

Aggregation and the Separateness of Persons

Iwao Hirose
McGill University and CAPPE, Melbourne

September 29, 2007

1 Introduction

According to some moral theories, the gains and losses of different individuals are combined and balanced in order to judge the rightness and 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 will always be justified by the fact that this brought 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 critical of aggregation. For them, to say that aggregation should be allowed is to say that we carve up a healthy person and transplant his organs to five needy patients, because this makes the outcome strictly better. Thus, critics of utilitarianism rule out aggregation from their proposed moral theories.

Criticism of aggregation, however, appears to go too far. Suppose that we are faced with a choice between (a) rescuing one stranger and letting five different strangers die, and (b) rescuing the five strangers and letting the one stranger die. Suppose further that there is no morally relevant difference between these six strangers. Many people believe that it is right to choose (b), i.e. to save the greater number of individuals. However, it is pointed out that the case for saving the greater number cannot be made unless we appeal to aggregation. That is, the judgement implicitly balances the combined losses of the five strangers' lives and the loss of one stranger's life. If we rule out aggregation, it is claimed that we must agree with John Taurek (1977), who holds that we should flip a fair coin to decide either (a) or (b). Taurek believes that there is no perspective from which the combined losses of five lives can be said to be bad, and that we should give an equal

¹This definition is narrower than, and different from, that in economics. In economics (or more specifically, social choice theory) aggregation means a collective decision making rule that maps a set of individual preference orderings over states of affairs to an ordering such as social welfare ordering. There is another type of aggregation, i.e. *intrapersonal* aggregation. It is concerned with how the temporal parts of a person's life are combined within his life. I do not consider this type here.

and positive respect to each stranger's life. The coin-flipping, Taurek thinks, best captures his belief, because it gives each stranger the same chance (i.e. 50%) of being saved. But this appears counterintuitive to many people. Some philosophers try to establish the case for saving the greater number without appealing to aggregation. I have discussed this important issue elsewhere.² In this paper I will focus on the motivation for non-aggregative principles.

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*. Rawls criticized both classical and average utilitarianism, and put forward two principles of justice. Neither of the principles includes aggregation in any form. 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.³ 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 would be found in the idea of the separateness of persons.

The question I consider in this paper is this: Does the notion of the separateness of persons lead us to non-aggregative moral principles? I shall argue that it does not.

2 The nature and scope of aggregation

Before taking up the notion of the separateness of persons, let me first elucidate four basic properties of aggregation, and then make a remark on the scope of aggregation.

Aggregation is concerned with the moral trade-off between the gains and losses of different groups of individuals. According to aggregative moral theories, the gains of a group of individuals can *morally* outweigh the losses of 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 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 of a group do not morally outweigh the loss for another group or individual. When there is a conflict of interests, non-aggregative principles resolve the conflict without

²See Hirose (2007). See also Kamm (1993), Otsuka (2006), and Wasserman and Strudler (2003).

³See Nagel (1979, p. 125), Nozick (1974, p. 33), and Scanlon (1982, p. 123).

combining the gains and losses of different individuals.

Aggregation consists in four basic properties. The first feature is that aggregation requires the comparability of the gains and losses for different people. I call this the *unit-comparability*.

Unit-comparability: For any pair of persons, it is possible to compare the gain of one person and the loss of another different person.

It is easy to see that we cannot combine and balance the gains and losses unless we can compare the gains and losses of different individuals. I think unit-comparability to be uncontroversial.⁴

The second property is a version of impartiality.

Impartiality: Two alternatives are morally indifferent if they differ only with regard to the identities of people.

The basic idea of impartiality is that every person should receive an equal respect as others. To see the gist of this property, consider the two-person case, and compare two alternatives $x=(\text{losing an arm, losing no arm})$ and $y=(\text{losing no arm, losing an arm})$, where the brackets show what happens to persons 1 and 2. The only difference between x and y is that the personal identities are permuted. The other features remain the same. Impartiality holds that x is morally indifferent to y . That is, the permutations of personal identities do not affect our moral judgement. If a moral principle claims that either x or y should be chosen without judging these two alternatives are strictly indifferent, the person losing an arm could object to such a moral principle, because he is not shown an equal respect. There might be other ways to define the notion of impartiality, but it is this version of impartiality that aggregation appeals to.

