005), pp. 97–114; G. Arrhenius and W. Rabinowicz, 'Value and Unacceptable Risk', *Economics and Philosophy* 21 (2005), pp. 177–97; K. K. Jensen, 'Unacceptable Risks and the Continuity Axiom', manuscript (University of Copenhagen, 2010). Some philosophers reject continuity in some cases. The most famous one would be John Stuart Mill. In *Utilitarianism*, he distinguishes higher and lower pleasures, and contends that any amount of higher pleasure is better than any amount of lower pleasure. However, Mill does not reject continuity within higher pleasure or lower pleasure. He only rejects continuity between higher and lower pleasure. It would be accurate to say that Mill constrains continuity. For more extensive and precise discussion concerning Mill's idea of discontinuity in values, see G. Arrhenius and W. Rabinowicz, 'Millian Superiorities', *Utilitas* 17 (2005), pp. 127–46.

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

$w_2, \ldots, w_n) \geq \min (w'_1, w'_2, \ldots, w'_n)$, where w_i denotes the well-being of person i. Maximin satisfies interpersonal comparability (more precisely, interpersonal level-comparability), impartiality and continuity. But it violates Pareto. This is why Maximin is non-aggregative. Take another example, Leximin. Leximin first compares the worst-off across alternatives and, if the worst-off are at the same level, compares the second worst-off, and so on. More precisely, Leximin holds that for all $x = (w_1, w_2, \ldots, w_n)$ and $y = (w'_1, w'_2, \ldots, w'_n)$, if there exists a position k in $N = \{1, 2, 3, \ldots, n\}$ such that: (1) the well-being level in k is strictly higher in x than y; (2) the well-being level of every position $j<k$ is the same in x as in y, then x is strictly better than y. Otherwise, x and y are equally good. Leximin satisfies interpersonal comparability (more precisely, interpersonal level-comparability), impartiality and Pareto. But it is a discontinuous ordering. Leximin violates continuity and cannot be represented in functional form.[10] This is why Leximin is non-aggregative.

III. THE SEPARATENESS OF PERSONS

In *A Theory of Justice*, John Rawls criticizes both classical and average utilitarianism. In criticizing classical utilitarianism, he appeals to the notion of the separateness of persons. He criticizes average utilitarianism for different reasons.[11] I shall first consider what Rawls criticizes by invoking the notion of the separateness of persons.

According to Rawls, classical utilitarianism extends to society as a whole the principle of rational choice for a perfectly sympathetic, and ideally rational and impartial, spectator. The impartial spectator 'identifies with and experiences the desires of others as if these desires were his own' and organizes 'the desires of all persons into one coherent system of desire'. Within this system, 'there is no reason in principle why the greater gains for some should not compensate for the lesser losses of others; or more importantly, why the violation of the liberty of a few might not be made right by the greater good shared by many'. As 'many persons are fused into one', Rawls claims that '[u]tilitarianism does not take seriously the distinction between persons'.[12] This is the

[10] Note that the rejection of continuity does not entail Leximin. See E. Carlson, 'Aggregating Harms: Should We Kill to Avoid Headaches?', *Theoria* 66 (2000), pp. 246–55.

[11] Rawls appeals to psychological stability, self-respect and the strains of commitment when he criticizes average utilitarianism. See J. Rawls, *A Theory of Justice* (Cambridge, Mass., 1971), pp. 175–92. See also S. Scheffler, 'Rawls on Utilitarianism', *The Cambridge Companion to Rawls*, ed. S. Freeman (Cambridge, 2003), pp. 426–59.

[12] Rawls, *Justice*, pp. 26–7. For the general discussion on the separateness of persons, see D. Brink, 'The Separateness of Persons, Distributive Norms, and Moral Theory', *Value, Welfare, and Morality*, ed. C. Morris and R. Frey (New York, 1993), pp. 252–89; D. McKerlie, 'Egalitarianism and the Separateness of Persons', *Canadian Journal of*

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

separateness of persons objection to classical utilitarianism (hereafter, the *separateness objection* for short).

Following Rawls, other non-utilitarian philosophers appeal to the separateness of persons to criticize utilitarianism. For example, Nagel states,

> The conditions of choice corresponding to this [utilitarian] principle are that the chooser should treat the competing claims arising from distinct individuals as though they all arose from the interests of a single individual, himself. He is to choose on the assumption that all these lives are to be amalgamated into one life, his own. But this situation is unimaginable, and in so far as it is not, it completely distorts the nature of the completing claims, for it ignores the distinction between persons, as we observed before.[13]

The libertarian side-constraints theory is quite different from Rawlsian contractualism; however, Robert Nozick also joins the anti-utilitarian camp by appealing to the notion of the separateness of persons. He writes,

> There is no social entity with a good that undergoes some sacrifice for its own good. There are only individual people, different individual people, with their own individual lives. Using one of these people for the benefit of others, uses him and benefits the others. . . . To use a person in this way does not sufficiently respect and take account of the fact that he is a separate person, that his is the only life he has.[14]

Nozick appeals to the notion of the separateness of persons when he criticizes utilitarianism and attempts to constrain the domain of aggregation in social choice. He also appeals to the same notion when he criticizes Rawls's Difference Principle. Rawls's Difference Principle regards the distribution of natural talents as a common asset and allows the parties to share the benefit of other people's talents. According to Nozick, the pooling of people's talents does not take the separateness of persons seriously.[15] Thus, there is no consensus about the notion of the separateness of persons even among those who appeal to this notion in an important way.

