

A Defense of the Difference Principle beyond Rawls

Peter Koller

Abstract

The Difference Principle (DP) is one of the most original and disputed components of Rawls' theory of justice. The DP, which is contained in the second of Rawls' two principles, says that social and economic inequalities are to be to the greatest benefit of the least-advantaged members of society. The aim of my paper is to defend the DP, or at least its main idea. To this end, I will first recapitulate Rawls' explication of the DP, and then give a brief review of his theory in general and of his various arguments for the DP in particular. After that, I am going to scrutinize Rawls' justification of the DP, with the result that his arguments do not succeed. Since I nevertheless regard the DP as a sound principle of socio-economic justice, I finally want to make an attempt to defend it in a slightly modified form.

Introduction

The Difference Principle (DP) is certainly one of the most original, though greatly disputed components of John Rawls' conception of social justice, a component that has remained more or less unaffected by the revisions of this conception from its initial exposition in *A Theory of Justice* (1971, quoted as TJ) up to its final restatement in *Justice as Fairness* (2001, quoted as JF). Rawls' conception, designed to regulate and govern the institutional basic structure of a modern state-ruled society, is known to consist of two main principles: The first principle focuses on a society's *political order* and demands that every member ought to have equal basic civil rights and liberties insofar as they are compatible with the same rights and liberties for all. The second principle concerns society's *basic socio-economic system* in regard to which it aims to define the justifiable extent of social inequalities that may emerge from this system. Rawls' final formulation of the *second principle*, which contains the DP, runs as follows:

"Social and economic inequalities are to satisfy two conditions: first, they are to be attached to offices and positions open to all under conditions of fair equality of opportunity; and second, they are to be to the greatest benefit of the least-advantaged members of society (the difference principle)." (JF, 42 f.)

Although Rawls' two principles are closely connected and, therefore, form a unified whole, I will mainly deal with the second part of the second principle, the DP, and refer

to the other demands only when this is necessary to illuminate the features of the DP under consideration. In order to simplify the matter, I will assume throughout my paper that the first principle, the principle of equal liberty, is satisfied so that there is no need to discuss Rawls' highly questionable proposition that this principle has absolute priority over the second. The aim of my following considerations is to defend the DP, or at least what I regard as its main idea. With this aim in mind, I shall first recapitulate Rawls' explication of this principle in order to understand it correctly (1). Then, I want to give a brief review of Rawls' approach to social justice in general and of his various and repeatedly changing arguments for the DP in particular (2). After that, I am going to scrutinize Rawls' explication and justification of the DP through a critical discussion, which will lead me to the conclusion that Rawls' arguments do not succeed, and, even worse, cannot succeed since his approach suffers from considerable defects which undermine the DP (3). Since I nevertheless think the DP is, by and large, a sound principle of socio-economic justice, I finally want to make an attempt to defend the DP in a slightly modified form (4).

1. Rawls' Explication of the Difference Principle

Rawls' formulation of the second principle with the DP gives rise to a number of questions of interpretation, especially the following: What are the features of social and economic inequalities, or in other words, what are the properties of people that are to be compared in order to identify and measure these inequalities? What is meant by the statement that the inequalities are to be attached to offices and positions? What are the conditions of fair equality of opportunity? Who are the least-advantaged members of society? And how is the relationship between the inequalities under consideration and the position of the least-advantaged members to be understood so that the former are to the greatest benefit of the latter? As for these questions, Rawls' theory provides us with more or less definite answers, which, for the moment, I shall summarize briefly without scrutinizing them in detail. The problems which may result from some of these answers will be discussed later on.

As for the initial question concerning the nature of social and economic inequalities, Rawls defines them as the members' relative shares in certain sets of those societal values or resources of human well-being which he calls "primary social goods" (TJ, 90 ff.; JF, 58 ff.). *Primary social goods* in general are characterized by two features: first, they result from social cooperation and are distributed through the society's institutional basic structure; and, second, they are scarce goods of supreme order of which every individual reasonably wants to possess more rather than less, because "with more of these goods men can generally be assured of greater success in carrying out their

intentions and in advancing their ends, whatever these ends may be" (TJ, 92). Or as Rawls explains, primary social goods are "social conditions and all-purpose means that are generally necessary to enable citizens adequately to develop and fully exercise their two moral powers [i.e. their capacities for a sense of justice and for a conception of the good], and to pursue their determinate conceptions of the good" (JF, 57).

Rawls proposes a list of such goods which contains five categories: (i) basic rights and liberties, (ii) freedom of movement and free choice of occupation, (iii) powers and prerogatives of offices and positions of authority and responsibility, (iv) income and wealth, and (v) the social bases of self-respect (JF, 58 f.). While the first two kinds of primary social goods, i.e. the fundamental civil and political rights, are thought to be subject to the first principle requiring their unconditional equality, the last three kinds, that contain the most important socio-economic resources, fall in the domain of the second principle. This principle permits or even demands inequalities of their distribution to the extent in which such inequalities are necessarily or unavoidably connected to a socio-economic system that improves the position of the worst-off members of society, provided that they are attached to public offices and social positions open to all under fair equality of opportunity. This leads to the question of the significance of these offices and positions.

Since Rawls' principles of justice apply to the *institutional basic structure* of society rather than to the interpersonal interactions and relationships among its members, they address these members not as individuals with their own personal properties and fates, but "as representative persons holding the various social positions, or offices, or whatever, established by the basic structure. Thus in applying the second principle [...] it is possible to assign an expectation of well-being to representative individuals holding these positions. This expectation indicates their life prospects as viewed from their social station." (TJ, 64) So the DP does not strive to raise the well-being of particular individuals who, for whatever reasons, are worst-off at a certain time of their life, but it rather demands to improve the *social position*, namely the life-time prospects of well-being of the members of the least advantaged social group to the greatest possible extent in a way that meets the condition of fair equality of opportunity (cf. Schaller 1998, 370 ff.).

