The role of the sense of justice in Rawls' theory

O lugar do senso de justiça na teoria de Rawls

PABLO AGUAYO WESTWOOD

Abstract

In this paper, I discuss the relevance of the sense of justice for Rawls's theory of justice. In the first part, I analyse the interpretation that Rawls offered of moral psychology for evaluating its role in the generation of the sense of justice and the acceptance of the principles of justice. In the second part, I discuss the difficulties of accepting Rawls's proposal. Finally, I conclude that the research on moral psychology developed by Rawls allows him to recognise a way of supporting his principles of justice that is different from the traditional forms of deductive justification.

Keywords: Social justice. Social stability. Moral Psychology. Sense of justice.

Resumo

Este artigo analisa a importância do sentido de justiça para a teoria da justiça de Rawls. Na primeira parte, analiso a interpretação da psicologia moral oferecida por Rawls para avaliar seu papel na formação do sentido de justiça e na aceitação dos princípios de justiça. Na segunda parte, discuto as dificuldades para aceitarmos a proposta de Rawls. Finalmente, concluo que a abordagem em psicologia moral desenvolvida por Rawls permite-lhe reconhecer um modo de defender seus princípios de justiça que difere de formas tradicionais de justificação deductiva.


1 This paper is part of the project FONDECYT nº 11170230.

a Universidad de Chile, Santiago, Chile. PhD in Philosophy, e-mail: paguayo@derecho.uchile.cl
Introduction

It will be useful in understanding why the capacity for a sense of justice is the fundamental aspect of moral personality in the theory of justice.

John Rawls

These, it seems, supplement or substitute for what was more often in Theory called the “sense of justice.” These “very great virtues” are, as specified in various places: reasonableness and a sense of fairness, a spirit of compromise and a readiness to meet others halfway, toleration and mutual respect, and a sense of civility. Rawls stresses frequently how important it is that citizens of a well-ordered society develop these virtues.

Susan Moller Okin

One of the main concerns of the Rawlsian theory of justice was to determine the moral and political conditions that would make social stability possible. In this search, Rawls considered several strategies to justify our social practices through different moral principles and criteria. As he showed in “Justice as Fairness” (1957), and “The Sense of Justice” (1963), his reflections were strongly influenced not only by the analytical philosophy of his time but also by logical positivism. The above does not mean that Rawls assumes some of the emotivist positions typical of the Anglo-Saxon moral philosophy of the mid-twentieth century. Instead, his search was aimed at finding — following the title of Stephen Toulmin's work — a place for reason in ethics. In his discussion of the conceptual framework offered by logical positivism, Rawls deeply examined our moral psychology. In this search, the sense of justice appears as a central element in the explanation of our moral behaviour. For Rawls, our sense of justice would ensure “the integrity of the agreements reached in the original position” (RAWLS, 1971, p. 145), as well as the strict compliance of the principles that emanate from it.

Rawls said that social stability depends in great measure on the possibility that people can develop their sense of justice in a normal way and can support and maintain their basic social institutions. But, what is the psychological mechanism...
that drives individuals to develop their sense of justice as the basis of their moral actions? For Rawls, the answer to this question relies on a particular way of understanding our moral psychology in combination with certain principles of justice. It is because of the above that in this paper I analyse the interpretation he offers of our moral psychology to evaluate (i) its role in the acceptance of the principles of justice and (ii) the reasons and difficulties of founding social stability on certain moral feelings. Before developing the points mentioned above, I offer a framework for discussion so that Rawls’s ideas and arguments can be evaluated from the context in which they emerged.

