Some Remarks on the Contractarian Justification of the Principles of Justice as Set Forth in John Rawls's Theory

Paweł Polaczuk

I. Introduction

In this paper I address two aspects of the theory of justice formulated by *John Rawls*.¹ These are the general and hypothetical natures of an agreement concerning the principles of justice. In relevant literature it is claimed that these two formal features (generality and hypotheticity) do not allow one to employ the normative dimension of the agreement argumentatively.² When analyzing the generality condition and the hypothetical nature of the agreement, I also refer to the situation in which rational individuals consent to the principles of justice set forth in the agreement and then to non-contractual circumstances. In the critical analysis of the objections expressed in the literature, I additionally introduce the so-called circumstances of justice. My analysis is of a theoretical nature. It focuses on reasoning schemes and structures as well as on the relationships between specific elements of the theory of justice. The meanings of particular concepts are reconstructed from *John Rawls's* deliberations.

To initiate my discussion, I need to recall that the generality condition and the hypothetical nature as well as the conditions of the agreement appear in the model of justification adopted in the theory of social contract³. The idea of social contract entered into by rational individuals is also the leading idea of *Rawls's* theory of justice. The aforementioned agreement is formulated by rational beings who are situated behind a veil of ignorance as concerns their actual living situation. This is a hypothetical situation in which the parties put forward suggestions of specific principles of justice and agree

¹ Rawls, Teoria sprawiedliwości [A Theory of Justice], 2009, 42 f.

² Kersting, Die Grenzen des Vertrages. Systematische Probleme der Vertragstheorie, in: Geiger/ Merle/Scarano (eds.), Modelle politischer Philosophie, 2003, 138 (140–141).

³ Some remarks concerning the fact that *Rawls's* conception belongs to the theories of social contract should be added. The most frequently mentioned criterion is the scheme of arriving at the agreement upon the principles of justice for the basic structure. This refers to the procedure in which agreement on the content of the principles of justice is reached. A broader understanding of this criterion indicates that his conception belongs to the theories of social contract is proved by the leading thought of the conception of justice, according to which political institutions will be considered as just and will deserve recognition if they are shaped according to the notion of rationally acting citizens, adopted under hypothetical and reasonable contract conditions. Relevant literature indicates one more, supplementary or additional criterion. It is stated that *Rawls's* conception belongs to the theory of political justice must be a means of conveying political liberalism. However – what is extremely important – it should refer to the historical tradition of modern constitutional democracies and cannot include or express premises referring to citizens' activities beyond the political sphere. I mention this so that there is no doubt that not each theory that copies the mentioned argumentation scheme must necessarily be a means of conveying political values of liberalism.

upon their choice. Ignorance as to their actual living situation leads them to accept the principles that are the subject of the theory of justice. The agreement is binding for those individuals who entered into it. Thus the principles of justice agreed upon under such circumstances and in such a form would be in effect in a particular society after the agreement has been made, when the parties become aware of their factual social positions and the convictions about justice determined by these positions. The agreement on the principles of justice is at the same time general. It concerns the principles of justice adopted as the foundation of the social order rather than particular institutions. It is supposed to be the basis for the reconstruction of a chain of agreements adopted for specific institutions. Consequently, it is an agreement of a very high hypothetical level.⁴

II. Generality

The previous paragraph provides a very approximate definition of the generality of the agreement concerning the principles of justice. This agreement refers only to the principles and these principles are the foundation of the social order. A more thorough explanation of this element needs to embrace the premises of the generality of the agreement and the theory of justice as well as the consequences of the generality condition for the hypothetical nature of the agreement and for the circumstances in which the parties enter into it. In the context of the first mentioned element, one should draw attention to *Rawls's* assertion that the departure points for the theory of justice are intuitive convictions that each person holds. These convictions are of a general nature and contribute to the perception of justice in a way that is open to interpretations. Thus intuitive convictions of justice can be formulated into an agreement-based conception of justice that will reconcile different convictions and will justify a set of interrelated principles.⁵ The principles of justice are to reflect these intuitions, while preserving the necessary generality. Moreover, Rawls connects the generality condition with the view that justice is currently the first virtue of social institutions. This emphasizes the consequences of the previous social changes that revealed individuals' expectations, motivated by a sense of justice, as to social institutions. In this sense, the quoted view concerns the primacy of justice among other values that lie at the foundation of the principles of cooperation between society members. The view stressing the primacy of justice results in the general nature of the theory of justice. Due to this feature, the principles of justice may become the foundation of the organization of social institutions. The relationships between these premises of generality of the contract theory, that forms the core of the author's theory, and the generality of the principles of justice, are revealed by Rawls in his argumentation concerning the reception of the idea of social contract. This contractual idea can serve to justify the entire ethical system that would embrace principles for