This condition, Rawls tells us, requires "not merely that public offices and social positions be open in the formal sense, but that all should have a fair chance to attain them. To specify the idea of a fair chance we say: supposing that there is a distribution of native endowments, those who have the same level of talent and ability and the same willingness to use these gifts should have the same prospects of success regardless of their social class of origin [...]. " (JF, 43 f.) But this explication is

ambiguous, since it may be interpreted in various ways ranging from a very weak to a greatly strong interpretation: Understood in a weak sense, it may mean that people from different social classes who *actually have* the same qualifications and ambitions in spite of their unequal starting positions ought to have equal prospects of success. However, at the other end of the scale, it may mean that people from different social classes who, due to their natural endowments, potentially *could achieve* the same level of qualification and ambition if they enjoyed similar promotion of their talents and abilities, should have equal prospects of success. Whereas the first interpretation appears much too weak, because it demands nothing more than formal openness of desirable social positions with a fair selection procedure, the second seems too strong, since it would require a complete equalization of the children's starting positions, including educational opportunities and material means (cf. Pogge 1989, 161 ff.). So a plausible interpretation should lie somewhere between these extremes.

A further question is how to *identify* the group of the least advantaged members of society whose prospects of well-being could be improved to the greatest possible extent through social and economic inequalities. This question calls for a measure of people's well-being in terms of their share of the relevant primary social goods, including powers and prerogatives, income and wealth, and the social bases of self-respect. Such a measure could be easily found, if the distributive patterns of these various kinds of primary goods among different social groups would always coincide, so that all members of society without influence and power were also poor and lacking other bases of self-respect. Where this is not the case, the matter becomes more difficult. Yet, Rawls thinks it is possible to aggregate people's relative shares of the various primary goods to an appropriate index that may be used to identify the least advantaged group of society at least in an approximate way. In *A Theory of Justice* we find two suggestions for ways of defining the least fortunate group: "One possibility is to choose a particular social position, say that of the unskilled worker, and then to count as the least advantaged all those with the average income and wealth of this group, or less. [...] Another alternative is a definition solely in terms of relative income and wealth with no reference to social position. Thus all persons with less than half of the median income and wealth may be taken as the least advantaged segment." (TJ, 98) A more refined strategy is offered in the *Revised Edition* of TJ: "To fix ideas, let us single out the least advantaged as those who are least favored by each of the three main kinds of contingencies. Thus this group includes persons whose family and class origins are more disadvantaged than others, whose natural endowments (as realized) permit them to fare less well, and whose fortune and luck in the course of life turn out to be less happy, all within the normal range [...] and with the relevant measures based on social primary goods." (Rawls 1999, 83)

If we assume that there is a way to determine the least fortunate group, we face the following question: How is it possible that social inequalities may amount to the benefit of the least fortunate members, and under what circumstances are they to their greatest benefit? First of all, Rawls assumes that a society is a complete system of social cooperation whose overall product is not a constant sum, but variable depending on the distribution of its members' benefits and burdens. Consequently, it appears to be possible to increase the social product through schemes of division of labor including inequalities of power, influence, income and wealth. "By varying wages and salaries", Rawls notes, "more may be produced. This is because over time the greater returns to the more advantaged serve, among other things, to cover costs of training and education, to mark positions of responsibility and encourage persons to fill them, and to act as incentives." (JF, 63) Such inequalities are to the benefit of all members of society, if the benefits of the growing social product which they stimulate are "spread throughout the system and to the least advantaged" (TJ, 78). And if the social inequalities are even to the *greatest benefit* of the latter, they meet the DP and are, in Rawls' view, therefore just. In this context, he differentiates between "perfectly just" and "just throughout": Accordingly, a *perfectly just* scheme is obtained when "no changes in the expectations of those better off can improve the situation of those worst off", while an arrangement is merely *just throughout*, when a further increase of inequality by improving the position of the more advantaged "would raise the expectations of those in the lowest position" (TJ, 78 f). What the latter condition exactly means, however, is still ambiguous. I will return to this point later on (in section 3).

So much to Rawls' explications concerning interpretation of his DP. On this basis, I want to take a look at his arguments for its justification (see Koller 1998).

2. Rawls' Arguments for the Difference Principle

Since Rawls' reasoning in favor of the DP is deeply embedded in the highly complex architecture of his entire theory of justice, which I cannot treat here at length, I restrict myself to a rough sketch of its main points.

It is known that Rawls models the moral viewpoint, from which we are to make our judgments on matters of justice, as a fictitious initial situation, the "original position", into which all members of a society, considered as free and equal citizens, enter in order to reach an unanimous agreement on the principles of social justice which should govern their social life. In the original position, people are constructed as persons who, on the one hand, are *equally rational* in the sense that they all exclusively pursue their own well-considered interests, while, on the other hand, they are forced to *impartial reason*,

since they are deprived of any particular information about their individual natural properties and social positions through a so-called “veil of ignorance”. However, they are familiar with the general facts relevant for the choice of principles of justice, such as the features of human nature, the problems of social order, and the regularities of economic systems. In view of these facts, people in the original position conceive of their respective society as a *complete system of social cooperation* whose main rules and institutions, its *basic structure*, will greatly affect their ways of life and prospects of well-being. Consequently, they regard the society's basic structure as the primary subject of social justice, whose principles are to be designed to regulate the distribution of primary social goods in a fair, generally acceptable way.