Few people would disagree with impartiality in the two-person case. But some people would disagree when the number of people is greater than two. Compare $x = (5, 3, 1)$ and $y = (3, 1, 5)$, where the brackets show the level of three people's well-being. Impartiality requires us to judge that x is indifferent to y . Some people would disagree with this. The simple majority rule would choose x , because x is better for two people and worse for one person. Those who wish to minimize the maximum loss for each person would judge that we should choose y .

The third property is the *positive responsiveness* to the gain of every person.

⁴This is not the case in economics. In so-called new welfare economics, it is believed that there is no scientific basis for the comparison between one person's preference satisfaction and another's.

Responsiveness: If an alternative benefits some person and harms no person, it should be chosen.⁵

In plain words, responsiveness holds that the gain of a person counts as strictly positive, and that moral judgements should register the additional gain of this person. If a moral principle does not acknowledge the improvement of a person's condition, then there would be a legitimate reason for this person to object to this principle, because he could complain that it does not show him a positive respect.

The fourth property is what I call *continuity*.

Continuity: There is a sufficiently large number of people, n , such that the gains for n -people outweigh the losses for another different group of people.

Continuity implies that there is an Archimedean point where the gains of a sufficiently large group of individuals morally outweigh the losses of another group of individuals if either the number of benefited individuals is sufficiently large or the benefit for each individual in this group is sufficiently large (or both). Some people would disagree. For example, J. S. Mill's distinction between the higher and lower pleasure can be seen as a case against continuity: the losses in the higher pleasure cannot be outweighed by the gains in the lower pleasure, no matter how great they would be.

Let me make a short remark on the scope of aggregation. Those opposed to aggregation are critics of utilitarianism, but not *vice versa*. That is, the rejection of aggregation is a sufficient condition to refute utilitarianism, but not a necessary condition. For example, advocates of egalitarianism are opposed to utilitarianism. Utilitarianism is not concerned with how well-beings spread across individuals. But egalitarianism is concerned with the distribution of people's well-being. There are several versions of egalitarianism. Some versions hold that equality is a good-making property, and that a more equal distribution is better than a less equal one, other things being equal. For example, consider what is known as the *Gini social welfare function*. The goodness of a state 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.⁶ This version would claim that the relative goodness of states of affairs is given by people's average well-being and the disvalue of inequality. This version

⁵If a moral principle is understood within the framework of preference satisfaction, this condition is equivalent to the Pareto principle.

⁶Notice that this equation is equivalent to:

$$\begin{cases} 1/4w_1 + 3/4w_2 & \text{if } w_1 \geq w_2 \\ 3/4w_1 + 1/4w_2 & \text{if } w_1 \leq w_2. \end{cases}$$

It is easy to verify that this function is symmetric (thus, satisfying impartiality), strictly increasing (thus, satisfying the responsiveness condition), and continuous, and assumes interpersonal comparability of utility.

of egalitarianism thus includes aggregation. Some versions of egalitarianism are perfectly compatible with aggregation. This means that some critics of utilitarianism do not need to rule out aggregation. There are many types of criticism of utilitarianism, and criticism of aggregation is just one type amongst others. It is not necessary to rule out aggregation in order to refute utilitarianism.

3 The separateness of persons objection

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.⁷ 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”.⁸ This is the *separateness of persons objection* to classical utilitarianism (hereafter, the *separateness objection* for short). Following Rawls, other philosophers such as Thomas Nagel and Robert Nozick also appeal to the separateness of persons to criticize utilitarianism.⁹

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 believe that Rawls and other critics of utilitarianism are concerned with this kind of question. 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 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

⁷See Scheffler (2004).

⁸Rawls (1971: pp. 26-27). For the general discussion on the separateness of persons, see Brink (1993), McKerlie (1988), Parfit (1984), Raz (1986), and Scheffler (2004).

⁹See Nagel (1970, p. 142) and Nozick (1974, p. 33). Bernard Williams (1973) appeals to a similar, but distinct, notion of “integrity”.