Different people live different lives. This is a truism. There might be a metaphysical question whether different people really live different lives, but I do not think that Rawls and other critics of utilitarianism are concerned with this kind of question.

Philosophy 18 (1988), pp. 205–26; D. Parfit, *Reasons and Persons* (Oxford, 1984); J. Raz, *The Morality of Freedom* (Oxford: 1986); and Scheffler, 'Rawls'.

[13] T. Nagel, *The Possibility of Altruism* (Princeton, 1970), p. 138.

[14] Nozick, *Anarchy*, p. 33.

[15] Nozick, *Anarchy*, p. 228.

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

When Rawls appeals to the notion of the separateness of persons, he is concerned with something more substantive. But what exactly are his concerns?

The first interpretation is that the separateness objection is concerned with impartiality. On this interpretation, we should not fuse different people into one entity through the impartial spectator's point of view. According to Rawls's diagnosis, the impartial spectator's point of view disconnects the desire from the identity of its holder, and fuses the desires of many people into one coherent system of desire. It is the impartial perspective that undermines the separateness of persons. The separateness of persons is respected when and only when the impartial perspective is given up. If this interpretation is correct, the separateness objection would rule out aggregation because aggregation assumes impartiality.

As far as Rawls's theory of justice is concerned, this interpretation is incorrect for two reasons. First, when Rawls justifies his two principles of justice, he appeals to the rational decision-making of the parties in the original position. In the original position, the parties are deprived of information about their place in society, physical and psychological dispositions, and their own conception of the good. They only have a general knowledge about the basic laws of sociology, economics and psychology. They are supposed to choose the principles of justice independent of their actual identity. This means that the parties in the original position are supposedly impartial in the sense that the actual identities would not affect the parties' rational choice. If the separateness of persons hinges on impartiality, Rawls's justification for two principles of justice does not take the separateness of persons seriously. Thus, Rawls does not mean that the separateness objection aims at impartiality.

Second, Rawls's Difference Principle implies impartiality. Suppose that society consists of two mutually exclusive and jointly exhaustive groups. Compare two states of affairs, $x = (10, 5)$ and $y = (5, 10)$, where the brackets show the level of primary social goods of the two groups. The Difference Principle judges that x is indifferent to y, as impartiality requires. If the interpretation under consideration is correct, Rawls's Difference Principle does not take the separateness of persons seriously. For these two reasons, Rawls's separateness objection is not concerned with impartiality. Note that I am not attempting to defend the notion of impartiality here. All I claim is that we cannot criticize the notion of impartiality by invoking Rawls's notion of the separateness of persons because his proposed principles imply the same notion of impartiality.

The second interpretation is that the separateness objection is concerned with the combination of impartiality and Pareto. According

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

to this interpretation, impartiality and Pareto jointly dissolve the boundary between separate individuals, and reduce individuals to a mere means to maximize the goodness of a state of affairs. Suppose that we accept impartiality and Pareto. Compare two alternatives in the two-person case, $x =$ (losing an arm, losing no arm) and $z =$ (losing no arm, losing two arms). The combination of the two notions holds that we should choose x. To illuminate this judgement, suppose another possible state $z' =$ (losing no arm, losing an arm). By Pareto, when we compare z and z', z' should be chosen. By impartiality, x is indifferent to z'. Consequently, the combination of impartiality and Pareto judges that x should be chosen when we compare x and z.

This judgement may be challenged. It can be claimed that to judge that x should be chosen is to say that person 2 keeps both arms at the cost of an arm of person 1. There is a sense in which person 1 is used merely as a means to spare the two arms of person 2, or to maximize the goodness of the state of affairs. Arguably, this 'sense' picks out an instance where the separateness of persons is violated.[16]

On the face of it, this challenge seems to make sense. However, it does not make sense. First, even if x is chosen, this does not mean that person 1 is used *merely* as a means to benefit person 2. If there is a conflict of interests, one person will benefit and one or more persons will lose. To solve the conflict of interests is not to treat some person as a mere means. Second, if we choose z instead of x, there is also a sense that person 1 saves one arm at the cost of two arms of person 2, in which case person 2 is used merely as a means to benefit person 1. Therefore, there can always be a sense in which we treat one person as a means to improve the condition of the other, in so far as we choose either course of action. The only way of respecting the separateness of persons, according to this interpretation, is to choose neither.[17] That is, we do not help any person. If the separateness

[16] When Jens Timmermann and Weyma Lübbe reject the conjunction of impartiality and Pareto as a solution to the Rescue Case, they seem to appeal to this sense. See J. Timmermann, 'The Individualist Lottery: How People Count, but not their Numbers', *Analysis* 64 (2004), pp. 106–12; W. Lübbe, 'Taurek's No Worse Claim', *Philosophy and Public Affairs* 36 (2008), pp. 69–85.