On the basis of these assumptions, Rawls develops his argument for his conception of justice in general and the DP in particular. The argument starts from the consideration that the parties in the original position would agree on an *equal* distribution of all social values subject to the regulation of the society's basic structure, if these values were forming a constant sum, such as, for instance, a given cake. As they do know, however, that this is not generally true, because it may be possible to increase the product of social cooperation through permitting unequal shares of its benefits to everybody's advantage, they accept the following *general conception of justice*: “All social values – liberty and opportunity, income and wealth, and the bases of self-respect – are to be distributed equally unless an unequal distribution of any, or all, of these values is to everyone's advantage.” (TJ, 62)

Yet, this general conception, Rawls argues, will not be regarded as the final word by the parties in the original position, since it applies to the various categories of primary social goods in an indiscriminate way without taking their different significance and weight into account. First of all, the parties will insist on an equal distribution of basic liberties, for there is no reason to expect that, at least under normal social conditions, inequalities of these liberties would benefit even those with lesser liberties. Secondly, “if the parties assume that their basic liberties can be effectively exercised, they will not exchange a lesser liberty for an improvement in economic well-being” (TJ, 151 f.). So they arrive at a more differentiated conception of justice consisting of *two principles*. The first requires an appropriate arrangement of equal basic civil rights and liberties, whereas the second, subordinate, principle concerns the society's institutional arrangement regulating the distribution of the other primary social goods, including income and wealth.

Rawls' first, provisional statement of the second principle states that “social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be to everyone's advantage, and (b) attached to positions and offices open to all”

(TJ, 60). In order to transform this rather vague statement into the principle's final version with the fair equality of opportunity condition and the DP, he pursues his argument, as it appears in *A Theory of Justice*, along two parallel paths, one formal or deductive, and the other more informal or deliberative.

While following the *formal* path, Rawls makes a number of more or less daring assumptions in order to move from the rather weak requirement that inequalities shall be to *everyone's advantage* to the DP's much stronger requirement that they have to be to the *greatest benefit* of the *least-advantaged*. One of these assumptions is the empirical thesis that socio-economic inequalities are, in general, both "chain-connected" and "close-knit", so that any raise of more favored social positions that improves the prospect of the lowest position also raises the prospects of all positions in between (TJ, 80), and, furthermore, that any change in the prospects of the better off affects the prospects of the worst off in some way (TJ, 81 ff.). Due to this thesis, Rawls thinks it possible to proceed from a state of affairs which is to the advantage of the worst off to one which is to everyone's advantage in order to replace the reference to the latter by that to the former. A second assumption is his normative premise that the parties in the original position will apply the *maximin rule* for choice under uncertainty. This rule tells us to maximize the minimum, or more precisely, "to rank alternatives by their worst possible outcomes: we are to adopt the alternative the worst outcome of which is superior to the worst outcomes of the others" (TJ, 152 f.). Rawls regards this highly conservative rule as plausible under three conditions, which he deems to be satisfied by the original position: (i) one cannot take account of the respective likelihoods of the possible alternatives in consideration, (ii) one does not care very much to gain more than the minimum which the maximin rule grants; and (iii) one can hardly accept the worst outcomes of the other alternatives (cf. TJ, 154). With this premise and further reasonings dedicated to the comparison of his principles with alternative conceptions of justice, Rawls narrows the indefinite requirement of some advantage for everyone down to that of the greatest benefit of the worst off. So he eventually arrives at the DP in a way for which, in *A Theory of Justice*, he claims deductive stringency.

Beside this alleged formal proof, Rawls also pursues a more *informal*, deliberative way of reasoning in favor of his principles, including the DP. While this argument was not yet highly developed in *A Theory of Justice*, it becomes dominant in his subsequent writings, including his monograph *Political Liberalism* (1993, quoted as PL), where he even admits that his previous attempt to derive the principles from the original position with the means of decision theory was misleading.

The fundamental elements of his informal argument include the following ideas: first, the idea of a *political conception of justice* according to which his conception is not a

comprehensive moral doctrine claiming universal validity, but rather a “moral conception worked out for a specific kind of subject, namely, for political, social, and economic institutions” of modern societies in the form of constitutional democracy (PL, 11 ff.); second, the idea of *society as a fair system of cooperation*, whose members are characterized by two moral powers, namely a capacity for a sense of justice on the one hand, and a capacity for a conception of the good on the other (PL 15 ff.); and third, the idea of the *original position*, whose parties are now understood as persons who not only have the two moral powers mentioned, but are also thinking of themselves and each other as free citizens in the sense that they all have particular conceptions of the good, are equally entitled to make claims on common social institutions, and are capable of taking responsibility for their ends (PL, 22 ff.). On the basis of these ideas and in view of the fact of reasonable pluralism of individual conceptions of the good, Rawls assumes that people enter into a forum of public reason in order to achieve a generally acceptable conception of justice that is both neutral to and supported by the various conceptions of the good and the surrounding comprehensive moral doctrines, or, as he puts it, a conception finds support in an “overlapping consensus” (PL, 133 ff.). Rawls contends that his two principles, namely, the demand for an adequate scheme of equal basic liberties and their priority, and the requirement of fair equality of opportunity combined with the DP, should be preferred to any other conception, since only these principles could guarantee all citizens the social conditions essential for the full development and exercise of their two moral powers as free and equal persons, particularly their capacity to realize their respective conception of the good within the limits of the public conception of justice. Yet, while Rawls seeks to substantiate this contention extensively for the principle of equal liberty, he abstains from providing specific grounds for the particular features of the DP (cf. PL, 289 ff.).

However, in the final exposition of his theory in *Justice as Fairness: A Restatement* (2001), Rawls again takes up some pertinent elements of the original formal argument for the DP, in particular the maximin rule, and integrates them, albeit in a slightly modified manner, into the sketched deliberative considerations. As far as I can see, the resultant reasoning for the principles differs from the original formal argument mainly by the fact that it now proceeds on the way of *balancing* the reasons that the parties in the original position may have for and against the two principles in comparison with alternative conceptions rather than claiming deductive strength. Accordingly, Rawls now calls the maximin rule “simply a useful heuristic device” not essential for the parties’ reasoning in the original position (JF, 99).