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. Firstly, when Rawls justifies two principles of justice, he appeals to the parties' rational decision-making in the original position. In the original position, the parties are deprived of information such as their places in society, physical and psychological dispositions, and their own conceptions 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 independently 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 picks out 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.

Secondly, Rawls's Difference Principle implies impartiality. Suppose that the 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 the 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.

The second interpretation is that the separateness objection is concerned with the combination of impartiality and responsiveness. Impartiality and responsiveness 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 responsiveness. Compare two alternatives in the two-person case, $x = (\text{losing an arm, losing no arm})$ and $z = (\text{losing no arm, losing two arms})$. The combination of the two notions holds that we should choose x . To illuminate this judgement, suppose another possible state $z' = (\text{losing no arm, losing an arm})$. By responsiveness, when we compare z and z' , z' should be chosen. By impartiality, x is indifferent to z' . Consequently, the combination of impartiality and responsiveness judges that x should be chosen when we compare x and z .¹⁰

¹⁰Hirose (2001) shows that impartiality and responsiveness jointly suffice to justify the case for saving the five strangers in the saving-five-or-saving-one example. Needless to say, the combination of impartiality and responsiveness is weaker than aggregation because it does not require continuity. Some non-aggregative consequentialist principles satisfy impartiality and responsiveness, but judge that x should be chosen. An example is the

This judgement may be challenged. It would 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 such that 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.

On the face of it, this challenge seems to make sense, but in fact it does not. Firstly, 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. Secondly, if we choose z instead of x , there is also a sense such 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 that we treat one person as a means to improve the another, insofar as we choose either course of action. The only way of respecting the separateness of persons, according to this interpretation, is to choose neither. That is, we do not help any person. If the separateness objection is concerned with the combination of impartiality and responsiveness, we should make no moral judgment 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 the moral theory 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 responsiveness. It is concerned with a stronger notion. It is the combination of impartiality, responsiveness, and continuity; that is, aggregation.

The separateness objection seems to make sense, 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,

lexicographic extension of the maximin rule (leximin, for short). Leximin first compares the level of the worst off person across alternatives, and if the worst off persons are at the same level, then compares the second worst off person, and so on. In the comparison between x and z , the worst off person in x is person 1, who loses one arm, whereas the worst off person in z is person 2, who loses two arms. Leximin holds that x is strictly better than z , and hence that it is right to bring about x . Leximin, however, is non-aggregative. Leximin holds that we should benefit the worst off person even if we incur some loss to other people, however many. That is, a small benefit of the worst off person, however small, cannot be outweighed by the combined losses of the large number of people, however large. Leximin does not combine the gains and losses of different people into the gains and losses of a group of individuals.

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”.¹¹ 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 (1970: p. 138) claims that utilitarianism “ignores the distinction between persons”. Nagel (1979: p. 123) later proposes what he calls *pairwise comparison*, which is not aggregative. 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 between the gains and losses of two persons, it rules out the possibility that the combined gains of a group of different persons outweigh the combined losses of another group. Even if x is slightly better than y for a million people but if 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. It would be claimed that his non-aggregative principle takes the separateness of persons seriously. Likewise, Nozick (1974: p. 33) claims that balancing the gains and losses of different people “does not sufficiently respect and take account of the fact that he is a person, that his is the only life he has”. 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. 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 egalitarianism does not take the separateness of persons seriously. This is a surprising consequence of the separateness objection. 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.

¹¹Rawls (1971: p. 26).

4 Two arguments against 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 would 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 objections to this point.

Firstly, non-aggregative principles may take the separateness of persons seriously, but *ignore* many persons. Non-aggregative principles do not combine the gains and losses of different people, instead neglecting the gains and losses of many people; they must focus on only one person, and do not consider the gains and losses of other people. Consider the maximin rule. According to the maximin rule, the relative goodness of states of affairs is judged by the relative level of the worst off person in each state.¹² This means that the judgement about the relative goodness of states of affairs is not affected by the gains and losses of the non-worst off people. The state of the non-worst off does not affect our moral judgement at all, and practically it is not taken as a morally relevant consideration. Thus, the maximin rule ignores many people.