[17] If the argument is that there is more goodness in x than y, then there is a sense in which we are indeed ignoring the separateness of persons. The mere fact of reasoning in terms of goodness and badness of states of affairs seems to fail to take the separateness of persons seriously. On the other hand, saying that x would be preferred to y from each person's point of view (assuming they did not know who they were) would not violate the separateness of persons in the same way. John Harsanyi and other proponents of contractualist average utilitarianism would be happy to agree with this. This means that contractualist average utilitarianism respects the separateness of persons, and therefore that the separateness of persons does not lead us to non-aggregative principles. See J. C. Harsanyi, 'Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility', *Journal of Political Economy* 63 (1955), pp. 309–21.

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

objection is concerned with the combination of impartiality and Pareto, we should make no moral judgement or action when there is a conflict of interests. The separateness of persons would tell us nothing in the cases where we must make judgements concerning how we distribute the burdens and benefits for different individuals. The scope of a moral principle that respects the separateness of persons in this way is extremely limited. This is not a desirable consequence. Thus, I do not believe that the separateness objection is concerned with the combination of impartiality and Pareto. It is concerned with a stronger notion: the combination of impartiality, Pareto and continuity – that is, aggregation.

The separateness objection seems to make sense only if it is understood as a criticism of aggregation. By combining the gains and losses of different people into the gain and loss of a group of people, it is claimed that we reduce different people's perspectives into a single impersonal perspective, thus disregarding the boundary between the different people. In putting forward the separateness objection,

Rawls directly objects to aggregation, saying that:

> [in classical utilitarianism] there is no reason in principle why the greatest balance of gains of some should not compensate for the lesser losses of others; or more importantly, why the violation of the liberty of a few might not be made right by the greater good shared by many.[18]

He proposes two principles of justice, but none of his principles includes the aggregative structure. Therefore, I believe that Rawls objects to aggregation in his separateness objection.

This is the case for other critics of classical utilitarianism as well. For example, Nagel points out that utilitarianism ignores the distinction between persons, and later puts forward what he calls *pairwise comparison*, which is not aggregative.[19] The pairwise comparison compares one person's gain and loss with another's in order to identify which alternative would be less unacceptable, and chooses the least unacceptable one from individual standpoints. It compares each person's gains and losses with another's, and, through the series of pairwise comparisons, identifies the alternative which is least unacceptable to everyone. Given that the comparison is made pairwise, it avoids comparisons between groups of individuals. If x is slightly better than y for a million people but y is considerably better than x for one very badly off person, the pairwise comparison claims that y should be chosen. The pairwise comparison is non-aggregative, and thus, it is claimed, takes the separateness of persons seriously. Likewise, Nozick

[18] Rawls, *Justice*, p. 26.
[19] Nagel, *Altruism*, p. 138 and 'Equality', p. 123.

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

claims that balancing the gains and losses of different people and imposing a burden on one person 'does not sufficiently respect and take account of the fact that he is a person, that his is the only life he has'.[20] His proposed side-constraint theory restricts the domain of aggregation, thus respecting the separateness of persons.

The separateness objection thus construed claims that we should rule out aggregation. But this claim goes far beyond what critics of classical utilitarianism were originally trying to achieve. That is, if we rule out aggregation, we must reject not only classical utilitarianism but also other aggregative principles, including some versions of egalitarian principles with aggregative structures.[21] According to the separateness objection, it would be argued that aggregative egalitarian principles fuse many people's well-being into one coherent system of well-being, and therefore that these version of egalitarianism do not take the separateness of persons seriously. This is a surprising consequence of the separateness objection. Some critics of aggregation must claim that the nature of some moral values such as equality is essentially non-aggregative, and that these values cannot be weighed directly against other values. For example, Nagel thinks that a correct egalitarian theory must be non-aggregative because aggregation undermines the foundation of egalitarianism.[22]

IV. DEFUSING THE SEPARATENESS OBJECTION TO AGGREGATION

If my interpretation is correct, according to the separateness objection, to invoke an aggregative principle is to undermine the separateness of persons. It can be claimed that non-aggregative principles respect the separateness of persons and hence that non-aggregative principles are more plausible than aggregative ones. I have two worries about this point.

First, non-aggregative principles may take the separateness of persons seriously, but it is not clear whether non-aggregative principles in fact take every person (as distinguished from the 'separateness' of persons) really seriously. Consider Maximin. According to Maximin, the relative goodness of states of affairs is judged by the relative level of the worst-off person in each state. This means that the judgement about the relative goodness of states of affairs is not affected by the

[20] Nozick, *Anarchy*, p. 33.

[21] Some egalitarian principles combine the value of people's well-being and the disvalue of inequality. For example, if the relative goodness of states of affairs in the two-person case is given by $\frac{1}{2}(w_1+w_2) - \frac{1}{4}|w_1 - w_2|$, where w_i denotes the person i's well-being, this principle is both aggregative and egalitarian.

[22] Nagel, 'Equality', p. 125.

