

## Is It Possible to Understand John Rawls's *Theory of Justice* Independently of Kantian Moral Foundations?

### John Rawls'un *Adalet Teorisini* Kantçı Ahlakî Açıdan Temellerden Bağımsız Olarak Anlamak Mümkün Müdür?

Tuba İlhan Dalar

Öğr. Gör. Dr., Kilis 7 Aralık Üniversitesi/İnsan ve Toplum Bilimleri Fakültesi/Felsefe Bölümü,  
tuba.ilhandalar@kilis.edu.tr,  
ORCID: 0009-0003-7284-5395

#### Article Information

##### Article Type

Research Article

##### Date Received

21.11.2025

##### Date Accepted

19.12.2025

##### Date Published

31 December 2025

**Plagiarism Checks:** Yes, Turnitin.

##### Ethical Statement

It is declared that scientific and ethical principles have been followed while carrying out and writing this study and that all the sources used have been properly cited. (Tuba İlhan Dalar)

**Conflicts of Interest:** The author(s) has no conflict of interest to declare.

Licensed under CC BY-NC 4.0 license.

#### Makale Bilgisi

##### Makale Türü

Araştırma Makalesi

##### Geliş Tarihi

21.11.2025

##### Kabul Tarihi

19.12.2025

##### Yayın Tarihi

31 Aralık 2025

**Benzerlik Taraması:** Evet, Turnitin.

##### Etik Beyan

Bu çalışmanın hazırlanma sürecinde bilimsel ve etik ilkelere uyulduğu ve yararlanılan tüm çalışmaların kaynakçada belirtildiği beyan olunur. (Tuba İlhan Dalar)

**Çıkar Çatışması:** Çıkar çatışması beyan edilmemiştir.

CC BY-NC 4.0 lisansı ile lisanslanmıştır.



Cite As | Atıf

Dalar, Tuba İlhan (2025). Is It Possible to Understand John Rawls's *Theory of Justice* Independently of Kantian Moral Foundations. *Mebadi International Journal of Philosophy*, (2) 2, 29-53. <https://doi.org/10.5281/zenodo.18080536>

**Abstract**

The purpose of this article is to analyse the position of John Rawls's theory of justice as fairness vis-à-vis Kantian moral foundations. Rawls's major work, *A Theory of Justice* (Rawls, 1971), is of interest not only in the field of political philosophy, but also in the debates over whether his position can be considered a Kantian. In this context, the paper begins by asking a crucial question: Is Rawls a Kantian or not? To answer this question, the paper delves deeply into a discussion of Rawls's Kantian position, drawing extensively on secondary sources. These secondary sources are re-examined and classified for further evaluation, as they shed light on improving our understanding of Rawls's views on Kant's moral foundation. In particular, the relationship between Rawls's theory of justice as fairness and the impact of Kant's principal work, the *Groundwork* (*Groundwork of the Metaphysics of Morals*), is discussed. The Kantian legacy becomes even more apparent in Rawls's construction and testing of his principles of justice. These principles are considered as part of a thought experiment (the original position) behind a veil of ignorance. Especially in section §40 of *TJ*, Rawls's Kantian interpretation regarding the genesis of the principles of justice as fairness determines the trajectory of this inquiry. While Rawls is clearly successful in linking the features of the person and justice as fairness to the Kantian categorical imperative and in establishing a parallel between the purely rational/autonomous person (in the Kantian sense) and the participant in the original position, a non-Kantian reading of *TJ* reveals some significant disagreements about the precise nature of Rawls's Kantianism. Nevertheless, the relationship between Rawls's theory of justice as fairness and its Kantian foundation remains a highly controversial issue among scholars. Therefore, this article aims to outline these disagreements between interpretations concerning Rawls's intellectual debt to Kant.

**Keywords:** Rawls, Kantianism, Justice as Fairness, Moral foundations, *Groundwork*.

**Öz**

Bu makalenin amacı, John Rawls'un hakkaniyet olarak adalet teorisinin Kantçı ahlaki temeller bağlamındaki konumunu analiz etmektir. Rawls'un başlıca eseri *Bir Adalet Teorisi*, (Rawls, 1971), yalnızca siyaset felsefesi alanında değil, aynı zamanda onun konumunun Kantçı olarak kabul edilip edilemeyeceği konusundaki tartışmalarda da ligi çekicidir. Bu bağlamda, makale kritik bir soru sorarak başlar: Rawls Kantçı mıdır, değil midir? Bu soruyu yanıtlamak için makale, kapsamlı bir şekilde ikincil kaynaklardan yararlanarak Rawls'un Kantçı konumunu derinlemesine incelemektedir. Bu ikincil kaynaklar, Rawls'un Kant'ın ahlaki temellerine dair görüşlerini daha iyi anlamamıza ışık tuttukları için yeniden incelenmekte ve daha iyi bir eleştirel analiz için sınıflandırılmaktadır. Özellikle, Rawls'un hakkaniyet olarak adalet teorisi ile Kant'ın temel eseri olan Temellendirme'nin (*Ahlak Metafiziğinin Temellendirilmesi*), etkisi arasındaki ilişkisi tartışılmaktadır. Kantçı miras, Rawls'un adalet ilkelerini inşa etme ve test etme sürecinde daha da belirgin hale gelir. Bu ilkeler, cehalet perdesi ardındaki bir düşünce deneyinin (orijinal durum, ilk konum) parçası olarak ele alınır. Özellikle AT'sinin §40 bölümünde, Rawls'un adalet ilkelerinin hakkaniyet olarak ortaya çıkmasına ilişkin Kantçı yorumu, bu araştırmanın gidişatını belirlemektedir. Rawls, bireyin ve hakkaniyet olarak adaletin özelliklerini Kantçı kategorik buyruğa bağlamada ve salt akıl/özerk kişi (Kantçı anlamda) ile orijinal durumdaki katılımcı arasında bir paralellik kurmada açıkça başarılı olsa da AT'nin Kantçı olmayan bir okuması, Rawls'un Kantçılığının kesin doğası hakkında bazı önemli anlaşmazlıkları da ortaya koymaktadır. Bununla birlikte, Rawls'un adalet teorisi ile Kantçı temeli arasındaki ilişki, akademisyenler arasında oldukça tartışmalı bir konu olmaya devam etmektedir. Bu nedenle, bu makale Rawls'un Kant'a olan entelektüel borcuna ilişkin yorumlar arasındaki bu anlaşmazlıkları ana hatlarıyla ortaya koymayı amaçlamaktadır.

**Anahtar Kelimeler:** Rawls, Kantçılık, Hakkaniyet olarak Adalet, Ahlaki temeller, *Temellendirme*.

## Introduction<sup>1</sup>

This paper purposes to critically review the secondary literature on the earlier Rawls of *Theory of Justice* (hereafter *TJ*, published in 1971)<sup>2</sup> and its relation to Kantian moral foundations. There is a significant reason for reviewing such secondary sources. Rawls's occasional attempts to show his own conception of justice as fairness as a crucial Kantian doctrine have not only attracted attention but also drawn criticism from other scholars. The general assumption among many readers of Rawls is that his theory of justice is excessively universalistic and abstract, and they see his theory as an extension of a Kantian comprehensive liberal doctrine where the features of the original position are similar to the Kantian categorical imperative principle; however, some argue the opposite position and regard his position as non-Kantian. Now, the following question arises: Does the Rawlsian intellection of justice depend on a Kantian moral interpretation or not? This question will bring us back to questioning the Kantian interpretation of justice as fairness expressed by Rawls in *TJ* and take us on to discuss the influence of thinkers with different Kantian interpretations on Rawls's Kantian position. So, my aim here is to contribute substantially to the literature by contending that these viewpoints are found upon incomplete and oversimplified interpretations of Kant's thought, a deficiency that cannot be attributed to Rawls. Since, in my view, one aspect worth examining is Rawls's own interpretation of Kant, rather than scrutinising him based solely on a one-sided reading of Kant. I will establish this more nuanced reading through the essay.