The revised argument for Rawls’ principles relies on *two fundamental comparisons*: The first comparison, in which the two principles, taken as a unit, are compared with the utilitarian principle demanding to arrange the society’s basic structure so as to maximize

the average welfare of all members, leads to a strong preference for a ramified conception of justice compound of two principles, the first of which calls for an appropriate scheme of equal basic liberties having priority over the striving for economic welfare (JF, 96, 101 ff.). Subsequently, the DP emerges from the second comparison in which the two principles “are compared with an alternative formed by substituting for the difference principle the principle of average utility (combined with a stipulated social minimum)”, in order to select a principle for regulating economic and social inequalities (JF, 96). In this second comparison, which, as Rawls admits, “turns on a less decisive balance of reasons” (JF, 97), the maximin rule is relevant only insofar as the parties compare the various principles under consideration in regard to their respective worst outcomes and find the level of welfare guaranteed by the maximin rule fully satisfactory (JF, 120).

To this effect, Rawls argues, the parties would prefer the DP to the principle of utility with a social minimum against the background of the *ideas of publicity, reciprocity and stability* for the following reasons: first, the DP would be more determinate than the utility principle and, therefore, less likely to create disputes and mistrust; second, the DP would be easier to accept than the utility principle, because it would ask less of the more advantaged citizens (to the benefit of the less advantaged) than the utility principle asks of the less advantaged members (for the sake of greater advantages for the more advantaged); and, third, the DP would cause less “strains of commitment” arising from its realization than the utility principle with a social minimum, since it would guarantee all members a share of goods owed them as free and equal citizens, while the minimum of the utility principle covers only the needs essential for a decent life owed persons in virtue to their humanity (JF, 126 ff.). Furthermore, Rawls’ exposition of this comparison presupposes the simplifying assumption that the society is divided in two groups only, i.e. the more advantaged on the one hand and the less advantaged on the other.

This concludes my summary of Rawls’ explication and defense of the DP. I shall now turn to its critical discussion which I will begin with some further questions concerning its interpretation that Rawls has left open. I will then proceed to various objections against his main arguments for its justification.

3. Critical Issues of the Difference Principle

A striking feature of the DP is its exclusive focus on the benefit of the *least advantaged* in judging social and economic inequalities. This leads to a number of problems (cf. Koller 1983). A simple problem already appears when we face a situation with several feasible distributive structures which equally provide the least advantaged group with

the greatest possible welfare, but differ in the welfare levels of the social classes above whose members also may fare badly, such as those of the second or of the third worst-off classes. According to the DP, if taken literally, all these distributive structures would have to be judged as equivalent, which is contradictory to wide-spread intuition. In order to cope with this situation, Rawls adopts a proposition by Amartya Sen (1970, 138; 1974) to the effect that the DP ought to be understood as a “leximin principle” that demands to proceed as follows: “first maximize the welfare of the worst-off representative man; second, for equal welfare of the worst-off representative, maximize the welfare of the second worst-off representative man, and so on until the last case, which is, for equal welfare of all the preceding $n - 1$ representatives, maximize the welfare of the best-off representative man” (TJ, 83). This interpretation, however, does not solve all problems arising from the feature of the DP mentioned.

As I said before, Rawls’ move from the rather weak general conception of distributive justice (according to which social inequalities are to be to the benefit of every member of society) to the much stronger DP (with its exclusive focus on the greatest benefit of the least advantaged) relies on his assumption that the society’s various social positions are interrelated in a special way characterized through what he calls *close-knitness* and *chain-connection*. This is clearly an *empirical* assumption, which, by and large, may have some plausibility for modern societies with money economy, but is certainly not a necessary truth. In fact, there seem to be cases which do not meet the conditions of close-knitness and chain-connection.

One such case is a situation where social inequalities have developed to such an extent that they have contributed to raise the welfare of worse-off social groups to its maximum level, so that a further improvement of their position is no longer possible, while it would still be possible to raise the welfare of the most advantaged through granting them further benefits that are not transferable and, therefore, would not reduce the welfare of the worse-off. This case gives rise to the question as to whether the DP admits or excludes an increase of social inequality resulting from a further raise of the welfare of the well-off that would neither improve nor worsen the welfare of the worse-off. Rawls does not provide a clear answer to this question, for his few remarks on the justified extent of social inequalities may be interpreted in two different ways leading to different variants of the DP (see van Parijs 2003, 202 ff.). On the one hand, Rawls says that inequalities are just only if they “are” to the advantage of the worst off members of society, or that they “must be” to the benefit of the least advantaged, which suggests that the inequalities must causally contribute to improving the lowest position. Understood in this more egalitarian way, the DP would exclude any increase of inequality by raising the prospects of the better off that would neither improve nor diminish the well-being of the least advantaged. On the other hand, Rawls claims that

the DP is compatible with efficiency, which implies that improvements of the better off should be acceptable, if they do not affect the well-being of the worse off. His previously quoted note on the leximin principle also supports this less egalitarian reading of the DP. In this case, however, I think there are reasons to prefer the first, more egalitarian interpretation of the DP.

A further case in which the conditions of close-knitness and chain-connection are not met is when a minor improvement of the position of a less advantaged group, e.g. unemployed people, may be achieved, but only at the price of a comparatively significant reduction of the welfare level of a social group whose members are slightly better off, such as, for example, unskilled workers with a job. Is it really clear that in such a case the only thing that counts is the maximum welfare level of the worse-off group with no consideration for the sacrifices that its achievement may impose on those who are slightly better off? I have doubts as to whether this is so easily possible without carefully balancing the advantages of the less advantaged against the costs of the more advantaged. These doubts may suggest a modification of the DP to the effect that it allows a *weighing up* between the additional benefits of less advantaged people and the losses of better-off groups in the case described. A promising guideline for such a modification appears to be the following proposition by Derek Parfit (2000, 101), which he names the *Priority View*: "Benefiting people matters more the worse off these people are". If the DP is modified in accordance with this view, it will take a weaker form that may be roughly stated as follows:

Social and economic inequalities are justifiable up to the point at which they are benefiting the members of less advantaged social groups more the worse off these people are (under condition of fair equality of opportunity).

I deem this variant preferable to Rawls' DP, for it leads, I think, to the same results, if social inequalities satisfy the conditions of close-knitness and chain-connection, while it is more flexible in cases in which these conditions are not fulfilled.