⁴ See also *Polaczuk*, Teoria sprawiedliwości Johna Rawlsa (zarys), in: Kruszewska/Polaczuk/Swito (eds.), Sprawiedliwość. Wybrane koncepcje, 2010, 61 (69–86).

⁵ Ibid., 64.

Whether a given theory of social contract will be included into the historical tradition of modern constitutional democracies is a very complex matter and largely depends on the adopted perspective. For some researchers rights and liberties will be of key importance, for others majority rule, and for the rest the rule of law. See further *Besch*, Über John Rawls' politischen Liberalismus. Zur Rolle des Vernünftigen in Rawls' Begründung einer politischen Gerechtigkeitstheorie, 1998, 9 f., and *Rinderle*, Politische Vernunft. Ihre Struktur und Dynamik, 1998, 25–34.

Paweł Polaczuk

all the virtues. Due to the generality of intuitive convictions concerning justice and the assertion that justice is the first virtue of social institutions, the reception of the idea of social contract refers exclusively to the principles of justice. This is how I understand *Rawls's* claim that limiting his contract theory to the principles of justice is connected with the general type of view that the theory of justice exemplifies.⁶

As a result, *Rawls* generalizes the theory of justice and social contract and transfers the concept of contract to a higher level of abstraction. Generality is materially expressed in the generality of the principles defining the fair terms of peoples' associations. These principles are to regulate all further agreements; they specify the kinds of possible social cooperation and institutions.⁷ In other words, the principles of justice must be first of all sufficiently general, if they are to serve as the foundation of the entire system of cooperation in which members of society function.⁸ Only general principles can be the basis for a multitude of social institutions. They can further become more detailed when assenting on particular social institutions.

III. Hypotheticity

Moreover, the principles must allow for criticism and reform of institutions from the point of view of justice. The criticism and change of the hitherto rules of cooperation requires the assessment of how these principles function. *Rawls* claims that such an assessment is not possible without the analysis of how they were employed in practice. For the sake of this analysis, hypothetical agreements that are the sources of the rules shaping this practice are reconstructed. The reconstruction begins with the agreement entered into in the original position in which general principles of justice are determined.⁹ In order for the principles of justice to become the foundation of the social organization, in *Rawls's* understanding, they must be agreed upon in a hypothetical agreement. Otherwise, the assessment of how the rules of cooperation function will not be possible, because it will be impossible to reconstruct the hypothetical agreements concerning these rules. The agreement in which the principles of justice are determined is the basis for this reconstruction. This agreement should be of a more hypothetical nature than the agreements reconstructed with a view to assessing how particular institutions function. In order to achieve that the conditions of the agreement on the principles of justice must be obviously formulated in a more abstract manner than the circumstances that shape the situation of the subjects who decide upon the structure or competences of the organs implementing specific rules in practice. The agreement entered into in the original position determines the binding basis of the terms of social cooperation.

Emphasizing the internal structure of the agreement, rather than its form, one may state that the circumstances in which the principles of justice are agreed upon must be

⁶ Rawls, Teoria sprawiedliwości [A Theory of Justice], 2009, 48.

⁷ Ibid., 40–42; *Pogge*, Hypothetische Gesellschaftsveträge: Drei Schwierigkeiten, in: Geiger/Merle/ Scarano (eds.), Modelle politischer Philosophie, 2003, 117–118: institutions that *Rawls* writes about are in fact particular rules of cooperation; they can refer to, e.g., ownership, division of labour or competition between the subjects of social interactions.

⁸ This is a condition for their choice, rather than those principles postulated by other conceptions, the condition of their fairness and a common knowledge of the principles which are consequently obeyed.