On the other hand, aggregative principles take every person seriously. If a person gets additional gains, then the value of combined gains would be increased. Aggregative principles show a positive responsiveness to the gains of every person, and negative responsiveness to the losses of every person. Every change in the gains and losses for every person is acknowledged in terms of the value of combined gains and losses. Thus, aggregative principles respect every person, although it would be claimed that it does 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 respect in which his theory precisely does not ignore any group in society. In the 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 ignores the non-worst off people. This claim, 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 every person's interests seriously. If this is correct, why does Rawls need to hold the non-aggregative Difference

¹²Strictly speaking, the maximin rule is distinct from Rawls's Difference Principle. The Difference Principle applies only to (a) the basic economic and social structure, (b) the primary social goods, and (c) the representative groups in society.

Principle? He could hold an aggregative principle insofar as he gives a lexical priority to the other principles that respects the separateness of persons and every person's interest.

Now, it should be asked whether ignoring many people would be more acceptable than not taking the separateness of persons seriously. I do not believe that we can possibly ignore the gain and loss of any person. When there is no conflict of interests between different people in a morally relevant sense, we do not need a moral principle for settling the conflict. But, if the conflict is not morally relevant, we do not need a moral principle. For example, when there is a conflict of interests between two card players, we do not call for a moral principle – we just let them play. When there is a conflict of interests in the morally relevant sense though, we do need some moral principle to settle the conflict. This means that every person's gains and losses are morally relevant whenever we call for some moral principle to settle the conflict. In other words, we are not allowed to ignore the gains and losses of any person when we recognize the gains and losses to be morally relevant. Thus, I believe that ignoring the persons is less plausible than not taking the separateness of persons seriously.

Secondly, it is not clear whether aggregative principles do not 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 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".¹³ As I discussed earlier, many versions of egalitarianism do not 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 sense. Consider an aggregative principle, 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 sep-

¹³Parfit (1984: p. 330).

arately. Classical utilitarianism would hold that Annie's well-being should count for one, and that Betty's well-being should count for one, too. On this view, to give the same weight to everybody'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.¹⁴

Notice that I do not contend that the idea of the separateness of persons directly entails one of these interpretations. 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. 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. It is concerned with the comparison between the competing moral views. I have been considering the logical relation between the separateness of persons and non-aggregative principles. 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 does not support 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 would accept the trivial fact that different people live different lives, and that each person's life should be respected separately from other people's lives. In the criticism of utilitarianism, it is claimed that, in virtue of the meaning of the separateness of persons, 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 used to state merely what they want to reject. But the separateness of persons is supposed to explain why we 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.

¹⁴It may be pointed out that my claim here is not about the separateness of persons, but the equal respect. This point can be correct, but does not undermine my claim. Classical utilitarianism cannot give the equal respect to Annie and Betty unless it acknowledges the fact that Annie and Betty live different lives.

5 The separateness of persons and contractualism

Here is another interpretation: The separateness of persons is concerned with the acceptability of moral principles to those affected by the actions they would 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 to bear that burden if every person including himself agrees to those principles in a certain hypothetical situation. Imposing a burden in this way, it would be argued, respects 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, "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 the 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 Difference Principle would ignore the non-worst off people. But these ignored people are supposed to have no legitimate complaint about how they are treated when the veil of ignorance is lifted. Insofar as the moral principles are agreed by every person in the original position, it respects the separateness of persons. Thus, the notion of the separateness of persons would lead us to contractualism in the original position.

The separateness of persons thus construed would 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”.¹⁵ For the parties in the original position, it would not be rational to choose classical utilitarianism. However, there is a “surprising contrast”¹⁶ between classical and average utilitarianism. According to Rawls, the parties in the original position may well choose average utilitarianism in the original position. Rawls does not rule out the possibility of average utilitarianism.

According to Harsanyi (1953, 1955), it is average utilitarianism that the parties would rationally choose 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 hence that the parties 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 takes the separateness of persons seriously.