In his reasoning for the DP, Rawls relies on the assumption that there is an appropriate *welfare index* that enables us to determine the overall welfare level of the various social positions in a satisfactory way. This assumption, however, is built on sandy grounds, since the search for such a welfare index encounters significant difficulties, even if the list of primary social goods that Rawls deems to be subject to the DP were plausible. The variety of these goods, which include social powers and prerogatives, income and wealth, and the social bases of self-respect, makes it far more difficult to reduce them to a common denominator, as Rawls suggests. But this difficulty is a minor problem. The main problem with his list of primary goods is not that it contains too many kinds of

goods, but rather that it covers too *few* because of the fact that it only refers to divisible goods attached to social cooperation and, thereby, misses a number of pertinent issues of social justice. The missing issues include the following: first, the *burdens of social cooperation*, i.e. the extent to which the members contribute to social production; second, *public goods* whose provision has significant distributive effects on particular social groups; and third, society's *common assets of its natural environment and cultural heritage*. These gaps in Rawls' design of the subject of social justice give rise to some grave objections to his approach on which the DP is based, irrespective of whether or not this principle itself is deemed to be right.

First of all, distributive social justice concerns the distribution of social goods *and burdens* rather than goods alone. It is true that Rawls' focus on primary social goods also implies some reference to social burdens, since these goods, when assigned to particular individuals by the rules of the basic structure, take the form of individual rights that go hand in hand with correlative duties. But this fact alone does not provide a sufficient basis for taking account of the individual burdens in social life. The cause of Rawls' failing to pay appropriate attention to these burdens lies in his conception of society as a system of social cooperation among its citizens who all are assumed to be "normal and fully cooperating members of society over a complete life" (PL, 20, 21, 183; cf. Rawls 1984/85, 234). Due to this assumption, he not only leaves aside all contingencies that "prevent people from being normal and fully cooperating members of society in the usual sense", such as physical disabilities, mental disorders and involuntary unemployment, but also omits to relate the distribution of the social goods subject to the DP to the members' relative contributions to social cooperation. And this omission is one reason for his peculiar refusal of the significance of individual merits, deserts and achievements for a basic principle justifying socio-economic inequalities (cf. TJ, 103, 305 ff.; JF, 72 ff.). I consider this a mistake, since arguments based on merit, desert and achievement, properly designed, easily fit with the basic idea of the DP and even support it, as I shall later proceed to argue.

Secondly, a plausible conception of social justice should not only deal with the distribution of *divisible* social goods which the society's basic structure may assign to its members in the form of individual rights, such as their rights to civil liberty, political participation and private property, but it should also deal with the provision of *public goods* whose ways of provision exert significant distributive effects on individual members depending on their social position and personal abilities, such as the systems of public services, public infrastructure, public transport and public media. If, for instance, a country does not have a sufficient system of public transport, poor people who cannot afford a car will suffer from this failure much more than the wealthy. A particularly striking case of the distributive effects of public infrastructure is the situation

of people with physical disabilities who need appropriate devices in streets, buses and buildings for their mobility. So, in my opinion, Martha Nussbaum (2007, 109) is correct when she argues that Rawls' "account of the primary goods, introduced, as it is, as an account of the needs of citizens who are characterized by the two moral powers and by the capacity to be 'fully cooperating,' has no place for the unusual social arrangements that will need to be made in order to include as fully as possible people with physical and mental impairments." This critique, I think, can even be extended to the more general objection that Rawls' conception of primary social goods as divisible private goods is completely insensitive to the fact that people on different levels of social stratification have divergent needs for public goods.

Thirdly, Rawls does not take the goods provided by the society's *natural environment* and *cultural heritage* into account, such as the atmosphere, land, waters, wild animals, mineral resources, language, technical knowledge, writing, social traditions, and artistic achievements. Even though most of these goods do not occur in pure form, but are contained in human artifacts, the regulation of their accessibility and use is a pertinent matter of social justice, since they are not only essential ingredients of people's means of freedom and welfare, but also, in a sense, society's common assets to whose benefits all members are equally entitled. The society's common ownership of its natural environment and cultural heritage does not exclude that their goods are part of individual private property, but it requires an arrangement of the possession and use of those goods that is to the benefit of every member of society, even though such an arrangement may allow inequalities, if they are expected to increase everyone's benefit, at least in the long run. Thus, the idea of the society's common ownership of its natural and cultural values significantly strengthens an egalitarian approach to social justice. It is therefore surprising that Rawls refrains from making use of this plausible and widely shared idea, by contrast to some exponents of left libertarianism, such as Hillel Steiner (1994) and Philippe van Parijs (1995), who have tried to build complete theories of justice based on this idea.

To sum up, Rawls' fundamental assumptions concerning the nature of society and its primary goods are much too weak to provide a solid basis for his principles of justice, particularly the DP. His view of society as a system of cooperation to the mutual advantage of "fully cooperating members over a complete life" fails to do justice to those members who, for whatever reason, are prevented from full participation in social cooperation. This view is also not sufficient to back up his rather demanding egalitarian account of socio-economic justice, even if applied only to the small set of primary goods that Rawls takes into consideration (cf. Sandel 1982, 66 ff.). For if society is conceived of as nothing more than a system of cooperation to the mutual advantage of its members, it appears hardly plausible that the demand of distributive justice should

extend to the totality of primary goods rather than only to that portion of those goods that David Gauthier (1974, 14) terms “the *social surplus*, it is that portion of the total quantity of primary social goods which would not be produced without cooperation”. So Rawls’ understanding of society as a mere system of social cooperation fails to back his conception of social justice, or even undermines it (cf. Barry 1989, 250 ff.).