**Context of the discussion**

In her article Modern Moral Philosophy of 1958, G. E. M. Anscombe extended an invitation to moral philosophers to remain in silence until they have an adequate philosophy of psychology. I can summarise the thesis defended by Anscombe as follows: moral philosophy requires that moral concepts that make up the core of your reflection be considered beyond conceptual analysis and receive due treatment in the field of moral psychology; only in this way, Anscombe thought, would moral philosophy have significant effects. Years later, in an Inaugural Lecture delivered at Bedford College, Bernard Williams denounced the simple philosophical development that moral emotions had had in recent moral philosophy. Among the reasons identified by Williams to support the

---

1 I take as a basis for my discussion the article “The Sense of Justice” of 1963 and those parts of *A Theory of Justice* (hereinafter TJ) in which Rawls addresses these same issues. The ideas embodied in 1963 reappear almost unchanged in several sections of TJ and *Political Liberalism* (hereinafter PL). In its original edition, the sense of justice is discussed on pages 12, 19, 45-50, 120, 145, as well as throughout chapter VIII, entitled “The Sense of Justice”. When Rawls refers to the sense of justice in the context of the original position is when he states that “his sense of justice move him to act on the principles of right that would be adopted in the original position” (1971. P. 45). In the reedition of 1999, the sense of justice appears on page xii of the new preface with a central role for the Rawlsian project. With the capacity for a conception of good, the sense of justice is considered as one of the two moral powers of human beings. Finally, in PL Rawls addresses the sense of justice primarily in the "Idea of Public Reason Revisited". In that chapter, Rawls considers the sense of justice as an element associated with civic development and as a result of processes associated with family relationships that make up the basic structure of society.
previous thesis, firstly, I can cite a particular way of understanding moral philosophy as a reflection about moral language —here lies the same intuition of Anscombe. Secondly, Williams identified the combination of two other things: a relatively naive view of emotions together with a Kantian vision of morality (see WILLIAMS, 1965).

The previous references constitute only two milestones of the profound debate that took place in the Anglo-Saxon moral philosophy during the decades in which Rawls was elaborating his moral philosophy. Although Rawls did not use conceptual analysis as a philosophical tool, his methodology was strongly marked by the elements of this tradition. Proof of the above is his article “Outline of a Decision Procedure for Ethics” (1951a), in which he sketched a reasonable procedure for deciding between conflicting moral judgments, a procedure that should be established by rational methods of investigation. Another point that supports the above is the review Rawls wrote of S. E. Toulmin’s book: *An Examination of the Place of Reason in Ethics* (1951b). In it, Rawls explained that the central enterprise of ethics is to explain the soundness of moral reasoning.

However, not only the analytical tradition influenced Rawls’s reflections on morality. As recent literature has shown (see BEVIR & GALIŠANKA, 2012), in his early work, Rawls rejected the reduction of ethics to empirical-observational statements and offered an essential place to moral feelings. During the next decade, he prepared and delivered several seminars on these topics, perhaps aware of the warnings of Anscombe and Williams. For example, in the autumn of 1958 Rawls offered a seminar on moral psychology where he showed the influence Wittgenstein had on him and in which he identified those points where he followed the analytical tradition. All of the above would be reflected in his article “The Sense of Justice” (1963).

**Moral Psychology**

In “The Sense of Justice” Rawls offered an interpretation of persons as rational beings with their own ends and capable of a sense of justice. Likewise, he
understood the meaning of justice as a moral capacity that would enable both the fulfillment of our agreements and the respect for the principles that organise the basic structure of society (see RAWLS, 1963, p. 291-293). But what would the sense of justice ensure? And more importantly, how would this be achieved? For Rawls, the sense of justice would enable and ensure both the integrity of the agreement reached in the original position and that the principles of justice were respected (see RAWLS, 1963, p. 300-301; 1971, p. 145). Likewise, the lack of a sense of justice would be associated with the lack of certain features that characterise our humanity. For Rawls “one who lacks a sense of justice lacks certain fundamental attitudes and capacities included under the notion of humanity” (1971, p. 488).