To address the question posed above, I will first consider the major interpretations and analyses of Rawls's philosophical reasoning, and his philosophical foundation referred to as "Kantianism." The fact is that Rawls's self-identification as a Kantian is not sufficient to provide clarity. Given this predicament, I argue that Rawls's Kantianism should be analysed in the context of today's

---

<sup>1</sup> This article is especially derived from the 'literature review' part of my doctorate dissertation entitled "Reappraising Rawls's Kantianism Through Hegel's Social and Political Thought," supervised by Prof. Dr. David Edward Rose and Dr. Michael Lewis, Ph.D. Dissertation, Newcastle University, UK, 2024.

<sup>2</sup> It must be noted that this paper focuses solely on Rawls's first major work, *TJ*, and examines the Kantian moral foundations of justice within this limited framework. Because Rawls's political turn, in his 1980 *Dewey Lectures* and subsequent works, requires the scope to be considered a non-Kantian Rawls and to be the subject of another research. In particular, Rawls's 1985 article "Justice as Fairness: Political not Metaphysical" clearly articulates how he distinguishes his political conception of justice from all metaphysical and moral doctrines (including Kantian doctrine). Evidently, following self-criticism, Rawls's theory in another major work, *Political Liberalism* 1993, shifted from a Kantian moral doctrine to a political theory. However, I have elaborated on this claim in more detail in my unpublished doctoral dissertation, which discusses it within a more contentious scholarly context. As mentioned earlier, this issue is beyond the scope of the current article.

leading philosophers. My primary goal is to gain a more nuanced understanding of Rawls's Kantianism. In doing so, I hope to provide a critical analysis of some of the most notable notions that have been debated in relation to Rawls's Kantianism. Non-Kantian interpretation of *TJ* and Kantian readings of *TJ* provide the motivation for the debates. I desire to reveal the Kantian basis in Rawls's theory and demonstrate it through various scholarly discussions. In the past five decades, many thinkers have discussed Rawls's position (without) relying on a Kantian background, and I shall group them around common positions. The first group says that Rawls is not a Kantian. On this point, a number of thinkers have criticised Kantian Rawls and find it odd that Rawls is recognised as a Kantian. I will highlight several scholars who have discussed or referred to the relationship between Kant and Rawls in their works and have concluded that Rawls is not a Kantian: Oliver A. Johnson, 1974 and 1977; Andrew Levine, 1974; Joseph M. Grcic, 1983; Otfried Hoffe, 1984; H.E. Mason, 2003, and Kerst Budde, 2007. For example, through questioning Rawls's *TJ*, these thinkers typically conclude that Rawls's theory cannot be labelled a type of Kantianism. To these scholars, the "comprehensive" foundation of Kantian moral conceptions does not preserve the integrity of Rawls's philosophy. They question Kantian Rawls and suggest evaluating him using an alternative philosophical framework. They argue that Rawls misinterprets Kantian morality, making it difficult to call him a Kantian.

The opposing view has been developed by several other scholars. More selectively, they all agree on Rawls's Kantianism: Stephen L. Darwall, 1976 and 1980; Robert P. Wolff, 1977; Arnold I. Davidson, 1985; Catherine Audard, 2007; Paul Guyer, 2018; Nicholas Tampio, 2007; Vadim Chaly, 2015; Jean Hampton, 1980; Michael Sandel, 1982; Modupe O. Adu, 2024 and Hong Yang, 2025. These scholars broadly acknowledge that Rawls improves Kant's status in contemporary moral and political philosophy (Wood, 1999, p. 337). I will also discuss Rawls's proponents who say his theory is based on Kantianism and that his philosophical methodology is Kantian, that Rawlsian justice depends on Kantian ideas of free and equal moral beings.

Later, in contrast to the Kantian interpretation of Rawls, I will show how Kant's explanation of self-legislating or moral agency appears to have shaped his theory. I argue that Rawls and Kant have both reached the same or similar presuppositions about the categorical imperative and the original position. In spite of the arguments of theorists who are critics of Rawls's Kantianism, I primarily maintain that Rawls's theory legitimately builds on and advances the massive legacy of Kantian moral assumptions. Let us now turn to a different type of challenge to Rawls's Kantianism and evaluate the most important criticism of Rawls's early work.

## 1. A Non-Kantian Interpretation of the *Theory of Justice*

In his early book *TJ*, Rawls claims that his view of justice as fairness is “fundamentally Kantian in nature” due to his interpretation of Kant’s ethical writings, particularly the *Groundwork of Metaphysics of Morals* (*GMM*, 1785)<sup>3</sup> (*TJ*, pp. vii, 11, 251).<sup>4</sup> This reliance on Kantian ethics is what the secondary literature has come to see as the problem with Rawls’s early work. In fact, Rawls (*TJ*, p. 221) remarks on “the content of the principle of equal liberty and the meaning of the priority of rights that defines.” It seems consistent at this point to state that there is a Kantian interpretation of justice as fairness from which this principle derives. This interpretation encompasses Kant’s conception of autonomy. Although Rawls himself states this viewpoint in *TJ*, whether his theory can be called Kantian or not has remained a matter of debate. Moreover, Hampton’s (1980) and Johnson’s (1974) arguments, taken together, lead the reader to question whether Rawls’s theory falls within the social contract and Kantian traditions (Corlett, 1991, p. 4).

In his essay, “The Kantian Interpretation,” Oliver A. Johnson (1974) examines the Rawlsian interpretation of Kant’s autonomy account and seriously discusses its Kantianism. Johnson first examines Rawls’s Kantian position and rejects Rawls’s Kantian interpretation of justice (1974, pp. 58-62). Johnson points out that individuals in the “original position behind” the “veil of ignorance” are motivated by what Kant refers to as heteronomous inclinations: “An action originally heteronomous is not rendered autonomous, even though performed under a veil of ignorance if the nature of motivation is unchanged” (Johnson, 1974, p. 62). Therefore, according to Johnson, Rawlsian principles of justice conflict with Kantian autonomy, the categorical imperative, and pure practical reason.

Interestingly, Rawls has realised the confusion that gives rise to Johnson’s claim, and he clearly expresses it earlier, as he states:

... since the persons in the original position are assumed to take no interest in one another’s interests...it may be thought that justice as fairness is itself an egoistic theory. It is not, of course, one of the three forms of egoism mentioned earlier, but some may think, as Schopenhauer thought of Kant’s doctrine, that it is egoistic nevertheless, now this is a misconception. For the

---

<sup>3</sup> Hereafter referred to as the *Groundwork* in the text. All citations will be from the following edition: Kant, I. *Groundwork of Metaphysics of Morals*, Gregor, M. (ed. and trans.) Cambridge University, 2012. Also, the *Groundwork* will be at the centre of this paper, since it plays a significant role in the development of Rawls’s early intellectual and philosophical thought. So here we will make further references to the *Groundwork*.

<sup>4</sup> Rawls frequently refers to Kant’s ideas in *TJ*; for instance, Rawls declares: “My aim is to present a conception of justice that generalises and carries to a higher level of abstraction the familiar social contract theory found, for example, in Locke, Rousseau, and Kant” (*TJ*, p. 11).

fact that in the original position the parties are characterised as not interested in one another's concerns does not entail that persons in ordinary life who hold the principles that would be agreed to are similarly disinterested in one another. Clearly, the two principles of justice and the principles of obligation and natural duty require us to consider the rights and claims of others. And the sense of justice is normally effective desire to comply with these restrictions. The motivation of the persons in the original position must not be confused with the motivation of persons in everyday life who accept the principles that would be chosen and who have the corresponding sense of justice...(TJ, pp. 147-8)

Rawls's theory is not egoistic or self-serving, as illustrated in this quote. More specifically, he focuses on such principles and links them to part §40 of TJ, "The Kantian Interpretation of Justice as Fairness", where he himself remarks that his method of describing justice is Kantian in origin.