⁹ Rawls, Teoria sprawiedliwości [A Theory of Justice], 2009, 41.

so abstract that the agreement entered into in the original position could justify a validity of the principles of justice expressed in it as the basis of the system of cooperation and assessment of its institutions. This is how I understand *Rawls's* assertion according to which the theory of justice generalizes the theory of social contract and transfers the notion of contract into a higher level of abstraction.

IV. The hypothetical nature of the agreement as an obstacle for the agreed upon principles to become binding

I need to begin my critical analysis of the justification adopted by *Rawls* with two remarks. First, my discussion is limited to a short characteristic of the formal features of the basic part of this justification. These features refer to the construction of the agreement on the principles of justice, shaped on the model typical of the theory of social contract. Thus, critical conclusions of the following analysis will refer to the theory of justice. Second, when investigating selected formal features of the construction of the agreement on the principles of justice one cannot forget that the idea of social contract is, at least theoretically, to serve the justification whose cogency is closest to the trust bestowed on intuitions concerning justice. I mention this because the dubiousness of the construction adopted by *Rawls* stems from the limitations that arguments based on social contract are subject to. In relevant literature it is claimed that a hypothetical social contract is a construction that makes it impossible to achieve its aims. The consequences of the agreement are conceived analogously to agreements entered into in actual social life. Yet only those contracts that are actually concluded and only those pledges that are factually made can be considered as binding. On the other hand, actually concluded contracts are not homogenous enough to treat the agreement as a source of justification for commonly binding principles. It is concluded *a contrario* that the idea of social contract cannot employ the normative dimension of the agreement argumentatively.¹⁰ These objections refer in particular to the theory of justice. Rawls elevates the theory of social contract to a higher level of abstraction as the previous paragraphs explain.

It should be emphasized that the criticism of the construction of social contract concerns the hypothetical nature of the agreement. Such arising doubts may be formulated into the question as to whether hypothetical agreements on the principles of justice can become the source of actual obligations, i.e., the foundation of rights and duties in the existing social structures. It is concluded that it is not so because actually concluded contracts that determine the rights and obligations of the parties have different contractual conditions and are effected in various non-contractual circumstances (e.g. factual, personal, social, individual). Thus it is not accurate to refer to being bound by a contractual relationship as a source for the common validity of the principles. The crux of the matter is not only that of making the conditions of the agreement on commonly binding principles homogenous, but also that of the circumstances that are in a way external to the contractual conditions. Without them, it is impossible to reinforce rights and duties resulting from the principles derived from hypothetical contractual conditions. In simpler words, based on the principles that we consented to in a situation that was only thought of (hypothetical conditions), but which we have not found ourselves

¹⁰ Kersting in: Geiger/Merle/Scarano (eds.), Modelle politischer Philosophie, 2003, 140-141.

in, it is impossible to determine rights and duties that are valid or obliging us in a specific social reality¹¹.

V. The analysis of the objection to the theory of social contract on the grounds of Rawls's theory of justice

Two remarks open this analysis. First, it needs to be stated that a critical argument concerning the idea of social contract in general refers also to the theory of justice. It may be deemed that if Rawls transfers the theory of social contract to a higher level of abstraction, the objection formulated by critics pertains in particular to the discussed theory of justice. The second remark concerns the assumptions that were adopted to formulate this objection. Two significant threads of the argumentative scheme based on the idea of social contract emerge in this objection. In this scheme, principles become binding as a consequence of a hypothetical agreement between the parties and this agreement is entered into under the assumed and homogenous circumstances. From the perspective of general reflections, critics of the theory of social contract are right. Some reservations need to be made, however, if these remarks are referred to Rawls's theory of justice. Applying this objection to the discussed theory of justice assumes a comparable diversity of contractual conditions and circumstances in which intuitions about justice are shaped. This is, however, a simplification that actually emerges from *Rawls's* view itself. He states that a theory of justice based on the idea of contract aims at justifying intuitive convictions of justice. If we limit ourselves to the main idea of the theory of justice, that is to the theory of social contract, then we must acknowledge that this justification is afforded by a hypothetical situation in which rational individuals agree upon the principles of justice. In other words, the simplification found in the quoted argument is rooted in the design of justification of the views that *Rawls* links with intuition shaped in diverse conditions. However, circumstances in which intuitive convictions of justice are shaped are transferred to the elements of the theory created for the sake of determining the subject of the principles of justice in such a way as to integrate a possibly wide spectrum of problems that appear in society. What I mean here is primarily the so-called basic structure of society. This scheme of thinking originates in a conviction that the source of preferences for the principles that are to become the basis for social cooperation must be specified.¹² Rawls's claim of the primacy of justice, already quoted here, indicates a need to justify the principles of justice in such a way that they should be recognized as a better alternative in some respects to principles postulated by other conceptions. The mentioned basic structure of society is meant to correspond to the actual conditions in which intuitive convictions about justice are shaped. These conditions are transferred from the level of their correlation with intuition to a theoretical sphere in which the subject of the principles of justice postulated by Rawls is shaped.¹³ The basic structure of society does not characterize specific situations in which various social demands are raised. Such demands preclude the formulation of the