The reason why the agreements reached in the original position are respected can be understood from two perspectives. According to the first one, the sense of justice, in its Kantian version, would make it possible for persons to consider each other not only as means but also as ends in themselves. Given the above, it would be morally unacceptable to defend principles that benefit them exclusively when this implies unjustified damage to others. A second perspective, in its psychological understanding, considers that the sense of justice would function as a principle of moral constriction of the will that would cause the principles of justice to be respected. In this line of argument, Rawls pointed out that: “We tend to feel guilty when we do not honor our duties and obligations, even though we are not bound to those of whom we take advantage by any ties of particular fellow feeling” (RAWLS, 1971, p. 474). For Rawls, the feeling of guilt would constrain our will, leading us to the fulfilment of our agreements.

In this context, Rawlsian moral psychology understood as the psychology of our moral feelings, place its attention on the development of our competencies as rational subjects, especially in those capacities that he considers as the essential features of moral persons. Amongst these characteristics, the capacity to have a sense of justice is one that the author discusses in detail. In other words, if both the capacity to propose and seek an idea of good, and that which allows us to have a sense of justice require an explanation of the psychological structure, it is the second feature that demands a detailed explanation of the principles that regulate our moral psychology.
As I have noted, these principles of moral psychology are introduced explicitly with the function of explaining and justifying certain behaviours that would promote social stability. In argumentative terms, Rawls makes an inference of the best explanation. Faced with the challenge of explaining the reasons that moral subjects would have to respect the principles of justice and ensure social stability — and rejecting at the same time any explanation based on metaphysical criteria — he considers that specific psychological principles would allow the development of the sense of justice and the respect for the highest social institutions.

In summary: (1) we need an explanatory hypothesis that shows the reasons why persons would have to respect and fulfil their agreements; (2) this hypothesis cannot appeal to the utilitarian principles; (3) neither this hypothesis should appeal to an apriori principle of morality with metaphysical character and (4); thus, the best explanation is to accept specific natural psychological laws that explain, at least in descriptive terms, the reasons we would have to respect our agreements. What these laws are and to what extent they could justify our moral behaviour is what I discuss in the next section.

**Stages of moral development**

The central thesis that describes the idea of moral development defended by Rawls is that this process is related to the appearance and development of the sense of justice. For Rawls, the youngest members of society gradually acquire a sense of justice to the extent that they are part of the different social instances that characterise their development. In what follows, I explain these stages to determine the role that moral feelings play in Rawls’ moral psychology, as well as their relevance to the stability of society.

**First stage: the morality of authority**

This stage is characterised by the relationships established within a family context in which the parental figures determine the behaviour of children, all of the
above surrounded by a moral framework. The first stage of moral development is determined by the first psychological law, which affirms that when the institution of the family is itself fair, children gradually develop a loving attachment towards their parents. The cause of this attachment lies primarily in the care parents have given their children. As a result of the above, the sense of self-esteem and self-respect develop in children. Over time, the parents' love and trust lead the child to follow and accept the mandates they hold. This situation would not be the product of the fear of a sanction—or for fear of losing love, as Freud would say in *Civilization and its Discontents*—but for the love and respect they have for their parents, as well as the desire not to disappoint them. In this way, the morality of the authority would allow us to be willing, without expectation of reward or punishment, to follow certain precepts that would not only seem arbitrary but also in no way correspond to our original inclinations (see RAWLS, 1963, p. 287-288, 1971, p. 466).

For Rawls, the morality of authority responds to the psychological principle according to which in our childhood we generate feelings of love and trust towards our parents when they have shown love and concern. Rawls said that love and trust are the causal elements that favour the appearance of feelings of guilt in the case of disobedience to orders and mandates that come from parental authority. That is why the absence of feelings of guilt would reveal a lack of love and trust from our parents. It is important to note that in this period of learning, infants do not have critical thinking and thinking skills that would allow them to have reasons to accept or reject possible mandates. Their participation in the moralisation process is rather passive.