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

gains and losses of the non-worst-off people. According to Maximin, the state of the non-worst-off does not affect moral judgements. Practically, information about the non-worst-off is redundant for the purpose of reaching a moral judgement. Maximin does not combine the gains and losses of different people. But practically, it takes the non-worst-off as irrelevant.

What about Leximin? Leximin gives complete priority to the lexically worse-off individual. Leximin satisfies impartiality, Pareto and interpersonal comparability. However, like Maximin, Leximin does not take many people seriously. Imagine a choice between saving the life of one person and sparing the arm of many people. Leximin judges that it is right to save the person's life, no matter how many people would lose an arm. According to Leximin, if an extra person was added to the group of arm-losers, his state would not affect our moral judgement at all. Thus, practically, Leximin takes many people as irrelevant.

On the other hand, aggregative principles can take every person seriously. If a person receives additional gains, then the value of combined gains increases. Aggregative principles show a positive consideration to the gains of every person, and negative consideration to the losses of every person. Every change for every person is acknowledged in terms of the value of combined gains and losses. Thus, aggregative principles take every person as relevant, although it could be claimed that they do not take the separateness of persons seriously.

Rawls's Difference Principle is slightly different, though it is also non-aggregative. Rawls could contend that there is a way in which his theory precisely does not ignore any group in society. In its first and second principles, Rawls's theory guarantees equal basic liberties to all as well as equal opportunities to all, and gives a lexical priority to these principles over the Difference Principle. Rawls could claim that these two lexically prior principles take every person seriously, even though the Difference Principle takes the state of the non-worst-off people as irrelevant. This interpretation, however, does not undermine my claim. Given that the two lexically prior principles are chosen by the parties in the original position, these two principles are supposed to respect the separateness of persons as well as take every person's interests seriously. If this is correct, why does Rawls need to hold the non-aggregative Difference Principle? He could hold an aggregative principle in so far as he gives a lexical priority to the other principles that respect the separateness of persons and every person's interest.

Nagel's pairwise comparison also takes many people as irrelevant. Compare $x = (10, 6, 6)$ and $y = (5, 8, 8)$. The pairwise comparison judges that x is less unacceptable than y from individual standpoints, and hence that it is right to choose x. This judgement will not change even if we add a large number of people, who are affected in the same

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

way as persons 2 and 3. For example, the pairwise comparison would choose $x' = (10, 6, 6, 6, 6, 6, 6)$ over $y' = (5, 8, 8, 8, 8, 8, 8)$. The presence of four extra people does not make any difference for what it is right to do. These extra people are irrelevant.

Nagel and his followers would claim that the pairwise comparison compares everyone with everyone else, and hence that it takes everyone into account. They could claim that the pairwise comparison takes every person seriously. I do not disagree with this interpretation. But I do not think that this interpretation shows that aggregative principles do not respect the separateness of persons. There is a sense in which the pairwise comparison does not take some people seriously, while there is also a sense in which it takes every person seriously. Likewise, there is a sense in which aggregative principles do not take the separateness of persons seriously, while there is also a sense in which they take the separateness of persons seriously, as I will argue later in this section. If it is claimed, from one particular sense, that the pairwise comparison takes every person seriously, it can also be claimed, from one particular sense, that aggregative principles do take the separateness of persons seriously.

Second, it is not clear whether aggregative principles do not really take the separateness of persons seriously in any sense. If aggregative principles respect the separateness of persons in a different sense from non-aggregative principles, we must consider which 'sense' is more plausible or important than others. Unless advocates of non-aggregative principles offer an argument for the superiority of their sense of the separateness of persons, the separateness objection cannot be used to refute aggregative principles. As I said earlier, anyone, whether utilitarian or non-utilitarian, would agree that different people live different lives. So, there is no disagreement as far as this fact is concerned. Critics of aggregation hold that the notion of the separateness of persons leads us to reject aggregation. But the separateness of persons may lead us to a different claim, which is not necessarily non-aggregative.

Consider egalitarianism. Parfit suggests one way in which the separateness of persons might lead us to egalitarianism. He writes: '[s]ince it is a deep truth that we live different lives, it is an ultimate moral aim that, in so far as we are equally deserving, the lives of each should go equally well'.[23] Some versions of egalitarianism do not necessarily exclude aggregation. Therefore, the separateness of persons might lead us to an aggregative egalitarian principle. Even classical utilitarianism would agree with the separateness of persons in another

[23] Parfit, *Reasons*, p. 330.