As a brief recap, in the context of the Kantian deontological doctrine, Rawls, claims that justice cannot be derived; it is a requirement of duty for duty's sake for moral persons because it is the right thing to do, even if it has nothing to do with the general good, utility, interests, or an ideal of human perfection (Audard, 2007, p. 43). So, rights are prior to welfare or pleasure, and they are unconditional and precede the preferences of the majority. This priority right over the good becomes a central feature of understanding justice as fairness. As Rawls admits, this is clearly inspired by the Kantian view of justice (TJ, p. 16n). Rawls, in his discussion of utilitarianism, is concerned with the question of the priority of "good" or "right." The question is: in a moral or political theory which basic moral concept – the good or the right – should have priority?" The question is significant because the demands imposed by right, in the manner of a duty, might simply conflict with the results of increasing good (Cekić, 2022, p. 43). This is the context for the issues raised by the commonly held notion of *fairness*. So, Rawls takes the view that the priority of right is a fundamental feature of Kant's ethics. Hong Yang, who sees Rawls as a proponent of Kantianism, expresses this situation as follows: "goodness is prior to rightness... because the sense of justice can be considered goodness by the citizens of a well-ordered society" (Yang, 2025, p.44). In the original position, the parties, who are hypothetical personalities, must choose the principles of justice in the criterion of rightness.<sup>5</sup>

In TJ, Rawls presents ideal justice by drawing a distinction between "ideal" and "nonideal" theory, or what he calls "strict compliance" and "partial

---

<sup>5</sup> However, in his article, Yang later argues that goodness is necessarily compatible with rightness and resolves the problem of priority by taking Rawls's later work, *Political Liberalism*. For more information, see his article titled "From Moral Philosophy to Reflective Judgment: Rawls as a Successor to Kant."



compliance" (*TJ*, §25, §39).<sup>6</sup> Rawls delineates his theory of justice as fairness as a universal moral ideal that is desired by all societies (Freeman, 2003, p. 2). In this vein, in *TJ*, Rawls provides his conception of the world that is based on an ideal perfect society, and his ideal of the person in the original position is also elaborated within the framework of his own Kantian interpretation of justice as fairness. Rawls explicitly makes clear that all characteristics of the original position must be evaluated regarding the "moral powers" ascribed by these ideal persons.<sup>7</sup> Scanlon (1973, p. 1022) highlights that imagining behind the "abstractions," Rawls constructs the original position as a particular ideal of the person, connected to an ideal of a well-ordered society. Firstly, it is important to remember that the original position is not a justification for the justice, since it is purely hypothetical, a sort of a "thought experiment" or a "device of representation."<sup>8</sup> The hypothetical contract, the original position, for Rawls, is not an actual thing, but a device for thinking in the correct way (Dworkin, 1989, pp. 17-8). Thus, Rawls's conception of person and society appear to be quite abstract.

Furthermore, Rawls has assumed that the persons in the original position are *rational* and do not have their own conception of the good (*TJ*, p. 123). This means that his account of the parties in the original position is theoretically defined as rational individuals in a thin sense insofar as they choose principles merely to promote their particular ends and interests. It is also important to note that Rawls's parties can be described as purely rational and liberal individuals, but it would be a mistake to say they are egoists. As Freeman states, "they are not egoists any more than chess players who play to win or buyers who shop for the lowest price are egoists" (Freeman, 2003, p. 13). Indeed, their moral interests are among the interests they propose to protect in their choice of the principles of justice (*TJ*, p. 125). The chief point here is that the parties are supposed to be clearly non-egoistical since they have a capacity for an effective sense of justice – "a desire to act not just according to but also for the sake of justice" (Freeman, 2003, p. 14). Therefore, Rawls stresses that the motivation of the person in the original position should not be confused with the motivation of the person in ordinary life (*TJ*, p. 126). For the fact that the parties are characterised as mutually disinterested does not entail that a person is in a just society. Then he adds, "the parties can rely on each other to understand and to act in accordance with whatever principles are finally agreed to" (*TJ*, p. 125). At that point, recalling Kant's motivation of duty – "duty for duty's sake,"

---

<sup>6</sup> The character '§' points to the section numbers indicates the book chapters.

<sup>7</sup> See T. M. Scanlon's 1973 essay "Rawls's Theory of Justice," which shrewdly analyses that point.

<sup>8</sup> As we will see in the following, many critics of Rawls have accused the original position of being abstract, missing its nuances in so doing.

in a similar sense, Kant's *Groundwork* explicitly identifies the conception of the person as a rational being who makes moral law. He examines this especially closely in sections I and II of *Groundwork*, along with the ideas that a law must be universal, and persons are ends in themselves.

According to a shallow reading of Kant, being autonomous or free in the moral sphere is to be able to follow "reason," not desires or external things. For instance, in the *Groundwork*, Kant holds that we, as rational beings, act on our presentation of law and can make and legislate the law. This line of thought leads to what Kant describes as the principle of the will of every rational being as a will giving universal law (*GMM*, 4:432). In other words, in his formulation of humanity, Kant lays out a conception of autonomy in which rational beings are capable of autonomous moral motivation. In Kant's words, autonomous actions are motivated by moral reasons. In addition, Kant says that moral motivation must be autonomous, not heteronomous, as he believes that we rationally bind ourselves to the law. The principle that we give universal law through our maxims suggests that moral motivation is autonomous. If we are motivated to obey a law heteronomously by a sanction, then the imperative we follow in obeying that law is a hypothetical imperative. However, according to Kant's principle of autonomy, we are able to make moral law and legislate it, and this principle, as he says, "would be very well suited to be the categorical imperative" (*GMM*, 4:432). Finally, Kant assumes that this moral legislation must be accepted under conditions that characterise men as free and equal rational beings. At that point, Rawls introduces a similar standpoint, in particular his view of the rationality and the motivation of the parties, as just seen, aligns with Kant's motivation of duty, which is performing the right actions motivated *from duty*, not from *immediate* inclination (*GMM*, 4:397). This view provides Kant's formulation of the categorical imperative that commands us to act only for the sake of duty as the only way that an action has moral value. In order to understand this claim, it is necessary to understand the image behind them: Kant wants us to think of someone who does not sympathise with the suffering of others and is not inclined to help them; as Kant puts it:

Suppose that now, when no longer incited to it by any inclination, he nevertheless tears himself out of this deadly insensibility and does the action without any inclination, simply from duty; then the action first has its genuine moral worth. (*GMM*, 4:399)

For Kant, reflection on this fact leads us to realise that the moral value of an action is not derived from its aim, but rather from the "maxim" on which it is done, the principle on which the agent acts (*GMM*, 4:399). In short, Kant holds that moral action is the action done from duty that must be for duty's sake; otherwise, it is not duty. In Kant's view, the moral value of human action



depends upon the motivation from which it is done. This moral motivation is not dependent on any external or outer results. Even if the consequences or effects turn out to be bad, the action and the agent's act should not be judged negatively if the intention is morally sound. Kant's motive of duty contrasts with the motive of inclination and the motive of self-interest. This standpoint provides us with the Kantian view. This is precisely how I show that Rawls's Kantian status, at a first glance, is justified on the grounds of this common and stereotypical reading of Kant.

This moral motivation is also ultimately fundamental to Rawls's argument for the principles of justice and their stability. Like Kant's argument for acting for the sake of duty, in accordance with the principle of duty, Rawls assumes that the parties do not act from their personal desires, inclination, or aims, deliberating on the principles of justice for the basic structure of society. They "take no interest in one another's interest" as contracting agents but are concerned only with promoting their own interests (Freeman, 2003, p. 14). Rawls believes they make a rational decision from the standpoint of the original position. In this way, Rawls suggests the Kantian conception of a person is based on an ideal of the person. Rawls's conception of the free and rational participant in the original position can be described as Kantian. At that point, Rawls borrows from Kant's work *Groundwork*, where Kant shows how rational individuals reach moral decisions. Like Kant, Rawls also starts from "the idea that moral principles are the object of rational choice" (*TJ*, p. 221) and that justice is the result of a rational agreement.