Rawls raises a set of criticisms against average utilitarianism. I do not reiterate his criticisms, because these criticisms are independent of the idea of the separateness of persons. However, let me mention one criticism in order to point out that there is at least another principle that could be chosen in the original position. According to Rawls (1971: p. 154), the worse off people in an actual situation would not live up to the principles chosen in the hypothetical situation, unless these principles assure the parties of a satisfactory minimum. Average utilitarianism would, Rawls claims, fail to assure such a satisfactory minimum, and would thus be unstable. If Rawls’s criticism is correct, the parties would not choose average utilitarianism. However, this does not imply that the parties would choose non-aggregative principles. It merely implies that the parties would choose a principle that guarantees the satisfactory minimum. Such a principle can be aggregative. For example, the parties would choose the principle that maximizes the average utility and guarantees a minimal level of utility to all. This principle does not allow any person fall below a certain level of well-being. But, if there is no person below the threshold level, it is equivalent to average utilitarianism. This principle is another aggregative principle that could be chosen in the original position.

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.¹⁷ I do not intend to answer this question. It seems to me that there is no conclusive argument that supports either aggregative or non-aggregative principles as far as the choice of moral principles in the hypo-

¹⁵Rawls (1971: p. 163).

¹⁶See Rawls (1971: p. 189).

¹⁷For example, see Arrow (1973) and Hare (1973).

thetical situation is concerned. This is because it is not clear what proves Rawls's claim to be true or Harsanyi's claim to be true. More precisely, it is not clear how we could know the principle that the rational parties would unanimously agree to choose behind the veil of ignorance. This is why I do not intend to, and cannot, give the answer to the question.

Some people would appeal to the literature of experimental economics that focus on what subjects *actually* choose in the controlled laboratory. Many results in experimental economics suggest that the overwhelming number of subjects, who are placed in the situation similar to Rawls's original position, actually choose the maximization of average *income* with a guaranteed minimal *income*, and that few choose non-aggregative principles such as Rawls's.¹⁸ I do not believe, however, these experimental results establish the case for aggregative principles and against non-aggregative principles. This is because, as Konow (2003, p. 1196) 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”. Thus, I do not intend to, and cannot, give the 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. 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.

6 Scanlon's contractualism

Before I conclude this paper, let me briefly discuss the possibility of another version of contractualism: that is Thomas Scanlon's (1982). 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 separateness of persons and non-aggregative principles. If (a) the sep-

¹⁸See Frohlich, Oppenheimer, and Eavey (1987), Frohlich and Oppenheimer (1992) and Konow (2003).

arateness 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 obtain the conclusion 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) at this stage. In this section, I will focus on (c).

According to Scanlon (1982), the following reasoning underlies the Rawls-Harsanyi version of contractualism. 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. A person's judgment that the principle is acceptable must be impartial. To judge impartially that a principle is acceptable is to judge that it is one which he would have reason to accept no matter who he was (i.e. to judge that it is a principle which it would be rational to accept if he did not know which person's position he occupied). 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 it 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 identification of moral principles and the justification 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 to only (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. the maximin rule. If this is correct, the maximin rule would not be justifiable to person 1. Thus, Scanlon believes that the principle that would be chosen by the rational parties in the hypothetical situation turns 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.

Scanlon wants to appeal to (P), and grounds the 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.¹⁹

This formula identifies what it is to be wrong, and also explains why it 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". When there exists even one person who has a legitimate reason to reject a principle, the same principle is not seen as a basis of judging the rightness and 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 unrestricted aggregation.

The following question arises: Does the individualist restriction rule out aggregation? To put it another way, are non-aggregative principles 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 could *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 the veto power to individuals. It can constrain the cases, where aggregative principles can be the basis for judging the rightness and wrongness of an act. It is not supposed to rule out any particular type of principles. 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. The individualist restriction can eliminate some aggregative principles in some contexts. But it is perfectly consistent to say that some aggregative principles need not be rejected in other contexts. To say that, given the individualist restriction, the process of the justification of moral principles is not necessarily aggregative is one thing. To say that the principles that

¹⁹Scanlon (1982, p. 110).

would pass the non-rejectability test are always non-aggregative is quite another. I have pointed out that Scanlon need not to commit to the second claim. I believe that Scanlon can contend the first claim. But this first claim would be supported by Rawls and Harsanyi. This is because the process of the justification of moral principles are not aggregative in their version of contractualism, 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.²⁰

7 Conclusion

Many critics of utilitarianism rule out 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 rule out aggregation *because* aggregation does not take the separateness of persons seriously. The grounds for rejecting aggregation must be something else. The idea of the separateness of persons does not quite add up.