As for Rawls’ previously sketched arguments for the DP in particular, none of them are successful. The formal argument, relying on the maximin rule, is not effectual, because it presupposes certain assumptions about people’s utility functions and risk-aversion that appear hardly acceptable (cf. Harsanyi 1975; Waldron 1986). This seems to be the reason why Rawls distanced himself from this argument in his later writings. Although his deliberative argument appears somewhat more plausible, it is, in my opinion, no more conclusive. Recall that Rawls seeks to substantiate the DP through comparing it with the utility principle combined with a social minimum in light of the ideas of publicity, reciprocity and stability. Yet, the reasons that he advances in favor of the DP are hardly convincing: Neither is the DP, if we consider all its problems of interpretation and application, more determinate than the utility principle with a social minimum, nor is it clear that its realization causes less “strains of commitment” because of the fact that it guarantees all members of society a larger share of goods than they need for a decent life. And as Rawls, in this comparison, assumes a society consisting only of two classes, his argument can only apply to more complex patterns of social inequality if these patterns meet the conditions of close-knitness and chain connection. Whether or not these conditions are actually met, however, is a contingent matter that cannot be settled without careful empirical research.

My critical discussion of Rawls’ conception of justice in general and the DP in particular should not give the impression that I am rejecting them as a whole. On the contrary, in spite of the raised objections to Rawls’ reasoning for the DP, I do subscribe to this principle, although in a slightly modified form. So I now come to present my own argument in its favor.

4. A Defense of the Difference Principle in a Modified Form

In order to circumvent Rawls’ elusive assumption that social positions are interrelated through a chain connection, I propose to modify his DP according to Parfit’s Priority View so that it allows a gradual balancing of the relative welfare of different social positions rather than focusing on the last advantaged only. Accordingly, I suggest putting the *modified DP* as follows:

Social and economic inequalities are justifiable if and only if they are necessary ingredients or unavoidable side-effects of a society's socio-economic order which, in general, increases the welfare of less advantaged members more the worse-off these members are (under condition of fair equality of opportunity).

As far as I can see, this variant of the DP is equivalent with Rawls' version under the condition that social positions are interrelated through close-knitness and chain-connection, while it leads to different results, if this condition is not met. The social and economic inequalities to which this principle refers may be broadly understood in terms of the members' shares in certain kinds of primary social goods. In addition to those on Rawls' list (i.e. powers and prerogatives of social positions, income and wealth, and the social bases of self-respect), I would also count *natural and cultural values* and *public goods with distributive effects*. Furthermore, when the members' relative shares of these goods are compared, their relative *burdens* in social cooperation need also to be taken into consideration.

The contention that the totality of the primary goods just mentioned is subject to distributive justice requires *a more substantial conception of society* than Rawls' cooperation view. In my view, a sufficiently ordered society, such as a modern state, combines three *communal features* that give rise to various demands of distributive justice (see Koller 2003). First of all, it is to be understood as an *ownership community* in regard to the members' common ownership of two kinds of assets, namely the society's natural resources on the one hand, and its cultural heritage on the other. Because of this fact, every member of society has a claim to a fair share of the benefits emerging from its natural and cultural goods (Steiner 1994). Secondly, a society is a *cooperation community* whose members are not only required to play by its rules for the sake of a peaceful social order, but also expected to partake in the production of social welfare, provided that they are capable of doing so. This is true of every society, but the extent in which a society embodies a cooperation community increases with its development, since members' mutual interdependence is growing with social differentiation and division of labour. Consequently, all members of society have a claim to equal access to its system of social cooperation and to a fair share of its benefits in consideration of their contributions. Insofar as individual members are prevented from realizing this claim through unavoidable effects of a generally beneficial socio-economic system, they have a derivative claim to appropriate compensation, as, for example, people who fall victim to unemployment in a market economy. And thirdly, societies are also to be regarded as *solidarity communities* whose members are mutually responsible for taking care of each other and providing those in need with appropriate support, so that everyone is insured against natural, social or individual contingencies, such as disasters, unemployment, sickness, old-age, and disability (cf. Walzer 1983, 64 ff.).

Thus, all members of society who are prevented from satisfying their needs to a sufficient extent have a claim to appropriate social support that enables them to lead a decent life. As a result, combining these three communal features gives rise to a rather demanding conception of distributive social justice that applies to all the primary social goods mentioned above.

For a first approximation to this conception, I suggest starting from a *general principle of distributive social justice* that is similar, though not identical, to Rawls' general conception. It runs as follows: All members of society ought to have an equal share of primary social goods, unless inequalities appear justifiable for reasons acceptable to every member from an impartial viewpoint. This principle, which simply results from applying the general basic demand of distributive justice to organized societies, leads to the question as to which reasons, if any, may justify social inequalities to what extent. Even though there are far-reaching disagreements on this question, most people think there are various reasons which, in principle, make social inequalities acceptable. The reasons that are usually advanced to that effect center around three general arguments: these are the arguments from desert, from liberty, and from need.

The *argument from desert* refers to the relative merits, achievements or contributions of individuals to the system of social cooperation. According to this argument, social inequalities are deemed to be justifiable, if they are in proportion to the individuals' relative merits, achievements or contributions (cf. Miller 1999, 131 ff.). Clearly, this argument appears plausible only under certain presuppositions: first of all, that there is sufficient agreement on the criteria by which people's performances are to be valued in order to determine their relative merits, achievements or contributions; and furthermore, that a larger share of the product of social cooperation is due only to those individuals whose performances generate a surplus value which is not taken by them alone, but is also to the benefit of all or most other members. One familiar argument why people exhibiting better or even extraordinary capacities may rightly claim a relatively larger share is the thesis these individuals need appropriate incentives that stimulate them to bring forth desirable or outstanding achievements to the general benefit. This thesis, which Rawls repeatedly maintains, may have some plausibility in view of the fact that, in social reality, people usually are more led by their selfish interests than by impartial considerations, but it is no good argument in the context of impartial moral discourse, as G. A. Cohen (1992; 2000, 117 ff.) made clear. Thus, appropriate reasons for privileging people whose performances appear particularly valuable should rather refer to the special burdens of these people, such as the tasks, expenses, costs, hardships, responsibilities and risks of their jobs.