Most evidently, in §40 of *TJ*, Rawls's attempt to connect his theory to Kant's moral philosophy is a Kantian interpretation of justice as fairness in which he asserts his theory in the original position behind the veil of ignorance. The fundamental idea is that the deliberations of the persons in the original position are analogous to those of the deliberations of an individual with a good will who tests his maxims in light of Kant's categorical imperative (Pogge, 2007, p. 189; Wolff, 1977, pp. 101-6). Rawls finds his point of contact with Kant in the improved conception of the original position as a condition of rational choice behind the veil of ignorance. As he says the Kantian interpretation is ultimately intended to demonstrate the following:

the description of the original position resembles the point of view of noumenal selves, of what it means to be a free and equal rational being.....the original position may be viewed, then, as a procedural interpretation of Kant's conception of autonomy and the categorical imperative within the framework of an empirical theory. (*TJ*, pp. 225-6)

The idea is that when we choose by isolating from or ignoring our own particular abilities, characteristics, and personal background, we choose as if

we were noumenal rather than phenomenal agents. In the Kantian sense, since each of us freely selects the principles by which we live, each of us has the capacity to make a law for ourselves and is therefore autonomous: “subject only to laws which are made by (oneself) and yet are universal” (*GMM*, 4:432). It is possible for us to adopt the view of the original position, and our decision to do so “expresses our nature as free and equal rational persons” (*TJ*, p. 256). The link between Kant and Rawls appears to stem from Rawls’s commitment to incorporating the Kantian notion of autonomy into his own theory of justice. Namely, the Kantian conception of the autonomous person Rawls invokes in *TJ* is a philosophical view of moral agency. It assumes that we are free and rational agents because we have moral capacity for practical reasoning. It is assumed that Kantian autonomy is exercised under conditions of freedom that allow agents important opportunities to figure out the right thing to do, and this is the core value of freedom. The claim that moral legislators proceed from laws given by the rational being means that, in terms of moral motivation, every individual is acting on their own faculty as a rational and autonomous being. At that point, Rawls adopts Kant’s method on the doctrine of autonomy: Each individual is free and equal if, and only if, they are autonomous persons. The assumption under the principle of justice is that we should treat persons as moral beings acting in relation to a categorical imperative for the human beings (*TJ*, pp. 222-3). In other words, Rawls accepts the validity of the Kantian definition of autonomy and applies it to support his own initial position. The claim appears to be that all humans are moral beings. Thus, all individuals are entitled to equal justice and freedom, which is considered to be an aspect of the Rawlsian political conception of equality. These persons in the initial position know that they also have the capacity to demonstrate a sense of justice. Rawls advances views of autonomy as self-realisation on this basis (*TJ*, p. 221). This view leads Rawls to associate his explanation of the original position with the kingdom of ends and to say that the party in the original position is like the Kantian ideal/noumenal person (Wolff, 1977, p. 114).

According to Kant, there are two ways for human beings to be motivated. The first is when humans are determined in accordance with laws from a phenomenon or an external world, and then their actions are heteronomously motivated. In contrast to this, when humans establish moral laws from their own self-reflection, they are autonomous and free persons. Rational beings, that is, act on a law that they have provided for themselves (Korsgaard, 1996, p. 22). Regarding this, in *TJ*, Rawls explicitly states: “a person is acting autonomously when the principles of his action are chosen by him as the most adequate possible expression of his nature as a free and equal rational being” (*TJ*, p. 222). Rawls adds: “the principles of justice are also analogous to categorical imperatives. For by a categorical imperative Kant understands a principle of conduct

that applies to a person..." (TJ, p. 222). Importantly, we have seen how the rationality of characteristics of the parties of Rawls's approach are built on a manoeuvre that bears a striking resemblance to Kant's ideal of a rational moral being.

Returning to Johnson's discussion, he observes the incongruity in equating the Rawlsian conception of the person with the Kantian one of a moral autonomous individual. He supposes that the parties in the initial position and their choices are not based on autonomous choice; they stimulate the interest of each. Additionally, the decisions of the parties in the starting position are enforced by heteronomous principles, not autonomous ones, as their decisions derive from their interests rather than regard for moral rules. That means, in Johnson's view, that there are inconsistencies between the Rawlsian and Kantian views of individuals as autonomous moral beings (Johnson, 1974, p. 58). Johnson then argues that Rawls's theory of rationality is inconsistent in the Kantian sense and claims that Rawls's point of view cannot be given a Kantian interpretation. Nevertheless, the Kantian legislator may be ethically independent, and Rawls's original parties may be rational choosers. Johnson's critique of Rawls's Kantianism fails because of his general a priori interpretation of autonomy. So, I think that Johnson's criticisms are based on his reading of Kant, namely that Kant's categorical imperative *generates* moral imperatives. Rawls does not describe the original position as a means of establishing a state or society. Rawls uses the contract to test our moral motivations and principles of justice. Significantly, the original position expresses the idea of moral equality, that each person's moral personality should be respected (Rawls, 1999a, p. 254). Johnson overlooks this crucial issue, which is the Rawlsian original position, in which Kant's categorical imperative *tests* moral motivation rather than generating it.

In his 1974 essay "Rawls's Kantianism," Andrew Levine follows the line of Johnson's claims. He misappraises Rawls, arguing that Rawls's notion of justice as fairness is connected to Kant in a polemical way. Levine interrogates Rawls's Kantian interpretation, arguing that it is based on "a systematic confusion of an anthropological understanding of Kant's notion of rational agency (replete with contingent assumptions about human nature) and Kant's own non-anthropological understanding" (Levine, 1974, p. 48). In the Rawlsian original position, the basic principles of justice are considered to free our choice of principles from what Kant calls *empirical* or *heteronomous tendencies*. However, Levine believes that the things we think about in the original position are not "pure" and autonomous motivation in the Kantian sense. According to Levine, Rawls attempts to reconcile the Kantian notion of universality with Hobbesian egoistic reason, which leads to inconsistency. Levine then argues that instru-

mental rationality used by Rawls involves an empirical element and that corresponds to heteronomous in Kant's words. From this perspective, Levine claims that Rawls's autonomy and categorical imperative cannot be interpreted as Kantian. Levine believes that Rawls does not discuss his parties' personalities in the original position with Kant's pure practical reason. Levine then claims to show that Rawls seeks Hobbesian egoistic rationality rather than Kantian universality, but his reinterpretation is incoherent because the parties in original position are not selfish. As Rawls states:

It should be noted that I make no restrictive assumptions about the parties' conceptions of the good except that they are rational long-term plans. While these plans determine the aims and interests of a self, the aims and interests are not presumed to be egoistic or selfish. Whether this is the case depends upon the kinds of ends which a person pursues. If wealth, position, and influence, and the accolades of social prestige, are a person's final purposes, then surely his conception of the good is egoistic. His dominant interests are in himself, not merely, as they must always be, interests of a self. There is no inconsistency, then, in supposing that once the veil of ignorance is removed, the parties find that they have ties of sentiment and affection and want to advance the interests of others and to see their ends attained. But the postulate of mutual disinterest in the original position is made to ensure that the principles of justice do not depend upon strong assumptions. Recall that the original position is meant to incorporate widely shared and yet weak conditions. (*TJ*, p. 111)

The passage indicates that the Rawlsian original position models a Kantian moral agent. Levine argues that, in Rawls's original position, "we express our nature as bundles of appetites for primary goods endowed with a capacity for instrumental rationality; not as bearers of pure practical reason" (Levine, 1974, p. 57). This view of human nature is influenced by external factors, heteronomous motivation. Levine adds that we should remember that the central point of Kant's moral philosophy – and the criterion by which it must eventually be assessed – is an attempt for an independent motivation for the moral life, distinct from human nature as a whole. In order for the suggested Kantian interpretation to be effective, the motivation that derives from pure reason would need to be the same as the motivation that stems from the assumptions regarding human nature in the original position (Levine, 1974, p. 52; Cekic, 2022, pp. 48-9).