¹¹ Ibid., 141.

¹² The primacy of justice is of a general nature with respect to intuition. With regards to the theory, one needs to demonstrate that the postulated conception of justice would be chosen out of many other ones.

¹³ Rawls, Teoria sprawiedliwości [A Theory of Justice], 2009, 157–158.

principles of justice. The basic structure is created by general and objective conditions of social cooperation, reduced to the level that determines the circumstances deciding on life prospects. It should be stressed that this refers to universal conditions of a structural nature.

The outcomes of the foregoing deliberations can be reduced to a claim that a conviction of the primacy of justice creates a theoretical problem of how to choose the principles postulated by *Rawls* from among other, alternative conceptions. The issue of choosing the principles of justice justifies the integration of specific social conditions within this theory. Hence it is difficult to agree with the opinion that a diversity of circumstances is reduced exclusively to hypothetical conditions. It should be stressed, however, that the imprecision in the criticism of the theory of justice indicated by me is actually justified by *Rawls's* line of reasoning in which the idea of social contract becomes the main idea of the theory of justice. To conclude, an amplified hypothetical nature of the agreement on the principles of justice is balanced by objectified conditions in which intuitive convictions about justice are shaped.

1. Circumstances of justice as an argument in the analysis of the objection based on the hypothetical nature of the agreement

In the foregoing discussion it is claimed that the conditions in which the agreement on the principles of justice is made are more abstract that the ones that shape the situation of the subjects who decide on the functioning of various organs that participate in the social cooperation in various ways. I also presented the opinion expressed in relevant literature according to which the agreement on the principles of justice cannot be employed to justify the principles of justice, because it is only a hypothetical contractual obligation. This view was supported by the non-factual coherence of the contractual circumstances (as well as the subject of the agreement itself). This argument proved to be too categorical and simplified. The presented view should be complemented by one more argument: the reasons that allow one to recognize that rational subjects entered into an agreement on the principles of justice. These reasons are part of the theory of justice itself.¹⁴ They can be exemplified by the so-called circumstances of justice that *Rawls* refers to the original position, in other words, the situation in which individuals suggest particular principles of justice and make their choice.¹⁵ The circumstances of justice are necessary so that the original position could be interpreted in a manner allowing for choosing always the same principles.¹⁶ The description of the circumstances of justice facilitates a claim that this element of the theory can replace specific circumstances in which intuitive convictions about justice are shaped. Rawls argues for this element of the theory, stating that the knowledge of specific circumstances concerning one's society would complicate the initial situation in which the principles of justice are agreed upon.¹⁷ These particular circumstances are diversified, and some of these are of an arbitrary and contingent nature. Consequently, reaching an agreement as to the prin-

¹⁴ Ibid.: according to *Kersting*, this mode can be considered in epistemological and moral dimensions.

¹⁵ *Rawls*, Teoria sprawiedliwości [*A Theory of Justice*], 2009, 184: the original position precedes the initial situation that is its philosophical interpretation. *Rawls* also claims that this interpretation is the result of a hypothetical course of reflection.

¹⁶ Ibid., 199, 212.

¹⁷ Ibid., 209.

Paweł Polaczuk

ciples of justice would be impossible¹⁸ without negotiations whose outcome will be always from some point of view insufficiently convincing. Intuitive convictions about justice assume that the rights are secured by justice and are not subject to political bargaining or to the calculus of social interest.¹⁹ The circumstances of justice are then to some extent argumentative indicators that become substitutions for particular conditions connected with intuition. The basic structure of society serves this function with

respect to the indications based on fundamental and structural conditions that refer to the choice of the principles postulated in *Rawls's* conception. Indications contained in the theory of justice would then serve as arguments making it possible to convince anyone to consent to the principles adopted by the subjects placed in a particular situation.