**Second stage: the morality of the association**

This stage is characterised by the learning and development of emotional bonds connected with moral standards of specific social roles—like friendship, companionship, neighbourhood, among others—that would correspond to the different associations in which we participate. By belonging to these associations, the adolescent is guiding his actions by a series of ideals and principles that
characterise these practices, in turn learning the virtues of the character of each of these roles. One of the essential features of this stage is that, to the extent that adolescents begin to understand the different associative schemes, they acquire a conception of the cooperation system that defines the associations and the purposes they promote. In this sense, the learning acquired by the subjects in their different roles—or in the fulfilment of their different roles—is central to their moral formation and functions as one of the first approaches to civic virtues, typical of those subjects who have acquired the sense of justice. Observing the different roles, as well as the division of functions and duties of a cooperative system, the adolescent would gradually develop a more complex vision of associative and collaborative work. They could understand and differentiate the various issues in which are committed regardless of their position, as well as their role. The foregoing would enable an even more significant step, namely, being able to empathetically understand the reasons why others are conducted in and out of the game. Empathy would then depend on both the development of self-esteem and the ability to know and appreciate within a system of cooperation, as well as the recognition that other persons go through a similar process.

For Rawls, this stage would correspond to the manifestation of a second psychological law according to which, if the partners fulfil the obligations that their roles demand, then feelings of friendship, loyalty and trust will be generated in the individual. As a consequence, the failure to fulfil their role would result in the appearance of feelings of guilt. These feelings would be manifested both in the attempt of reparation and forgiveness and in the fact of accepting as fair the application of a punishment for the breach of their role. For example, we can take the case of an athlete who once expelled or marginalised from a selection for not having responded to the demands that his role implies—arriving at training, not using drugs, not being involved in scandals, among other requirements—, consciously accepts this decision. Other less severe situations could be treated by timely recognition of the offence and its corresponding apology.
Third stage: the morality of the principles

This stage presupposes the previous two. In other words, it is assumed that the person has experienced attitudes of love and trust within his family, as well as feelings of friendship and mutual trust within the associations in which they have participated. Rawls argues that we develop a desire to support just institutions once (a) we recognise that our loved ones and we benefit from these institutions and (b) these institutions are publicly recognised as fair. This moral motivation is revealed by the desire to fulfil our duties and obligations, as well as by the willingness to do our part within the institutions. Also, this motivation would lead us to do everything possible to modify our institutions when they are not fair. Once the sense of justice has been acquired in the third stage, a fundamental feature that characterises its motivational force is that it no longer depends on any particular bond, be it love, friendship or sympathy. In the previous stages of development, respect and support for norms were determined by the specific relation between the subject and his family, or between the subject and a particular group.

In contrast, in the third stage, the motivation that leads to respecting moral principles is entirely independent. In this way, the sense of justice appears as a moral feeling that depends on principles [principle-dependent] and not on objects [object-dependent]. According to Rawls, this stage is governed by the third psychological law, which determines that once attitudes of love and trust are generated, as well as feelings of friendship and loyalty together with the recognition that we are beneficiaries of a fair, stable, and lasting institution the corresponding sense of justice is produced.

With regard to the fulfilment of our commitments, this stage reflects the transition that goes from a relation with real people —our parents and friends, for example— towards a relationship with certain principles that we could autonomously give ourselves\(^2\). In other words, this stage would reflect the agents' connection of the will with certain moral principles of justice that would allow the

constriction of their will. Let us remember that for Rawls the virtue of justice is a form of righteousness, understood as the general framework of his moral theory (see RAWLS, 1971, p. 396).

To conclude this section, it should be noted that in this third stage, the desire to be fair implies acting according to principles of justice commonly accepted. The violation of some of the principle of justice would give rise to feelings of guilt in a strict sense to the extent that the principles are a representation and actualisation of our nature as rational, free and equal beings.