Levine and Johnson agree that Rawls is unclear about what "rationality" means. They acknowledge that Rawls's rationality and Kant's rational agency are distinct and presumably irreducible. On this view, Johnson and Levine disapprove of a Kantian interpretation of Rawls but ignore Rawls's own later writings. Thus, both Kantian objections to Rawls are narrowly evaluated. Rawls agrees with Kant that autonomy is freedom and rationality. Darwall contends that Rawls's account of reason must be used to define the parties in original

position's rationality. At the end of *TJ*, Rawls discusses his concept of rationality, which is not limited to economics: "within the framework of justice as fairness we can formulate and demonstrate Kantian themes by using a properly comprehensive theory of rational choice" (*TJ*, pp. 583-4). Johnson and Levine generally focus on motivated assumptions about the parties in the original position, which casts doubt on Rawls's Kantian features.

In his book, the *Understanding of Rawls*, Robert Paul Wolff makes the same claims as Johnson and Levine. He contends that it is a very unusual approach to read Kant as claiming that the good (or any goods) is the basis of moral motivation, but Rawls consciously or unconsciously, does. In contrast, Wolff believes that Kant has always maintained a clear stance on this matter, asserting that a material end lacks moral significance. The observation that Rawls's account of "primary goods" is characterised by its generic nature and lack of particular adaptation to individual desires remains rather consistent. The nature of the chosen principle of justice is unaffected by that fact:

[The] veil of ignorance, in fact, only guarantees that the principles will be... generally heteronomous rather than particularly heteronomous. The choice of principles is motivated by self-interest, rather than by the Idea of Good. (Wolff, 1977, p. 115)

Wolff concludes that in Rawls's original position, participants cannot achieve independently willed principles in the Kantian perspective "because their choice of principles must be driven by self-interest to have the bargaining game continue" (Wolff, 1977, p.115). Even in ignorance, they reach only generally heteronomous principles, a hypothetical rather than a categorical imperative – a theory of rational prudence, but never an ethical theory (Wolff, 1977, p. 117).

Joseph M. Grcic's account of Rawls's Kantianism is also influential. Grcic discusses Rawls in his essay, "Kant and Rawls: Contrasting Conceptions of Moral Theory." Grcic believes that Rawls's theory of justice suggests an interpretation of Kant's second categorical imperative. He notes at least three ways in which he views Rawls's theory as not being Kantian, claiming instead that Rawls's theory is Kantian only in "its articulation or expression, not in its foundation of justification" (Grcic, 1983, p. 235). He argues that Rawls's two principles of justice (liberty principle and distributive justice)<sup>9</sup> are an acceptable ver-

---

<sup>9</sup> Rawls formulates the principle of equal basic liberty:

First: "Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all." (*TJ*, p. 53)

Second: "Social and economic inequalities are to be arranged so that they are both: (a) to the greatest benefit of the least advantaged, and

sion of Kant's second formulation of the categorical imperative, but his foundation is actually "a synthesis of Kantian and Utilitarian ideas." Grcic claims that Rawls shows his general ideas on moral justification in the last chapter of his *TJ*.

According to Grcic, Rawls opposes the foundational perspective and believes that a moral theory should be justified like any other theory: "justification is a matter of the mutual support of numerous factors, of everything fitting together into one coherent position" (*TJ*, p. 579). Grcic remains the coherentist approach by quoting *TJ*: "what is required is a formulation of a set of principles which, when conjoined to our beliefs and knowledge of the circumstances, would lead us to make these judgments with their supporting reasons were we to apply these principles conscientiously and intelligently" (*TJ*, p. 46; Grcic, 1983, p. 236). In other words, a moral theory is "true" if it "matches" (*TJ*, p. 579) our "considered judgments" or "judgments in which our moral capacities are most likely to be presented without distortion" (*TJ*, p. 47). For Grcic, Rawls's moral theory differs from Kant's. Nevertheless, Rawls claims that Kant's fundamental contribution was the rational choice theory of morality. "When the principles of his conduct are chosen by him as the most adequate possible expression of his nature as free and equal rational being," a person acts autonomously, according to Rawls (*TJ*, p. 252). He argues that the original position might be regarded "as a procedural interpretation of Kant's theory of autonomy and the categorical imperative" (*TJ*, p. 256). Grcic asserts that the categorical imperative arises from the hypothetical person's "decision" in the original position, not from a priori deduction as in Kant. Rawls exploits Kant to create a nihilistic political philosophy (Bloom, 1975, p. 656). For Bloom, Rawls cannot reconcile Kantian ethics and the utilitarian social contract tradition. So, Rawls seeks to keep Kantian freedom and rationality without accepting the procedure of universality. Rawls wants the "glow of Kantian moral nobility" without the heroic sacrifices of Kant's ethical procedure.

Also, in the same vein, Höffe (1984), in his article "Is Rawls's *Theory of Justice* really Kantian?," notes that although Rawls claims that his conception of rationality is Kantian, he could not properly understand the Kantian sense of reason. Höffe observes, according to Kant, "a human being really finds himself a capacity by which he disguises himself from all other things, even from himself insofar as he is affected by objects, and that is *reason*" (*GMM*, 4:397). It is crucial that reason is appropriately grounded: "[T]he ground of obligation [to moral laws] must be looked for, not in the nature of man nor in the circumstances in

---

(b) attached to offices and positions open to all under conditions of fair equality of opportunity." (*TJ*, p. 53)



which he is placed, but solely a priori in the concepts of pure reason..." (GMM, 4:397). Höffe has interpreted this citation in the following way:

Justice or the moral concept of right cannot be based upon assertions about human nature, i.e. on a practical or empirical anthropology, but must be given a purely rational (*a priori*) foundation in terms of pure practical reason. (Höffe, 1984, p. 104)

Yet, Rawls's conception of primary goods is a proof where justice as fairness is referenced in "practical or empirical anthropology" (Höffe, 1984, p. 105). Nevertheless, Höffe maintains, is that the significant idea here is that this rational and prudential choice is not reliant on Kant. As he puts it, "Prudential precepts represent (pragmatic) hypothetical imperatives, not categorical imperatives; since they are heteronomous and arise from considerations of our own well-being, they represent the very opposite of Kant's moral principle of autonomy" (Höffe, 1984, p. 105).

Nevertheless, Rawls maintains the idea that the principle of justice is to be applied to institutions solely on the basis of general information: "we try to work out what rational legislators suitably constrained by the veil of ignorance, and in this sense impartial, would enact to realise the conceptions of justice. Ideal legislators do not vote their interests" (TJ, p. 251). Strictly speaking, if the principles of justice are not the outcome of rational choice, these principles in the original position must not justify rational selection by free and moral individuals. Rawls modifies the concept of rationality in a broad sense; the crucial feature of justice as fairness is to think of the parties in the original position as rational and mutually disinterested. In addition to this, Audard (2007), Freeman (2007), and Pogge (2007), who are Rawlsian, explicitly mention that this claim is not so. For instance, as Audard clarifies it, "the parties in this initial situation...are artificial persons, clearly distinct from existing citizens...One common mistake made by critics is to treat them as real persons, not as constructs in a device of representation...The parties are representatives who act as trustees or guardians entrusted with citizens' most important interests" (Audard, 2007, p. 84). It is clearly important to note that there is a distinction between the parties in the original position and the actual persons.

In his essay, "On the Kantian Interpretation of Rawls's Theory," H.E. Mason (1976) again explains why Rawls's Kantianism is a problem. As we saw above, Johnson (1974, p. 58) claims that Rawls's theory is the "opposite" of Kant's, and though Mason makes some of the same arguments, he concludes that "rational persons behind the veil of ignorance in Rawls's original position cannot be regarded as Kantian noumenal selves autonomously imposing principles upon themselves" (Mason, 1976, p. 51). However, Rawls himself in many places in TJ acknowledges that:

The description of the original position interprets the point of view of noumenal selves, or what it means to be a free and equal rational being. Our nature as such beings is displayed when we act from the principles we would choose when this nature is reflected in the conditions determining the choice. Thus, men exhibit their freedom, their independence from the contingencies of nature and society, by acting in ways they would acknowledge in the original position. (*TJ*, pp. 255-6)

The Rawlsian deliberation on the original position fits Kant's categorical imperative "maximisation" test. This is because, in the same way that a moral agent using Kant's moral law tests a method of reasoning that tests an agent's maxim by reflecting on what it would be like for him if "all" people acted that way, Rawls's method tests justice by forcing thinkers behind the veil of ignorance to reflect on what a society ruled by this notion would be like for them if they were anyone in that society (Hampton, 1980, p. 337). Overall, Rawls bases his political liberal theory on Kant's moral philosophy. Kantian practical reason will clarify this claim. Thus, both strategies modify one person's interest by considering a rational person's interests. Kant and Rawls also believe this method of testing describes the right procedure of practical reason.