The explanation of the conduct of the subjects who agree upon the principles of justice will be comprehensible on the basis of the arguments contained in the elements of the theory. The aforementioned arguments are meant to correspond to the circumstances in which the views about justice are shaped, but they are objectified for the sake of the function they serve in the adopted scheme of justification. It should be stressed that if the structural social circumstances are transferred to the basic structure of society, then the moral context of the convictions of justice is formed by the circumstances of justice. *Rawls* claims that the parties of the agreement are not bound by prior moral ties to each other.²⁰ He further states that the account of the conditions of justice

"involves no particular theory of human motivation. Rather, its aim is to reflect in the description of the original position the relations of individuals to one another which set the stage for questions of justice."²¹

It should be added the persons situated in the original position know that the circumstances of justice obtain.²²

2. The problem of the relationship between the circumstances of justice and moral needs. Conditions imposed on the concept of right

The foregoing argument against the objection raised by *Kersting* is not exhaustive. It requires the relationships between the formal circumstances of justice and moral needs shaped in social relations to be analyzed. This problem actually refers to whether these circumstances correspond sufficiently to moral needs. The answer to this question is negative. Although *Rawls* claims that the circumstances of justice reflect relations of individuals to one another, he in fact limits the need to determine these circumstances to two factors. He states that men are not indifferent as to how benefits are distributed and they want cooperation that makes possible greater benefits than any would have if each were to try to gain them by his own effort²³. The evident reduction of moral needs to the benefits of cooperation raises most doubts. The scheme of rational behaviours of identical subjects placed in the original position is a correlative of this reduction. Their knowledge is constrained to a specific set necessary to assent to beneficial principles of

- ¹⁹ Ibid., 30.
- ²⁰ Ibid., 197.
- ²¹ Ibid., 199.
- ²² Ibid., 197.
- ²³ Ibid., 195.

¹⁸ Ibid., 214.

justice, rather than those that correspond to moral motivations differentiated due to the life circumstances of particular individuals.

However, *Rawls* introduces one more element that is meant to justify the circumstances of individuals situated in the original position. He claims that their knowledge and alternatives as to the choice of principles open to them are subject to constraints resulting from the conditions imposed on the concept of right. He claims that fulfilling these conditions allows for eliminating egoism from human behavior, and so he derives the conviction of the "moral force" of constraints of the concept of right. The propriety of these conditions is evident in adjusting the claims that persons make on institutions and one another.²⁴ It should be added that conditions imposed on the concept of right are objective because they apply when all ethical principles are chosen. In this sense they form specific moral frames for the circumstances of the principles of justice. The already discussed circumstances of justice also contribute to them.

Out of the five conditions imposed on the concept of right three are of key importance. The first condition is that of generality. Its fulfillment allows for formulating the principles without the use of proper names, personal pronouns and the so-called rigged definite descriptions. Then the principles express general properties and relations rather than the egoistic needs of the individual. It should be added that the condition of generality corresponds to the idea that principles are to be universal in application, that is they must hold for everyone in virtue of their being moral persons. Thus they will not be framed in such a way that based on their wording it would be possible to limit their application to a restricted class of individuals, singled out by special biological or social characteristics. Thus socially unethical behaviours, motivated by a concern to strengthen the position of particular groups, whose members possess these particular characteristics, are eliminated. The universality of the principles results in the standardization of life circumstances of the individuals. *Rawls* treats their rights and duties as the consequence of the first principles that hold for all.

Finally, the condition of an ordering on conflicting claims springs directly from the role that the principles of justice are to play. Here another reservation should be made. *Rawls* explains that he does not mean a complete ordering, that is such that would be able to order all the claims that can arise in practice. It is only about such claims whose ordering can be established based on the criterion of relevance for persons and their situation which are independent from their social position, or their capacity to intimidate and coerce.²⁵ The fulfillment of the ordering condition makes it possible to exclude the so-called general egoism that is revealed in that everyone is authorized to do whatever he wants to advance his aims or interests, according to his judgment.²⁶ The condition of an ordering of conflicting claims on the basis of the principles that fulfill this requirement is a significant obstacle for the egoistic attitudes of society members.