According to the *argument from liberty*, social inequalities appear to be justified if they

emerge from free individual activities of people who all make use of their privately owned means within the limits of the respective rights granted to them by the social order. When, for instance, two siblings inherit the same amount of money, and the first uses her money for a secure investment, while the other carries his to the casino where he gambles it away, the second has certainly no reason to complain about the resultant inequality and to claim that the first ought to share with him her money plus returns. Of course, the validity of this argument also has limitations as it relies on the condition that the individual activities leading to inequality are framed by a societal order that, by and large, operates to the benefit of all members, regarded from an impartial point of view. The paradigm example of individual liberties which generate social inequalities, are those based on private property rights, but many other individual liberties also result in differences. It is fairly obvious that a social order granting such rights and liberties may be to the benefit of all members, for they are not only necessary for an efficient economic system that promotes social welfare, but also required for individual autonomy and self-determination. In this context, the reference to incentives is less problematic, since, here, the incentives mainly result from the individuals' opportunities to use their own resources to their benefit in a proper way (see Dworkin 2000, 120 ff.).

Finally, the *argument from need* maintains that inequalities are justifiable, if they arise from rendering appropriate support to fellow-members who would otherwise suffer deprivation, so that they are able to meet their needs to a sufficient extent. Note that providing support to these members amounts to treating them *unequally* rather than equally, for they get a larger share of supportive means than those who are required to contribute to the provision of these means without needing support. This argument is also subject to various constraints. First, the relevant needs must be conceivable against the background of the society's current standard of living; and furthermore, the extent of support to which members in need are entitled cannot be unlimited, but must be balanced against the sacrifices of those who have to render support. In the case of some unfortunate contingencies which people may suffer, such as requiring health care when falling ill, most people find themselves on both sides of the equation over the course of their lives, that is, both on the side of the needy as well as on the side of those contributing to the support of those in need. So, with regard to such cases, there may be prudential reasons for agreeing on an arrangement of mutual support after having balanced the benefits and costs, in order to insure oneself against the contingencies under consideration. But if we take an *impartial moral perspective* in which we have to imagine that we could be in the situation of any member of society, we are required to decide on a scheme of mutual support by balancing its benefits and costs for all cases of possible unfortunate contingencies that may concern us. Thus, from this perspective, we are required to choose a social insurance system under complete uncertainty of our own personal benefits and costs (see Braybrooke 1987;

Miller 1999, 203 ff.). Understood in this way, a balanced scheme of mutual support will be to the benefit of every member of society.

Even though the three arguments considered refer to different grounds that may make social inequalities generally acceptable, they have, upon closer inspection, a basic idea in common: this is the idea that social inequalities appear to be justifiable, if and only if they are to the *benefit of all members* of society, regarded from an impartial perspective. Thus, the basic idea underlying all three arguments may be put in the form a *general principle for justifying social inequalities* as follows: Social inequalities are justifiable, if and only if they are necessarily or unavoidably connected to a social order that is to the benefit of all members of society, including the worse-off. This principle is, by and large, equivalent with Rawls' first, unspecified version of his inequality principle leading to the DP.

Now, the question arises of how it may be possible to specify this relatively vague general principle in a way that leads to the DP in its modified shape according to the Priority View. In what way is it possible to move from the general principle's indeterminate and weak demand that inequalities are to be to the benefit of *every* member to the much more specific and demanding requirement that they ought to benefit the members of *less-advantaged* social groups to a *greater extent* the worse-off they are? In my view, this is not a very difficult task, for I think there is a number of well-grounded assumptions which together create a good case for the Priority View and, consequently, for the modified DP. These assumptions focus on the following facts: (1) the decreasing marginal value of increasing shares of primary goods, (2) the significance of relative differences between unequal social positions, and (3) the negative effects of socio-economic inequalities on the worth of equal liberty to people.

First of all, the assumption concerning the *decreasing marginal value of increasing shares of primary goods* means that one's benefit from possessing or using a portion of some particular primary goods gradually decreases as one's share of these goods increases. This assumption resembles the so-called "law of diminishing marginal utility", which has played an influential role in the history of economic thinking (cf. Samuelson/Nordhaus 2010, chap. 5). Yet, it differs from this 'law' in an important respect: While the principle of diminishing marginal utility, understood as an empirical thesis, is not generally true, because people's actual utility functions are usually much more complex and often even highly idiosyncratic, it appears highly plausible in the context of *impartial moral reasoning* where it refers to the fundamental interests of individuals in general rather than to their contingent individual preferences. Therefore, the assumption that the marginal value of an individual's share in primary goods decreases with the share's increase is a sound reason in favour of the Priority View in

general and the reading of DP according to this view in particular.

The second assumption, the *significance of relative differences between unequal social positions*, refers to the fact that, in social reality, people are not merely interested in the absolute size of their shares in social goods irrespective of how large the shares of others are. They also pay attention to the *relative* differences between their own shares in social goods and those of their fellows. Even if it may be expedient to construct the people in the original position behind the veil of ignorance as mutually disinterested persons striving only for their own shares in primary social goods in absolute terms, as Rawls suggests, there is no reason to prevent them from knowing the general fact that, in reality, most people also pay attention to their relative social positions and levels of well-being in comparison to others. Thus, social goods in general and primary social goods in particular are not merely means for the pursuit of the individuals' separate life-plans, but also *positional goods*, for which people compete in order to distinguish themselves from others or not to fall behind their fellows (cf. Hirsch 1976). Due to this fact, people who, irrespective of their absolute level, fare much worse than their fellows tend to feel deprived and wish to keep the relative differences between their own position and that of the better-off as small as possible. And as long as this disposition does not amount to an envious attitude which denies others any differential benefits even at the price of one's own disadvantage, it appears fully reasonable from an impartial perspective. After all, larger inequalities will affect the individuals' starting positions in subsequent competition and, thereby, increasingly undermine a fair distribution of social benefits. This reason for constraining social differences to the extent necessary for the increase of the welfare of worse-off people, also supports the Priority View and the modified DP.