## 2. A Kantian Interpretation of the *Theory of Justice*

More fundamentally, in contrast Johnson's and Levine's main criticisms of Rawls's reliance on Kantianism, I want to note here some of the key thinkers who have strongly acknowledged that Rawls's principles of justice do indeed rely on Kantian ethics.

Rawls mentions that in making their choices, parties are debarred from many items of knowledge about themselves, such that they are not able to psychologically make any rational choice. The reasoning is that such parties can make a rational choice without possessing knowledge of their own primary end, or essential values and attachments. That is consistent with the Kantian idea of autonomy. In essence, Kant outlines this as, namely,

the idea of the will of every rational being as a *will giving universal law*...then although a will that *stands under law* may be bound to this law by means of some interest, a will that is itself the supreme lawgiver cannot possibly, as such, depend upon some interest; for a will that is dependent in this way would itself need yet another law that would limit the interest of its self-love to the condition of a validity for universal law. (*GMM*, 4:432)

Following that valuable quotation, we can understand more clearly what Rawls means by the veil of ignorance when determining the features of the participants in the original position. Since in achieving the principles of justice, they must forget their own particular interest. So, Rawls purposes that individ-

uals consider their identities behind the veil of ignorance so that society's principles of justice can be determined. This is also a political interpretation of Kant's idea of universalizability to arrive at universal and impartial principles (Adu. 2024, pp. 57-60). The position of the legislator or rational being is here to make a law: He or she legislates it, hence their choices are an act of legislation. On that point, I may say that one of the shortcomings of non-Kantian readings of *TJ* is that they do not scrutinise Kant and Rawls adequately. As Modupe O. Adu (2024, p. 60) says that Rawls reformulates of Kantian ethics by "addressing its perceived limitations and offering a more workable framework for applying moral principles to complexities of the contemporary world."

In §40 of *TJ*, Rawls explicitly admits the Kantian interpretation of the original position within its veil of ignorance form. It is true when he starts to expand the argument of the initial position, he uses Kant's philosophical argument. Rawls himself develops his link with Kant's thought in the revised notion of the original position as a condition of rational choice under the veil of ignorance (Wolff, 1977, p. 112). Rawls remains faithful to the Kantian interpretation until the end of the section: "the original position may be viewed, then, as a procedural interpretation of Kant's conception of autonomy and the categorical imperative" (*TJ*, p. 226). The main point here is that in the original position, he posits a moral person who constructs a "device of representation" designed to impart the principles of justice. In Rawls's original position, such people should forget about their particular identities and do not know who they are in the initial position, even being ignorant of their conceptions of the good. But these persons in the initial position know they also have a capacity for a sense of justice. In addition to this, in "Fairness to Goodness," Rawls (1999b, p. 536) points out that the parties' conditions in the original position are constructed on the grounds of two basic elements, which are that (a) the initial agreement must be unanimous, and (b) the parties, with their conceptions of the good, must be treated fairly. Nonetheless, Rawls's social contract theory is not *modus vivendi* since he has a different perspective on the social contract tradition; thus his procedure of justice as fairness has stability. For this reason, Rawls has failed to engage sufficiently with the communitarian objections. Rawls's reply is that "the significance of the original position lies in the fact that it is a device of representation or, alternatively a thought-experiment for the purpose of public and self-clarification" (Rawls, 2001, p. 17). That means that the original position is to be accepted as a device of representation in terms of Rawls's thought. Following the Kantian meaning, what Rawls proposes is that parties in the original position are deprived of any knowledge of themselves as particular agents. At this point, non-Kantian readings of *TJ* have misread the essence and intent of Rawls's original position by insisting that there is no fundamental correlation

between the Kantian sense of the autonomous person and the essential characteristics of these parties. Thus, Kantian moral agents would agree with this argument.

For instance, Sandel (*Liberalism and the Limits of Justice*, 1982) objects to Rawls's political method because it relies on Kantian morality and individuals. Sandel's fundamental thesis is that Rawls's theory of justice and its assumptions are based on the Kantian idea of the self as autonomous chooser of private ends and values because the parties have to choose a conception of justice and its principles in the original position. It does not matter what social position I find myself in because I know nothing about which specific identity I have: "I" could be anyone in my society. For Kant, human beings who have allowed their desires to define their objects of pursuit first have enslaved their will to these objects and are, thus, able to act only heteronomously, whereas the moral agent who acts solely from a law *he gives himself* is the complete determinant of all the actions he takes, and thus acts *freely and autonomously* (Hampton, 1980, p. 337). In her interpretation, Hampton acknowledges that, in this case, Rawls is following the Kantian attitude in identifying "justice" as Hampton finds a valuable similarity between Kant and Rawls. Firstly, in a similar way to Kant, Rawls states that a moral agent, according to the principles of justice, is to be described by looking at what sort of action an autonomous, rational person would regard as moral after pursuing the appropriate form of practical reasoning. In this respect, Hampton accepts that there is a similarity between the Rawlsian method of the original position and the "universalisation" method of the Kantian categorical imperative (Hampton, 1980, p. 337). Rawls's notion of justice forces the deliberator (under the veil of ignorance) to consider what a society ruled by this conception would be like for him if he were anyone in that society. Both Kant and Rawls regard this kind of universalisation procedure as descriptive of the correct operation of our practical reason. Hence, both procedures turn one's own interests into a single perspective that takes into account the interest of every rational agent. In *TJ*, Rawls himself says that he draws from Kant's approach in many respects, especially: "The person's choice as a noumenal self I have assumed to be a collective one" (*TJ*, p. 257). Here he suggests a Kantian method for understanding the contract argument, for instance:

The description of the original position interprets the point of view of noumenal selves, or what it means to be a free and equal rational being. Our nature as such beings is displayed when we act from the principles we would choose when this nature is reflected in the conditions determining choice. Thus, men exhibit their freedom, their independence from the contingencies of nature and society, by acting in ways they would acknowledge in the original position. (*TJ*, pp. 255-6)

Hampton concludes that Rawls was more Kantian than he realised. To interpret Rawls's Kantianism and its effects, Hampton (1980, p. 315) analyses the way in which Rawls's genuine, non-contractarian selection procedure provides a highly Kantian justification for his conception of justice.

Furthermore, I want to emphasise the importance of prioritising right over good in Kant's and Rawls's theory, which Johnson ignores when challenging Kantianism. Some modern liberals reject the utilitarian approach and follow Kant's argument that utilitarianism does not take into account distinctions between persons. Kantian liberals like Rawls do not sympathise with the feature of the utilitarian view of justice and prefer deontological ethics that values rights more. For Rawls and other Kant supporters speak more of the priority of liberty, emphasising "basic rights and liberties" species by a list: freedom of conscience, freedom of thought, association, that cannot be sacrificed for the general welfare or equality of opportunity. At that point, Rawls formulates two principles of justice: The first principle of justice is based on the principle of equal basic liberties that is to be expressed in the political institutions, whereas the second principle embodies priority to economic constitutions. That is why Rawls mentions that the theory of justice as fairness is not utilitarianism, but a deontological which does not specify the good independently from the right or does not interpret the right as maximising the good" (*TJ*, p. 26). It is assumed that justice as fairness within deontological theory characterises the rightness of institutions and acts independently from their consequences. According to Rawls, "each person possesses an inviolability predicated on justice that even the welfare of society as a whole cannot override" (*TJ*, p. 3). Justice's rights are not negotiable or susceptible to social interests (*TJ*, pp. 3-4). Maximising general prosperity has been replaced by morality that prioritises individual rights. Even Kantian liberals today need an explanation of rights without utilitarian assumptions. More importantly, Rawls, who has adapted that subject to contemporary political and social philosophy, expressly discusses it.