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

sense. Consider an aggregative principle, according to which the well-being of Annie and Betty jointly counts for one, and the well-being of everyone else counts for one. Proponents of classical utilitarianism would claim that Annie and Betty live separate lives, and hence that Annie's well-being and Betty's well-being should count separately. Classical utilitarianism would hold that Annie's well-being should not count together with Betty's well-being. On this interpretation, to give positive and equal weight to every person's well-being is to acknowledge the fact that everyone lives a different life from everyone else. Classical utilitarianism endorses the separateness of persons in this sense. It may be pointed out that my claim here is not about the separateness of persons, but equal respect or impartiality. This point does not undermine my claim here. Classical utilitarianism cannot give equal respect to, or be impartial between, Annie and Betty unless it acknowledges the fact that Annie and Betty live different lives.[24]

There is a difficult question regarding whether one interpretation is more substantive and plausible than other interpretations, but this question is broader than the question I have been considering. To ask whether interpretation A is more plausible than interpretation B is to ask whether the moral view underlying interpretation A is more plausible than the moral view underlying interpretation B. This question is concerned with the comparison of the competing moral views. I have been considering the logical relation between the separateness of persons and non-aggregative principles. In this section, I do not intend to answer the question of whether a moral view that rules out aggregation is more plausible than a moral view that permits aggregation. However, my analysis in this section rejects the claim that the moral view that rules out aggregation is more plausible than the moral view that permits aggregation *because* the latter does not take the separateness of persons seriously.

Any principle can accept the trivial fact that different people live different lives, and that each person's life should be respected separately. Critics of aggregation claim that aggregative principles violate the separateness of persons, and non-aggregative principles respect the separateness of persons. If the separateness of persons is invoked in this way, it is used to state a certain moral view, not to argue for it. The separateness of persons can be stated in other ways, as we saw in the cases of utilitarianism and a version of egalitarianism. In the separateness objection, the separateness of persons is merely used

[24] Notice that I am not claiming that the separateness of persons is best captured by a version of utilitarianism. All I claim is that there are many different interpretations concerning the separateness of persons, and that each interpretation is perfectly consistent with the idea of the separateness of persons in one way or another.

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

to state what critics of aggregation want to reject. But the separateness of persons is supposed to explain why we should rule out aggregation. The same notion can be used to state whatever we want to support or reject. Thus, the separateness of persons itself cannot be the grounds for rejecting aggregative principles. The rejection of aggregative principles does not follow from the idea of the separateness of persons.

V. THE SEPARATENESS OF PERSONS AND CONTRACTUALISM

Here is another interpretation: the separateness of persons is not concerned with a certain outcome, but with the acceptability of moral principles to those affected by the actions they permit and require. When there is a conflict of interests between different people, someone must bear some burden. It is not controversial that moral principles impose some burden on some people. However, moral principles must be acceptable to every person so that the imposed burden will be acceptable to any person no matter who bears that burden. Even when principles impose some burden on a particular person in an actual case, he would accept that burden if every person including himself agrees to those principles in a certain hypothetical situation. Imposing a burden in this way, it can be argued, takes every person's perspective separately. Therefore, the separateness of persons would require us to agree to a moral principle in a hypothetical situation; it would lead us to a version of contractualism.

When I speak of the acceptability of moral principles, I have one particular type of contractualism in mind, namely Rawls's. In Rawls's original position, the parties are supposed to place themselves under an informational constraint, the veil of ignorance. Behind the veil of ignorance the parties are deprived of information about their own particular situation. They know nothing about their positions in society, their physical and mental endowments, their conceptions of the good, or their psychological dispositions, including their attitude towards risk. They only know some general information, including the basic laws of sociology, economics and psychology. Under these conditions, the parties would rationally choose the two principles of justice Rawls defends. As every rational person is supposed to choose Rawls's principles of justice behind the veil of ignorance, every person would accept bearing some burden if it is what the two principles impose. The state of the non-worst-off people does not affect the judgement of the Difference Principle. But these non-worst-off people are supposed to have no legitimate complaint about how they are treated when the veil of ignorance is lifted. In so far as the moral principles are agreed upon by every person in the original position, they respect the separateness of

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

persons. Thus, the notion of the separateness of persons would lead us to contractualism in the original position.

The separateness of persons thus construed can rule out some aggregative moral principles. Take classical utilitarianism. If everyone agrees to choose classical utilitarianism in the hypothetical situation, the burden imposed by classical utilitarianism is acceptable to every person. However, Rawls rightly points out that the parties would *not* rationally choose classical utilitarianism in the original position. According to classical utilitarianism, other things being equal, when the number of persons in society is duplicated, the total utility is judged to be twice as great. However, according to Rawls, 'since the parties aim to advance their own interests, they have no desire in any event to maximize the sum total of satisfaction'.[25] For the parties in the original position, it would not be rational to choose classical utilitarianism. However, there is a 'surprising contrast'[26] between classical and average utilitarianism. According to Rawls, the parties in the original position may well choose average utilitarianism in the original position. Rawls accepts the possibility of average utilitarianism, and does not apply the separateness of persons objection to it.

According to John Harsanyi, who appeals to a similar method of justification to that of Rawls, the parties would rationally choose average utilitarianism in a hypothetical situation like Rawls's original position. In the hypothetical situation where each person does not know which person's position he would occupy, Harsanyi contends that a rational person would believe that he has an equal chance of being in any of these positions, and would rationally choose average utilitarianism. Thus, average utilitarianism is one of the moral principles that could be chosen by the rational parties. This implies that the contractualist version of average utilitarianism can indeed take the separateness of persons seriously. Rawls raises a set of objections to average utilitarianism. But these objections are independent of the idea of the separateness of persons.