²⁴ Ibid., 200–201.

²⁵ Ibid., 201–202, 204–205.

²⁶ Ibid., 207.

VI. The ethical dimension of the agreement as a motivation for being bound by the agreement

Finally, the ethical dimension of the agreement should be considered. I need to recall that *Kersting* negated that the individuals are bound by a hypothetical agreement, referring to the analogy with contracts concluded in practice. Such contracts differ from a hypothetical agreement with respect to both contractual conditions and non-contractual circumstances. In my view, these differences are eliminated at the theoretical level, although not to the full extent determined by the analogy to contracts concluded in practice. Consequently, some conditions of the agreement should be searched for in its ethical dimension.

It should be stated first that conditions imposed on the concept of right refer to the restrictions concerning the knowledge of individuals in the original position. They can also direct the choice made in such a way as to eliminate the alternatives that do not fulfill the discussed conditions. Such alternatives would be the source of an unethical organization of society that would be subordinated to the aims of one person or a particular group. It should be stressed, however, that according to the foregoing remarks, conditions imposed on the concept of right justify the circumstances of the subjects in the original position. The mentioned frames of moral circumstances of the principles of justice create non-contractual conditions. This refers also to the circumstances of justice. Critics claim that an agreement entered into in the original position is not autonomous.²⁷ Arguments in favor of a morally right choice made by the individuals placed in such a position come from the theory of justice. This can be exemplified by two elements of the theory discussed here: the circumstances of justice and the conditions imposed on the concept of right. Together they set the frames of moral circumstances for the principles. Argumentation contained in them is to enable the understanding and acceptance of the explanation as to the behavior of the subjects who decide on the principles of justice in a contractual situation.

It should be added, however, that the aforementioned construction eliminates from the justification moral convictions that are shaped in diverse life situations. What I mean is the standardization of the individuals' life circumstances and their claims. It is revealed that the scheme adopted by *Rawls*, based on arguments external with respect to the agreement, does not incorporate moral convictions shaped in diversified life situations. This ascertainment is not irrelevant for the justification of the principles and their application. The range of justification based on the contractual model is then limited, whereas the frames of moral circumstances for the application of the principles of justice are formed by non-contractual conditions. Consequently, this that would endow the agreement with an ethical dimension cannot be justified with the contract model, while from the theoretical perspective of justification – the agreement is of secondary importance. The moral dimension of the agreement is derived from the normative elements of the theory, following the interpretative action.²⁸

²⁷ Kersting in: Geiger/Merle/Scarano (eds.), Modelle politischer Philosophie, 2003, 142.

²⁸ Ibid., 142–144: *Kersting* also enumerates assumptions referring to equality and human rights.

VII. Conclusions

In this paper I discuss some objections raised with respect to the justification model based on the idea of social contract. These objections focus on the hypothetical nature of social contract. The character of the agreement would make it impossible to accept the assumption that it is binding for the subjects also in conditions other than those in which it was concluded. This objection supposedly also refers to *Rawls's* theory of justice, whose principal idea is the idea of social contract. In my analysis I have explained the assumptions, notions and complex relations from which the author of the theory of justice derives the hypothetical nature of the agreement on the principles of justice. Then I have indicated those elements in his theory (the circumstances of justice and conditions imposed on the concept of right) in which Rawls places the conditions, such as the conditions of contracts entered into in practice and their non-contractual circumstances. Such are similar to the conditions determined by the aforementioned analogy with contracts concluded in practice. Due to these elements of the theory of justice, I believe the assertion that a hypothetical agreement cannot serve at all as the source of justification for commonly valid principles of justice to be a bit previous and too categorical. Yet, it should be deemed inaccurate only if the agreement had an ethical dimension. The indicated elements of the theory of justice do not reveal, however, relevant relations with moral needs (motivations) of the individuals. The analysis of *Rawls's* line of reasoning has revealed the specificity of the moral argumentation in favour of the principles postulated by him. The author of the theory of justice reduces this argument to the benefits acquired by the subjects who agree on the principles and placed outside the contractual situation.

> Paweł Polaczuk, Uniwersytet Warmińsko-Mazurski w Olsztynie, E-Mail: pawel.polaczuk@uwm.edu.pl