After the first scepticism understanding Rawls's Kantianism, the first reactions to this scepticism came from Stephen L. Darwall. In his 1976 essay "A Defence of the Kantian Interpretation," Darwall provides his first investigation of Rawls's Kantianism and supports a Kantian interpretation of Rawls's theory of justice, contrary to Johnson's claim. Darwall argues for a Kantian interpretation of Rawls's theory and directly addresses Johnson's misinterpretation. Darwall believes that Johnson made a mistake in his identification of the Rawlsian usage of the categorical imperative and autonomy because the principles of justice may still connect with the Kantian interpretation. Darwall disagrees with Johnson's distinction between Rawls's party in the original position and Kant's autonomy, saying Johnson's conclusion is erroneous. Darwall believes that one may think that the choice of principles in the initial position may

be accepted as a heteronomous choice, but the decision of actual rational beings, not directly in the original position, to act within the frame of principles is based on an autonomous decision. Thus, these actions on the principles are autonomous. Darwall writes: "Thus, if a rational being chooses to act on principles which would be acceptable to him if he were under the veil of ignorance...such a choice is by no means a choice on the basis of his interests and thus is not, on those grounds, a heteronomous choice" (Darwall, 1976, p.166). The veil of ignorance provides a methodological instrument for abstraction, according to Rawls, as the original position is a device of representation. It is vital to emphasise that Darwall's argument simply shows that Kantian argumentation can still be applied to justice. There is a connection between knowledge about the conditions of justice and human beings. Although the concepts of justice may not be universal, they are applicable to all autonomous agents who are rational under justice. Darwall deals with Johnson's critique, asserting that Johnson misinterprets the Kantian pure practical reason and Rawlsian rationality. "The core of Rawls's invocation of Kant in support of his theory is that there is a Kantian justification for the limits on choice of principles imposed in the original position," Darwall adds (1976, p. 165). The reasonableness of Rawls's theory of justice stems from the universal understanding that self-interest serves as a fundamental starting point. Additionally, in his 1980 essay "Is there a Kantian foundation of Rawlsian justice," Darwall goes to extend on this claim in an unconventional manner:

The complaint that the parties are assumed to be self-interested is a red herring in any case. Because of the veil of ignorance, the original position is not a perspective of self-interest but rather of an interest in selves or individuals as such. The assumption of self-interested motivation plays no essential role. The same principles would be chosen, and the same arguments for them found convincing, were the parties not assumed to be self-interested, but to be completely other-interested. (Darwall, 1980, p. 340)

According to this quote, it holds that the outcomes (principles of justice) are the same regardless of whether the agent is self-interested or other-interested. In brief, Darwall's response to criticism of Rawls's misinterpretation of autonomy is acceptable. His fundamental contention is that subsequent judgments to uphold the principles of justice in ordinary life are autonomous in the Kantian sense; even if the decisions could be formed in the original position, they may have been seen as heteronomous. According to Darwall's defence, the autonomous decision to adhere to heteronomous principles is in line with Kant's viewpoint. Similarly, Chaly (2015, p. 148) notes that even while the people in the original position make judgments based on heteronomous personal interests, Kantians regard decisions to stay connected to justice in everyday life as autonomous. For example, Chaly (2015, pp. 151-2) points out that it is possible



to conclude that “rational beings would make the decisions protecting their rationality and autonomy (which are, of course, inseparable for Kantian beings) against possible claims of empirical inclinations that will later in various contingent proportions become part of their natures. This would certainly mean treating not only humanity, but also any other form of reasonable being, as an end in itself.”

More crucially, Paul Guyer's (2018) statement appears more essential than Johnson or Levine. He analyses Rawls's theory in relation to Kant's political writings and moral philosophy. In his article “Primary Goods and Categories of Right: Rawls and Kant,” Guyer (2018, p. 581) notes that Rawls's theory applies only to Kant's moral philosophy and not to his political thought as explained in his “Doctrine of Right” in *the Metaphysics of Morals*. Because, here, Guyer argues that Kant's political philosophy, which is associated with his principles of the innate right to freedom, private acquired right, and public right, clarifies the relationship between Rawls's principles of justice and Kant's idea of basic liberties and primary goods. This claim will be fascinating to discuss in my research. One reason commentators can dismiss Rawls's Kantianism is that they focus on his moral philosophy rather than the relationship between moral and political philosophy. Guyer (2018) argue that Rawls's political theory seeks to deepen Kantian equality. They argue that Kant's classical liberal political works might be understood as liberal egalitarian.

Nicholas Tampio's 2007 article “Rawls and the Kantian Ethos” contributes to this debate. Tampio attempts to explain how Rawls interprets and modifies Kant's legacy. He also examines how Rawls conceptualises four Kantian elements: “the identification of the problem, the engagement with common sense, the construction of principles, and the authentication of principles” (Tampio, 2007, p. 79). Like Kant, Rawls develops the scope of justice by drawing out a certain mode of reasoning, according to Tampio. The goal is to “uncover the concepts and principles latent in our conceptualisation of the individual as rational and reasonable” (Tampio, 2007, p. 93). Thus, “a theory of moral sentiments” might characterise the reasonable, according to Rawls (*TJ*, p. 44). Tampio adds that “Rawls is establishing a purely Kantian basis, it appears that his foundational is substantially beholden to Kant,” but this does not mean he is (Tampio, 2007, pp. 79-102). Overall, Tampio's discussion of Kantianism is unclear. In his essay, Tampio acknowledges Rawls's Kantianism and contrasts the early and late Rawls. He also considers Rawls's late political position, which no longer has a Kantian base. He also aims to defend Rawls's position from communitarians like Sandel and leading Kant scholars like Larry Krasnoff, and Allen Wood, who argue that Rawls misinterprets Kantian moral doctrine when he proposes ideas like the “CI-procedure.” Contemporary “Enlightenment” arguments are also influenced by Rawls's interpretation of Kant. According to Wood,

“Kant’s ethical thought...exercises such a strong and continuing influence on us that replacing commonly accepted ideas about it with more accurate and less oversimplified ones might help to transform our conception of our own history and of ourselves as heirs of the Enlightenment” (Wood, 1999, p. 14; Tampio, 2007, p. 82). At that moment, Rawls is mostly responsible for Kant’s misrepresentation (Krasnoff, 1999, p. 400; Wood, 1999, pp. 374-5). Tampio disagrees with Wood and believes that Rawls can revive the Kantian tradition for such heirs of the Enlightenment. Kant’s motto for the Enlightenment was “have the guts to apply your own understanding!” (Kant, 1784, 8:37). Tampio asserts that Rawls attempts to implement this discourse with bravery while he establishes political methodologies relevant to late modernity. More importantly, Tampio states that “Rawls maintains, however, that he is Kantian when he refashions Kant’s ideas or creates new ones. For Rawls, the Kantian ethos (or spirit) impels us to exercise in our time the philosophical courage that Kant exercised in the eighteenth century” (Tampio, 2007, p. 100). The fundamental point of Tampio’s defence is that “Rawls considers a critical intellectual sensibility (or ethos), rather than a specific doctrine (e.g., the categorical imperative), as the most valuable component of Kant’s legacy” (Tampio, 2007, p. 79).

Thus, Rawls builds his political theory on a Kantian moral basis. Ana Marta González (2005, pp. 152-3) states that Rawls attempts to make Kant’s moral theory more reasonable “partly by putting more emphasis on Kant’s ethical writings other than the *Groundwork* and partly by bringing Kant down to earth, relating his moral theory closer to modern culture.” The Kantian interpretation of Rawls’s theory allows us to see Kant’s ethics in liberal democratic society’s political philosophy. I claim that Rawls’s theory of “justice as fairness” fills the gaps in Kantian moral theory’s political philosophy. Rawls stays Kantian. Kant’s moral theory requires “that there is no such sequence of given objects establishing the initial principles of right and justice among free and equal moral persons,” according to Rawls (1999a, p. 305). He recognises free and equal people in a normative sense, following Kant. Rawls’s Kantian conception of the free and equal moral person has had a major impact on classical liberal understanding of individual freedom, particularly its role in political justification. Namely, Rawls says, “the notion of morality as based on the rational choice among free and equal persons is the true contribution of Kant” (Rawls, 1999a, p. 305).