**Evaluation of the Rawlsian analysis of our moral psychology**

**On the morality of authority**

In the constitutive process of the sense of justice, the morality of authority must be understood as that stage capable of offering the conditions for the development of love and trust. These natural attitudes are essential for the subsequent configuration of our moral feelings. Rawls did not support the idea that at this stage an individual could develop moral feelings in a strict sense. Rawls argues that when subjects claim to have certain moral feelings, they appeal in their explanation to certain moral concepts. For example, when faced with the question of why someone feels resentment or indignation, the type of explanation that person offers would be completely different from being asked why he feels rage or anger.³ In the case of the first feelings, the person would need to appeal to notions of moral correction. In this regard, Rawls points out:

³ Regarding the moral feelings Rawls said: “Resentment is our reaction to the injuries and harms which the wrongs of others inflict upon us, and indignation is our reaction to the injuries which the wrongs of others inflict on others. Both resentment and indignation require, then, an explanation which invokes a moral concept, say the concept of justice, and its associated principles(s) and so makes a reference to a right or a wrong. In order to experience resentment and indignation one must accept the principles which specify these rights and wrongs” (1963, p. 299).
In general, it is a necessary condition and a defining feature of moral feelings that the person’s explanation invokes a moral concept and its associated principle(s) and thereby makes a reference to an acknowledged right or a wrong. For example, a person feels guilty because he knows that he has taken more than his share and treated others unfairly (1963, p. 295).

However, as Rawls said, children who are in the context of relationships specified by the morality of the authority are not in a position to question the general precepts laid down by their mentor. For Rawls, the child does not have a developed critical sense because “He is not in a position rationally to reject parental injunctions, so that, if he loves and trusts them, he will accept their precepts” (1963, p. 288). Nevertheless, and as much as Rawls strives to point out that only in the third stage of development the feeling of guilt appears strictly, it is not clear why he maintains that this feeling is what would appear as a result of the violation of the mandates and precepts delivered by their parents. This idea would be ratified by the spontaneity of this feeling that would place it, according to the same classification that Rawls offers, in the category of natural feelings rather than a moral one. So, if it is only possible for Rawls to have feelings of guilt to the extent that we have a concept of guilt and we can offer an appropriate explanation of why one feels that way, what reason would he have to talk about feelings of guilt in relation to authority?

On the morality of the association

The conclusions Rawls obtains from the analysis of the morality of authority are related to the need to find a justification of the principles of justice. We cannot forget that these principles, which were presented by Rawls at the beginning of his 1963 article, are the aim of the discussion. As I argued, the objective pursued by Rawls is to justify of the principles of justice based both on our natural dispositions and moral feelings and on the normal development of our moral psychology. Thus, and if the conditions specified by Rawls are met - that is, if the subjects develop the feeling of guilt as a result of violating their commitments as constituent parts of cooperative activities - then that feeling of guilt would function as an
element capable of restricting our possible selfishness, tending to seek our own benefit and interest.

Rawls compares this result no less than with the capacity Hobbes assigned to the sovereign in Leviathan (see RAWLS, 1963, p. 290). Rawls believes that one way to interpret the Hobbesian sovereign is as an agency linked to unstable systems of co-operation. By maintaining vigilance and enforcing sanctions with the force of the sword, the sovereign would be able to inhibit those actions that produce social instability. Similarly, feelings of friendship and mutual trust would play a central role in the stability of society to the extent that “they tend to reinforce the scheme of co-operation” (1963, p. 291).

4

On the morality of the principles

The feelings of guilt that arise in the context of the associations have -like its predecessor the morality of authority- the same condition of being subject to natural relations and attitudes towards particular persons. In this sense, both forms of guilt respond to real relationships we have with our close ones. Unlike the above, guilt that appears in the third stage responds to a different structure. The reason is that it is possible to feel guilty even when those affected are not people with whom we have some form of sympathy. Following Rawls, in explaining the reasons why we feel principle guilt, we do not appeal to our relations with our neighbours, but rather to principles of justice that would adjust our interests to the conditions of equity, thus restricting our particular interests through the constriction of our will. Particularly, acting under the awareness of guilt in relation to the breach of the principles of justice would have the same effect as having a morality. For Rawls, having a morality implies reciprocal and impartial recognition of the application of the principles of justice. Thus, the recognition of equity—as a central feature of

4 One example of the feeling of guilt by association is when people recognize a fault and request to be reintegrated, accepting the reproaches and punishments associated with it. In this sense, it is possible to think that the motivational element that leads subjects to constrain their will is not the guilty as a moral feeling, but rather the fact of feeling excluded.
justice—implies the need to give explanations when someone acts against them, as well as the rise of feelings of guilt, shame and remorse.