Furthermore, the Priority View finds additional support in the fact that socio-economic inequalities exert *negative effects on the worth of equal liberty* for individuals, i.e. their actual capacities to exercise and enjoy their civil rights, individual liberties and political rights (cf. TJ, 201 ff.). Since these capacities greatly depend on people's respective socio-economic situation, the worth of equal liberty to individuals will diverge more, the larger the existing socio-economic differences are. In addition, growing socio-economic inequalities do not only diminish the worth of equal liberty to worse-off groups, but also create significant dangers for democracy, because well-off groups may use their social powers and economic means to influence the political decision-making process in order to advance their particular interests. Consequently, a just social order ought to keep socio-economic inequalities as small as possible by restricting them to the level at which they necessarily or unavoidably result from an institutional arrangement that raises the welfare of lower classes the more the worse-off their members are.

As a result, all these assumptions combined give strong support to the view that, within a generally beneficial socio-economic order that unavoidably generates inequalities, the benefits of less-advantaged members have increasingly more weight than those of better-off people the worse-off the former are. Consequently, this view gives rise to the demand that the socio-economic order of a society is to be arranged in a way that its emerging inequalities are not merely to some unspecified benefit of all members, but rather to the greater benefit of the less-advantaged groups, or, in other words, that the inequalities do not exceed the extent to which they are necessary to increase the life-time welfare of the worse-off members. So we arrive at the modified Difference Principle.

Bibliography

Barry, Brian (1989), *Theories of Justice*, London / Sydney / Tokyo: Harvester – Wheatsheaf.

Braybrooke, David (1987), *Meeting Needs*, Princeton, NJ: Princeton University Press.

Cohen, G.A. (1992), "Incentives, Inequality, and Community", in: Grethe B. Peterson (ed.), *The Tanner Lectures on Human Values*, vol. 13, Salt Lake City: University of Utah Press, 263-329.

Cohen, G.A. (2000), *If You're an Egalitarian, How Come You're So Rich?*, Cambridge, MA / London: Harvard University Press.

Dworkin, Ronald (2000), *Sovereign Virtue. The Theory and Practice of Equality*, Cambridge, MA / London: Harvard University Press.

Gauthier, David (1974), "Justice and Natural Endowment: Toward a Critique of Rawls' Ideological Framework", *Social Theory and Practice* 3, 3-26.

Harsanyi, John C. (1975), "Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls's Theory", *The American Political Science Review* 69, 594-606.

Hinsch, Wilfried (1998), "Rawls' Differenzprinzip und seine sozialpolitischen Implikationen", in: Siegfried Blasche / Diether Döring (eds), *Sozialpolitik und Gerechtigkeit*, Frankfurt / New York: Campus, 17-74.

Hirsch, Fred (1976), *Social Limits to Growth*, Cambridge, MA: Harvard University Press.

Koller, Peter (1983), „Rawls' Differenzprinzip und seine Deutungen“, *Erkenntnis* 20, 1-25.

Koller, Peter (1998), „Die Grundsätze der Gerechtigkeit“, in: Otfried Höffe (ed.), *John Rawls – Eine Theorie der Gerechtigkeit*, Berlin: Akademie Verlag, 45-69.

Koller, Peter (2003), „Soziale Gerechtigkeit – Begriff und Begründung“, *Erwägen Wissen Ethik* 14, Heft 2, 237-250.

Miller, David (1999), *Principles of Justice*, Cambridge, MA / London: Harvard University Press.

Nussbaum, Martha (2007), *Frontiers of Justice. Disability, Nationality, Species Membership*, Cambridge, MA / London: Harvard University Press.

Parfit, Derek (2000), „Equality or Priority?“ (1997), reprinted in: Matthew Clayton and Andrew Williams (eds), *The Ideal of Equality*, New York: St. Martin's Press, 81-125.

Pogge, Thomas (1989), *Realizing Rawls*, Ithaca / London: Cornell University Press.

Rawls, John (1971), *A Theory of Justice*, Cambridge, MA: Harvard University Press [quoted as TJ].

Rawls, John (1984/85), „Justice as Fairness: Political not Metaphysical“, *Philosophy and Public Affairs* 14, 223-251.

Rawls, John (1993), *Political Liberalism*, New York: Columbia University Press [quoted as PL].

Rawls, John (1999), *A Theory of Justice. Revised Edition*, Cambridge, MA: Belknap Press of Harvard University Press.

Rawls, John (2001), *Justice as Fairness: A Restatement*, ed. by Erin Kelly, Cambridge, MA / London: Harvard University Press [quoted as JF].

Samuelson, Paul A. / William D. Nordhaus (2010), *Economics* (1st ed. 1948), 19th ed. New York: McGraw-Hill.

Sandel, Michael J. (1982), *Liberalism and the Limits of Justice*, Cambridge: Cambridge University Press.

Schaller, Walter E. (1998), „Rawls, the Difference Principle, and Economic Inequality“, *Pacific Philosophical Quarterly* 79, 368-391.

Sen, Amartya K. (1970), *Collective Choice and Social Welfare*, San Francisco: Holden-Day.

Sen, Amartya K. (1974), „Rawls versus Bentham: An Axiomatic Examination of the Pure Distribution Problem“, *Theory and Decision* 4, 301-309.

Steiner, Hillel (1994), *An Essay on Rights*, Oxford: Blackwell.

Van Parijs, Philippe (1995), *Real Freedom for All: What (If Anything) Can Justify Capitalism?*, Oxford: Oxford University Press.

Van Parijs, Philippe (2003), "Difference Principles", in: Samuel Freeman (ed.), *The Cambridge Companion to Rawls*, Cambridge: Cambridge University Press, 200-240.

Waldron, Jeremy (1986), "John Rawls and the Social Minimum", *Journal of Applied Philosophy* 3, 21-33.

Walzer, Michael (1983), *Spheres of Justice*, Oxford: Blackwell.