There is a great deal of philosophical literature concerning the question of whether it would be rational for persons behind the veil of ignorance to choose Rawls's two principles of justice, or average utilitarianism, or something else.[27] I do not intend to answer this question because it seems that there is no conclusive argument that

[25] Rawls, *Justice*, p. 163.

[26] See Rawls, *Justice*, p. 189.

[27] For example, see K. J. Arrow, 'Some Ordinalist-Utilitarian Notes on Rawls's Theory of Justice', *Journal of Philosophy* 70 (1973), pp. 245–63; R. M. Hare, 'Rawls' Theory of Justice I', *Philosophical Quarterly* 23 (1973), pp. 144–55; and D. Lyons, 'Nature and Soundness of the Contract and Coherence Arguments', *Reading Rawls*, ed. N. Daniels (Cambridge, 1975), pp. 141–66.

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

supports either aggregative or non-aggregative principles as far as the choice of moral principles in the hypothetical situation is concerned. Let me explain.

Rawls contends that the parties behind the veil of ignorance would rationally choose Rawls's two principles of justice, which are non-aggregative. In contrast, Harsanyi contends that the parties would rationally choose average utilitarianism.[28] Rawls confines the notion of rationality to the standard notion of rationality in economics:

> the concept of rationality must be interpreted as far as possible in the narrow sense, standard in economic theory, of taking the most effective means to given ends. I shall modify this concept to some extent but one must try to avoid introducing into it any controversial ethical elements.[29]

Harsanyi also confines it to the basic axioms of Bayesian decision theory. In the original position, self-interested parties are supposed to have no information concerning their actual position and conception of good in the actual world. They only know the basic laws of economics, psychology and sociology. Harsanyi appeals to a similar hypothetical choice situation where the self-interested parties have no information concerning their actual situation. However, Rawls and Harsanyi disagree radically concerning the choice of basic principles that govern society.

Let us focus on the difference between the Difference Principle and average utilitarianism. If the Difference Principle is understood as Maximin applied to the bundle of primary social goods that representative members of social groups would possess, the difference is whether or not the rationally chosen principle satisfies Pareto. On this interpretation, the Difference Principle does not satisfy Pareto, whereas average utilitarianism does. Pareto is not a property of rationality or consistency. However, it is a property that is believed to be desirable for general rules. Rawls claims that it is not a desirable

[28] This is too quick. Harsanyi made two claims. The first claim is that if individual and social preferences satisfy the axioms of expected utility theory and the social welfare function satisfies the Pareto condition, then social utility must be represented by a weighted sum (precisely, an affine transformation) of individual utilities. This is known as the *Aggregation Theorem*. The second claim is that if individuals were placed behind the veil of ignorance, then they would rationally believe that each of them has an equal chance of being in any of these positions, and hence that the weight of each individual's utility must be $1/n$ in the n-person society. This is known as the *Impartial Observer Theorem*. The notion of the veil of ignorance is used for the second claim. See Harsanyi, 'Cardinal', and J. C. Harsanyi, *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations* (Cambridge, 1977). See also J. A. Weymark, 'A Reconsideration of the Harsanyi–Sen Debate on Utilitarianism', *Interpersonal Comparisons of Well-being*, ed. J. Elster and J. E. Roemer (Cambridge, 1991), pp. 255–320.

[29] Rawls, *Justice*, p. 14.

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

property, whereas Harsanyi believes that it is. But it is not a matter of rationality whether a Paretian or non-Paretian principle is chosen.

If the Difference Principle is understood as Leximin applied to the bundle of primary social goods that representative social groups possess, one difference between the Difference Principle and average utilitarianism lies in whether the principle satisfies continuity or not. The Difference Principle does not satisfy continuity, whereas average utilitarianism does. Again, continuity is not a matter of rationality. It is not irrational for self-interested parties to choose a discontinuous principle. Nor is it irrational to choose a continuous principle. Continuity is a property that is believed to be desirable for general rules. Rawls would claim that it is not desirable, whereas Harsanyi would claim that it is. Thus, either way, the difference between the Difference Principle and average utilitarianism cannot be explained by the rational choice of self-interested parties behind the veil of ignorance. This is why I believe that we cannot tell what self-interested parties would rationally choose behind the veil of ignorance.

Some people would appeal to the literature of experimental economics that focuses on what subjects *actually* (as distinguished from *rationally*) choose in the controlled laboratory. Many results in experimental economics suggest that very few subjects, who are placed in the situation similar to Rawls's original position, actually choose non-aggregative principles such as Rawls's.[30] I do not believe, however, that these experimental results establish a case against non-aggregative principles. This is mostly because, as Konow rightly observes, '[p]assing through the laboratory door is not necessarily equivalent to passing through a veil of ignorance, and previously formed knowledge and expectations might taint subjects' reasoning'.[31] Thus, there is no definite answer to the question of whether the parties would rationally choose aggregative principles or non-aggregative principles. For present purposes, however, it suffices to point out that there are several reasonable theories that support aggregative principles as the principles chosen in the original position, and that aggregative principles are not ruled out in the original position.