But if the restrictions of having a morality are necessary conditions for recognising and justifying the principles of justice, why does the sense of justice require accepting the principles of justice as its condition of possibility? In other words, why does Rawls rest the sense of justice in his third psychological law when at the same time the principles of justice require the condition imposed by the sense of justice? In this regard, it could be argued that either the sense of justice depends on the acceptance of the principles, or the justification of the principles of justice depends on having a morality, that is, on having developed our moral competencies. I have tried not to consider Rawls's arguments as following deductive patterns, but it is difficult not to see a confusing argument here, a circular one. The circularity that we can observe is given by the fact that the sense of justice is, in the first place, part of the conditions for the justification of the two principles of justice, but then it appears as an effect of the third psychological law that is presupposed as a framework for the acceptance of the principles of justice.

**Conclusions: moral psychology, sense of justice and social stability**

In this paper, I have shown that the conception of moral development presented by Rawls has as one of its objectives to give reasons for accepting that its principles are the most appropriate for the rise and maintenance of the sense of justice and social stability. In this line of argument Thomas Hill in his paper “Stability, A Sense of Justice, and Self-Respect”, said that for Rawls’ empirical studies of psychology could provide some elements that would allow make comparisons between different theories of justice, for example, regarding the development of those moral feelings associated with social stability. Hill stated that, unlike utilitarianism, “a stabilizing sense of justice would tend to develop in those who grow up in a society well-ordered by Rawls’s the principles of justice” (2014, p. 205). However, to defend a particular version of moral development as a condition
for the appearance of the sense of justice implies accepting that other commitments that should be attached with the moral-normative basis of his theory. Following this idea, we could argue that appealing to a theory of moral feelings as a condition for accepting the principles of justice would make it difficult to admit, as Rawls did, that they have the features of categorical imperatives (see Rawls 1971, p. 253).

Having said this, the way to get out of this apparent contradiction is to assume that the exposition of moral psychology that Rawls presented in Chapter VIII of the Third Part of TJ called ‘The sense of justice’, does not have the same argumentative characteristics as the kind of justification he offered of his principles in Chapter III of Part One called ‘The original position’. We know that the main objective of the third part of TJ was to argue for the stability of a well-ordered society based on the basic structure that satisfies the two principles of justice (see Rawls 1971, p. 347). Rawls considered his argument in favour of stability as an essential confirmation of the argument in support of the principles of justice. As is known, in the third part of *A Theory of justice* Rawls offers reasons why the parties in the original position and behind the veil of ignorance would choose the two principles of justice. However, Rawls considers necessary to complement this last argument because, under the mechanism of hypothetical representation, the parties were evaluating the merits of the principles of justice of an actual well-ordered society. Initially, the parties in the original position neither take into consideration societies that could be profoundly diverse and divided by justice issues, nor do they have insight on how such societies were created or how they could survive if a series of abuses and discords begin to threaten the order and structure. Rawls believes that the lack of an internal stabilising force to face such challenges can lead to the destabilisation of a society, making it doubtful that people have enough motivation to accept it.

As I have shown in this paper, Rawls defended the idea that a society structured by its principles is more stable than other societies. Thus, both from the points of view of the parties (the contract argument) and citizens (the argument of stability and coherence) the principles of justice could be affirmed as offering better guarantees than others. This would occur not only concerning their moral justification but also to the consistency between the principles of justice as fairness.
and laws of moral psychology, laws that are in favour of the development of the sense of justice and social stability.

References


RAWLS, J. Stephen Toulmin’s An Examination of the Place of Reason in Ethics. Philosophical Review, n. 60, p. 572-580, 1951b.


RECEBIDO: 23/01/2020
APROVADO: 09/04/2020