## Conclusion

In this article, I aimed to show Rawls’s link with Kant in its most basic form. In the light of the current literature, we obtain a general impression of Rawls’s Kantian framework. It appears that (a) there are disagreements between

interpretations of whether the first Rawls owes a debt to Kant or not. It is evident from the literature that some scholars have classified arguments as acceptable or inappropriate and addressed the Kantian basis of Rawls's principles of justice in their own writings. These arguments have been examined comparatively and respectively. This analysis of the literature primarily purposes to illustrate how ambiguous Rawls's Kantianism is among interpreters; and (b) the non-Kantian readings of *TJ* (e.g., Johnson, Levine, Hoffe, Wolff) have taken issue with Rawls's claim to Kantianism, which Rawls proclaimed. The original critique of Rawls's professed Kantianism revolves around his attempt to reconcile the egoistic-utilitarian motivation with the Kantian normative framework, and this was a frequent theme in Rawls's early criticism. Since they believed that Rawls had radically misunderstood Kant's theory in order to accommodate his views within the framework of Kant's moral philosophy.

I contend that these different perspectives are based on partial interpretations of Kant that are not nuanced, and I contribute significantly to the literature; Rawls is not responsible for this flaw. However, I think that this kind of interpretation is insufficient to support the Kantianism of Rawls's intellectual development. I should also note that there is a similar stereotyped view of Kantianism shared by both groups who believe Rawls is a Kantian and those who do not, as discussed in this study. Ultimately, the most important criticisms of Rawls's theory stem from their (mis)understanding of Kant.

## References

1. Adu, M. O. (2024). Immanuel Kant's Socio-Ethical Philosophy and Its Refinement by John Rawls. *APPON Philosophical Quarterly*, 3 (3). pp. 53-62. Available at. <https://acjoi.org/index.php/apponquarterly/article/view/6451/6244> (Accessed: 08 July 2025).
2. Audard, C. (2007). *John Rawls*. Montreal: McGill-Queen's University Press.
3. Budde, K. (2007). Rawls on Kant: Is Rawls a Kantian or Kant a Rawlsian? *European Journal of Political Theory*, 6 (3), pp. 339-358.
4. Chaly, V. (2015). An Interpretation of Rawls's Kantian Interpretation. *Con-Textos Kantianos, International Journal of Philosophy*, 1, pp. 142-155.
5. Cekić, N. (2022). Was Rawls a Kantian? *Prolegomena*, 21 (1), pp. 41-58. Available at: URL: <https://hrcak.srce.hr/file/404310> (Accessed: 05 April 2023).
6. Darwall, S. L. (1976). A Defence of the Kantian Interpretation. *Ethics*, 86 (2), pp. 164-170.
7. Darwall, S. L. (1980). Is there a Kantian foundation of Rawlsian justice. in: G Blocker and E. H. Smith (ed.), *John Rawls's Theory of Justice*, Athens: Ohio University Press, 311-345.

8. Davidson, A. I. (1985). Is Rawls a Kantian? *Pacific Philosophical Quarterly*, 66, pp. 48-77.
9. Dworkin, R. (1989,1973). The Original Position. *The University of Chicago Law Review*, 40 (3), pp. 500-533.
10. Freeman, S. (2003). Introduction. in S. Freeman (ed.) *The Cambridge Companion to Rawls*, Cambridge: Cambridge University Press, pp. 1-61.
11. Guyer, P. (2018). Principles of Justice, Primary Goods and Categories of Right: Rawls and Kant. *Kantian Review*, Volume 23 (4). pp.581-613.
12. González, A. M. (2005). John Rawls and the new Kantian Moral Theory. in Thom Brooks and Fabian Freyenhagen, (ed.) *The Legacy of John Rawls*, London: Continuum, pp. 152-176.
13. Grcic, J. M. (1983). Kant and Rawls: Contrasting Conceptions of Moral Theory. *J. Value Inquiry*, 17, pp. 235-240.
14. Hampton, J. (1980). Contracts and Choices: Does Rawls Have a Social Contract Theory? *The Journal of Philosophy*, 77 (6), pp. 315-338.
15. Hampton, J. (1989). Should Political Philosophy Be Done Without Metaphysics? *Ethics*, 99, 4, pp. 791-814.
16. Höffe, O. (1984). Is Rawls's *Theory of Justice* Really Kantian. *Ratio*, 26 (2), pp. 102-123.
17. Johnson, O. A. (1974). The Kantian Interpretation. *Ethics*, 85 (1), pp. 58-66.
18. Kant, I. (1785) [1998]. *Groundwork of the Metaphysics of Morals*. M. Gregor (trans. and ed.) Cambridge: Cambridge University Press.
19. Kant, I. (1784) [1996]. What is Enlightenment? *Practical Philosophy*, Mary J. Gregor (ed.) New York: Cambridge University Press.
20. Kant, I. (1797) [1996]. *The Metaphysics of Morals*. M. Gregor (trans.) Cambridge: Cambridge University Press.
21. Korsgaard, C. (1996). *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
22. Krasnoff, L. (1998). Consensus, Stability, and Normativity in Rawls's Political Liberalism. *The Journal of Philosophy*, 95 (6), pp. 269-292.
23. Levine, A. (1974). Rawls's Kantianism. *Social Theory and Practice*, 3 (1), pp. 47-63.
24. Mason, H. E. (1976). On the Kantian Interpretation of Rawls's Theory. *Midwest Studies in Philosophy* 1, pp. 47-55.
25. Pogge, T. (2007). *John Rawls: His Life and Theory of Justice*. Oxford: Oxford University Press.
26. Rawls, J. (1971) [1999]. *A Theory of Justice*. Oxford: Oxford University Press.
27. Rawls, J. (1999a) [1975]. A Kantian Conception of Equality. in John Rawls *Collected papers*, S. Freeman (ed.) London: Harvard University Press, pp. 254-267.
28. Rawls, J. (1999b) [1975]. Fairness to Goodness. in John Rawls *Collected papers*, S. Freeman (ed.) London: Harvard University Press, pp. 267-286.

29. Rawls, J. (2001). *Justice as Fairness: A Restatement*. E. Kelly (ed.) Cambridge, MA: Harvard University Press.
30. Sandel, M. (1982). *Liberalism and the Limits of Justice*. Cambridge: Cambridge University Press.
31. Scanlon, T. M. (1973). Rawls's Theory of Justice. *The University of Pennsylvania Law Review*, 121 (5), pp. 1020-1069.
32. Tampo, N. (2007). Rawls and the Kantian Ethos. *Polity*, 39 (1), pp. 79-102. Available at: URL: <https://faculty.fordham.edu/tampo/Tampo%20-%20Rawls%20and%20the%20Kantian%20Ethos.pdf> (Accessed: 06 August 2022).
33. Yang, H. (2025). From Moral Philosophy to Reflective Judgment: Rawls as a Successor to Kant, Reconceptualising Cross-Border Academic Mobility in the Interdisciplinary Research Agenda, pp 34-42. 40. Available at: [https://ir.library.osaka-u.ac.jp/repo/ouka/all/101334/RCBAMIRA\\_all.pdf#page=54](https://ir.library.osaka-u.ac.jp/repo/ouka/all/101334/RCBAMIRA_all.pdf#page=54) (Accessed: 15 September 2025).
34. Wolff, R. P. (1977). *Understanding Rawls: A Reconstruction and Critique of A Theory of Justice*. Princeton, NJ: Princeton University Press.
35. Wood, A. (1999). *Kant's Ethical Thought*. Cambridge: Cambridge University Press.