References

- [1] Arrow, K. (1973). "Some Ordinalist-Utilitarian Notes on Rawls's Theory of Justice," *Journal of Philosophy*, 70.
- [2] Brink, D. (1993). "The Separateness of Persons, Distributive Norms, and Moral Theory," in Frey, R. G. and C. W. Morris (eds.) *Value, Welfare, and Morality*. Cambridge: Cambridge University Press.
- [3] Frohlich, N., J. A. Oppenheimer, and C. L. Eavey, (1987). "Choices of Principles of Distributive Justice in Experimental Groups," *American Journal of Political Science*, 31.
- [4] Frohlich, N. and J. A. Oppenheimer, (1992). *Choosing Justice*. Berkeley: University of California Press.
- [5] Hare, R. M. (1973). "Rawls's Theory of Justice," in *Philosophical Quarterly*, 21.
- [6] Harsanyi, J. C. (1953). "Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking," *Journal of Political Economy*, 61.

²⁰Parfit (2003) makes a stronger claim. He argues that Scanlon should drop the individualist restriction because it limits the scope of his contractualism.

- [7] Harsanyi, J. C. (1955). "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy*, 63.
- [8] Hirose, I. (2001). "Saving the Greater Number without Combining Claims," *Analysis*, 61.
- [9] Hirose, I. (2007). "Aggregation and Non-Utilitarian Moral Theories," forthcoming in *Journal of Moral Philosophy*
- [10] Kamm, F. (1993). *Morality, Mortality. Vol. I.*, Oxford: Oxford University Press.
- [11] Konow, J. (2003). "Which Is the Fairest One of All? A Positive Analysis of Justice Theories," *Journal of Economic Literature*, 41.
- [12] McKerlie, D. (1988). "Egalitarianism and the Separateness of Persons," *Canadian Journal of Philosophy*, 18.
- [13] Nagel, T. (1970). *The Possibility of Altruism*. Oxford: Clarendon Press.
- [14] Nagel, T. (1979). "Equality," in his *Mortal Questions*, New York: Cambridge University Press.
- [15] Nozick, R. (1974). *Anarchy, State, and Utopia*. New York: Basic Books.
- [16] Otsuka, M. (2006). "Saving Lives, Moral Theory, and the Claims of Individuals," *Philosophy and Public Affairs* 34.
- [17] Parfit, D. (1984). *Reasons and Persons*. Oxford: Oxford University Press.
- [18] Parfit, D. (2003). "Justifiability to Each Person," *Ratio*, 15.
- [19] Rawls, J. (1971). *A Theory of Justice*. Cambridge, MA.: Harvard University Press.
- [20] Raz, J. (1986). *The Morality of Freedom*. Oxford: Clarendon Press.
- [21] Scanlon, T. M. (1982). "Contractualism and Utilitarianism," in Sen, A. K. and B. Williams (eds.) *Utilitarianism and Beyond*. Cambridge: Cambridge University Press.
- [22] Scanlon, T. M. (1998). *What We Owe to Each Other*. Cambridge Mass.: Harvard University Press.
- [23] Scheffler, S. (2004). "Rawls and Utilitarianism," in Freeman, S. (ed.) *The Cambridge Companion to Rawls*. Cambridge: Cambridge University Press.

- [24] Taurek, J. (1977). "Should the Numbers Count?" *Philosophy and Public Affairs*, 6.
- [25] Wasserman, D. and A. Strudler (2003). "Can a Nonconsequentialist Count Lives?" *Philosophy and Public Affairs*, 31.
- [26] Williams, B. A. O. (1973). "A Critique of Utilitarianism," in Smart, J. J. C. and B. A. O. Williams, *Utilitarianism: For and Against*. Cambridge: Cambridge University Press.