Here I am not trying to prove that aggregative principles *must* be chosen behind the veil of ignorance. My claim is not that the

[30] See N. Frohlich, J. A. Oppenheimer and C. L. Eavey, 'Laboratory Results on Rawls's Distributive Justice', *British Journal of Political Science* 17 (1987), pp. 1–21; N. Frohlich and J. A. Oppenheimer, *Choosing Justice: An Experimental Approach to Ethical Theory* (Berkeley, 1993); and J. Konow, 'Which Is the Fairest One of All? A Positive Analysis of Justice Theories', *Journal of Economic Literature* 41 (2003), pp. 1188–1239.

[31] Konow, 'Fairest', p. 1196.

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

separateness of persons will lead us to aggregative principles. Nor is it that non-aggregative principles will never be chosen behind the veil of ignorance. All I claim is that the separateness of persons does not definitely lead us to aggregative principles or non-aggregative principles. Critics of aggregation contend that the separateness of persons must lead us to non-aggregative principles. I merely attempt to defuse the supposed logical connection between the separateness of persons and the rejection of aggregation, which critics of aggregation have assumed. Critics of aggregation must establish that the principles chosen behind the veil are always non-aggregative, and that non-aggregative principles will never be chosen. Ironically, Rawls did not claim this much, as we saw earlier in this section. Clearly, the burden of proof is on critics of aggregation. Even if the idea of the separateness of persons is concerned with the acceptability of moral principles to those affected by the actions they would permit and require, non-aggregative principles do not follow from the separateness of persons. The separateness of persons can lead us to many principles, including aggregative ones. Therefore, the separateness of persons cannot be the grounds for rejecting aggregation.

VI. SCANLON'S CONTRACTUALISM

Before I conclude this article, let me briefly discuss the possibility of another version of contractualism that Tim Scanlon puts forward. Scanlon objects to the Rawls–Harsanyi version of contractualism and proposes his own version. He believes that the justification of moral principles should be made from each person's actual perspective, and that the Rawls–Harsanyi version is not the right way of capturing the justifiability of moral principles. He further believes that his proposed version of contractualism is non-aggregative. He objects to utilitarianism, but he does not appeal, at least explicitly, to the notion of the separateness of persons. Here is a possible way of bridging the logical gap between the separateness of persons and non-aggregative principles. If (a) the separateness of persons is concerned with the justifiability of moral principles from each person's perspective, (b) the correct version of contractualism is not the Rawls–Harsanyi version but Scanlon's version, i.e. moral principles must be justified from every person's actual perspective, and (c) only non-aggregative principles can be justified from every person's actual perspective (as Scanlon contends), then we can conclude that the separateness of persons leads us to non-aggregative principles. I have argued for (a) in the last section. I am neutral about (b). In this section, I will focus on (c).

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

According to Scanlon, the following reasoning underlies the Rawls–Harsanyi version of contractualism.[32] In order to think of a principle as a candidate for unanimous agreement, I must think of it not merely as acceptable to me but as acceptable to others as well. My judgement that the principle is acceptable must be impartial. To judge impartially that a principle is acceptable is to judge that it is one that I would have reason to accept no matter who I was (i.e. to judge that it is a principle which it would be rational to accept if I did not know which person's position I would occupy). The justification of principles is made from the point of view of a single rational individual in the hypothetical situation.

Scanlon points out that two distinct types of reasoning are conflated in the Rawls–Harsanyi contractualism. The first reasoning is (P) that the single individual is held to accept a principle because he judges that it is one he could not reasonably reject whatever position he turns out to occupy. The second reasoning is (Q) that this principle is supposed to be acceptable to a person in any social position because it would be the rational choice for a single self-interested person in the hypothetical situation. According to Scanlon, Rawls and Harsanyi start with (P), but implicitly move to (Q). Whenever we adopt the Rawls–Harsanyi version of contractualism (i.e. whenever we distinguish the justification of moral principle from the identification of moral principles), we have a 'subtle transition' from (P) to (Q). But Scanlon believes that this transition is not plausible, and that we should appeal only to (P).

To elaborate, suppose that P is a principle that it would be rational for a self-interested person to choose in a hypothetical situation. Does it follow that no one could reasonably reject P? Scanlon says 'no'. For example, compare $A = (12, 2)$ and $E = (8, 4)$. In this case, Scanlon claims that person 2 has a reasonable ground for complaint against the principle that would choose A, e.g. average utilitarianism. If this is correct, average utilitarianism would not be justifiable to person 2. Consider another example: compare $A' = (120, 2)$ and $E = (8, 4)$. Scanlon would claim that person 1 has a reasonable ground for complaint against the principle that would choose E, e.g. Maximin. If this is correct, Maximin is not justifiable to person 1. Thus, Scanlon believes that a principle that would be chosen by the rational parties in the hypothetical situation may turn out to be unjustifiable to some people in the actual situation. Scanlon contends that the justification of moral principles must be made from every person's actual perspective.

[32] Scanlon, 'Contractualism', pp. 120–23.

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

Scanlon wants to appeal to (P), and grounds rightness and wrongness in an act's being justifiable to actual individuals. His proposed contractualism is stated as follows.

> An act is wrong if its performance under the circumstances would be disallowed by any system of rules for the general regulation of behaviour which no one could reasonably reject as a basis for informed, unforced general agreement.[33]

His contractualism provides an account of what makes a certain act wrong and explains why such an act is wrong. The identification of moral principles and the justification of moral principles are compounded in one formula. This prevents Scanlon's contractualism from moving from (P) to (Q).

Scanlon claims that this version of contractualism is non-aggregative because principles can be rejected from 'individual standpoints'. If even one person has a legitimate reason to reject a principle, the principle cannot be used as a basis for judging the rightness or wrongness of an act, no matter how many people would benefit from adopting this principle. Scanlon calls this feature the *individualist restriction*, and believes that it prevents his contractualism from opening the door to aggregation.

The following question arises: does the individualist restriction rule out aggregation? To put it another way, are non-aggregative principles the only principles that no one could reasonably reject from individual standpoints? Even if we agree to the idea of the individualist restriction, it is not clear whether aggregative principles will *always* be rejected from individual standpoints. The basic idea of the individualist restriction is very modest: the supposed task of the individualist restriction is to warrant a sort of veto power to individuals over certain principles in certain contexts. It can constrain cases where aggregative principles can be the basis for judging the rightness or wrongness of an act. It is not supposed to rule out any particular type of principle. That is, it is not supposed to rule out all aggregative principles.

In order for Scanlon to say that his contractualism is non-aggregative, he must establish that all aggregative principles are rejected by someone in every context. But this does not seem to be what Scanlon wants to claim. Consider what Scanlon says about the Rescue Case. He believes that his contractualism can offer the case for saving the greater number of strangers without an appeal to aggregation. Suppose that we are to choose either saving A or saving B and C. Scanlon's argument goes as follows. If there was only one person on each side,

[33] Scanlon, 'Contractualism', p. 110.

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

tossing a fair coin might be the principle that no one could reasonably reject from individual standpoints. But if the presence of C's claim made no difference to how we are required to go about deciding what to do, this implies that C's claim to being saved would have no moral significance and hence that C would not be shown a positive concern. Therefore, C could reasonably reject the principle of tossing a coin. According to Scanlon, the principle of saving the greater number is the one that no one could reasonably reject. However, Michael Otsuka points out that Scanlon's argument appeals to a sort of aggregation.[34] This is because C's claim tips balance in favour of B and C only when C's claim is presented together with B's claim, and this can be seen as comparison between different groups of individuals. Thus, it is not clear whether Scanlon's contractualism could successfully avoid a counterintuitive claim like Taurek's coin-toss without appealing to some sort of aggregation. In any case, Scanlon later concedes that the gains and losses of different individuals can be aggregated within a broadly defined moral category, saying that 'it could be permissible or even required to forgo saving one person's life in order to prevent many people from becoming totally paralyzed'.[35] Of course, there is the general question of when an aggregative principle satisfies the individualist restriction and when it fails. But this is a theoretical problem for Scanlon's contractualism.

For the sake of argument, I can agree that the individualist restriction can constrain aggregation in some contexts. But it is perfectly consistent to say that some aggregative principles are allowed in other contexts. It is one thing to say that, given the individualist restriction, the process of justifying moral principles is not aggregative. It is quite another thing to say that the principles that would pass the non-rejectability test are always non-aggregative. I have pointed out that Scanlon need not commit to the second claim. I believe that Scanlon should support only the first claim. But this first claim would also be supported by Rawls and Harsanyi. This is because the process of justifying moral principles is not aggregative in their version of contractualism as well, although Harsanyi contends that the justifiable principle is aggregative. Thus, the principles that would meet the demands of Scanlon's contractualism need not be non-aggregative.[36]

[34] M. Otsuka, 'Scanlon and the Claims of the Many versus the One', *Analysis* 60 (2000), pp. 288–93.

[35] T. M. Scanlon, 'Replies', *Social Theory and Practice* 28 (2002), pp. 337–58, at 356.

[36] Parfit makes a stronger claim. He argues that Scanlon should drop the individualist restriction because it limits the scope of his contractualism. D. Parfit, 'Justifiability to Each Person', *Ratio* 16 (2003), pp. 368–90.

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press

VII. CONCLUSION

Many critics of utilitarianism reject aggregation because, according to them, aggregation does not take the separateness of persons seriously. I considered various interpretations of the separateness of persons. But the idea of the separateness of persons does not necessarily lead us to non-aggregative moral principles, and it is perfectly consistent with some aggregative moral principles such as utilitarianism and egalitarianism. Therefore, critics of utilitarianism cannot say that we should reject aggregation *because* aggregation does not take the separateness of persons seriously. Some critics of aggregation believe that if we judge x to be superior to y because x has more aggregated goodness, then there is a sense in which we are indeed ignoring the separateness of persons. But this is a simple statement of moral view, not an argument against aggregation. The grounds for rejecting aggregation must be something else, and the burden of proof is on critics of aggregation. Indeed, the separateness objection does not quite add up.

iwao.hirose@mcgill.ca

https://doi.org/10.1017/S0953820812000386 Published online by Cambridge University Press