

An Interpretation of Rawls' "Kantian Interpretation"

Una interpretación de la "interpretación kantiana" de Rawls

VADIM CHALY*

Immanuel Kant Federal Baltic University, Russia

Abstract

Calling Kant a liberal philosopher requires important qualifications. Much like his theoretical philosophy, his political transcendentalism was and remains a great enterprise of navigating between the extremes of liberalism and conservatism, of balancing the "empirical" and the "pure" in human society, as well as in human mind. Of all the attempts to enlist Kant among the classics of liberalism, John Rawls' is the most impressive and thorough. However, it is hardly a success. The reason for this lies in a profound difference in their answering the fundamental (and therefore vague) question "What is Man?". This paper is an attempt to revise the debate about the extent of Rawls' Kantianism and to compare the meanings of basic concepts of what could be called "pure political anthropology" in Kant and in Rawls.

Key words

Rawls; Kant; "Kantian interpretation"; Political Anthropology; Autonomy; Rationality vs. Reasonability; Freedom vs. Liberty; Categorical Imperative; Humanity; Liberalism.

Resumen

Considerar a Kant un filósofo liberal requiere de importantes matices. Como su filosofía teórica, su transcendentalismo político fue y sigue siendo una gran empresa de navegación entre los

* Vadim Chaly, PhD, Head of Department of Philosophy at Immanuel Kant Baltic Federal University, Kaliningrad, Russia. E-mail for contact: vadim.chaly@gmail.com.

extremos del liberalismo y del conservadurismo, del equilibrio de lo "empírico" y de lo "puro" en la sociedad humana, como en ocurre en el caso de nuestra mente. Entre todos los intentos realizados para incluir a Kant entre los clásicos del liberalismo, John Rawls es el más impresionante y minucioso. Sin embargo, difícilmente alcanza éxito en su pretensión. La razón para ello reside en una profunda diferencia en su modo de responder a la pregunta fundamental (y, por ello, vaga) "¿Qué es el hombre?". Este artículo es un intento de someter a revisión el debate sobre el alcance del kantismo de Rawls y de comparar los significados de conceptos básicos acerca de lo que podría llamarse "antropología política pura" en Kant y en Rawls.

Palabras clave

Rawls; Kant; "interpretación kantiana"; antropología política; autonomía; racionalidad vs. razonabilidad; libertad vs. independencia; imperativo categórico; humanidad; liberalismo.

Coherent political philosophies begin with what a human being is, proceed to examine what it can/cannot be, and aim at what it ought and/or ought not to become. Classics were not ashamed to embrace this approach: Plato deliberately mixes anthropology with theory of state, Hobbes meticulously develops one to proceed to the other, and Kant's famous question "what is man?" can easily be given a political turn. Contemporary political philosophies are often careful to avoid this grand question at all, because normative modes it implies can be, and some have proven to be, speculative and oppressive. Rawls is exceptional, among other things, in his readiness to systematically address this question - that is why in *A Theory of Justice* the explanation of anthropological presuppositions of his theory occupies far more pages than "theory of justice as fairness" proper.

The normative approach, however dangerous, is unavoidable, because political philosophies cannot afford being purely descriptive, they also have to prescribe aims and means for the development (or conservation) of humanity, they have to guide us, irrespective of our belief in the very possibility of such guidance. Some contemporary political philosophies feature anthropological presuppositions that are implicit, assumed, unquestioned and might even prove conflicting. Kant's philosophy is prime example of the opposite. It provides us with detailed explication of "pure" mechanisms that are taken to be essential to a human being (Kant even goes further than that, claiming that he has a scheme for any being endowed with reason), and his "empirical" anthropology, although it received much less attention both from him and his readers¹, augments the normative part with recognizable descriptive image of what an actual human being is. Kant's political

¹ Until recently: cf. (Louden 2000, 2011)

philosophy is a recipe for (potentially) aligning the latter to the former, for straightening out the “crooked timber of humanity” - or at least indefinitely trying to².

To current liberal standards this might seem unabashed universalism, so much of contemporary interpretation of Kant’s political philosophy consist in rounding this edge, which, it seems, cannot be done without altering the very essence of Kant’s thought. It is to illustrate this point that I suggest to examine Rawls’ “Kantian interpretation” of *A Theory of Justice*, because its formulas include - and dramatically alter - some of Kant’s most basic notions: autonomy, freedom, rationality, ends, means, and so on.

1. Rawls

No one has done more to reinvigorate Kantian ideas in contemporary political philosophy than John Rawls, and his account of Kant is among most detailed and sympathetic. Despite early doubts (Nagel 1973, Levine 1974, Johnson 1974, 1977, Hicks 1974, Pogge 1981), it came to be accepted by many that “Kantian interpretation” of Rawls’ theory of justice generally succeeds in its purpose (Darwall 1976, Rawls 1980, DeLue 1980, Tampio 2007, Taylor 2011). I would like to, first, restate these early doubts mentioned, second, advance them further, taking into account some of Rawls’ more recent texts. The focus, however, remains on *A Theory of Justice*, because the assessment of Rawls’ later modifications requires a separate effort, which has to proceed from solid conclusions regarding the original theory.

Let us begin by collecting Rawls’ definitions of basic terms used in his “Kantian interpretation” to incorporate Kant’s arguments into the new theory. Any of these terms could be the starting point, and Rawls chooses **autonomy**: “Kant held, I believe, that 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” (Rawls 1999³, p. 222). Rawls’ definition relies on **freedom** and rationality, and while his brief

² Questioning the extent of Kant’s liberalism is far beyond the scope of this paper, but a preliminary attempt can be found in (Chaly 2014), albeit in Russian. The argument in that paper proceeds by emphasizing limits of Kant’s anthropological and, therefore, political optimism ascribed to him, for example, by Howard Williams (Williams 1983, 1992), and relies on Robert Louden’s analysis of Kant’s anthropology (Louden 2000, 2011).

³ The references are to the edition of *A Theory of Justice*, revised in 1999.

treatment of freedom in *A Theory of Justice* as will's ability to act despite natural constraints seems in line with Kant, his definition of rationality does not.

Rawls takes on **rationality** first in §25 "The Rationality of the Parties", and then in more detail in Chapter VII, "Goodness as Rationality" (§§60-68). His first definition is markedly instrumental: "The concept of rationality invoked here ... is the standard one familiar in social theory", where parties attempt "to win for themselves the highest index of primary social goods, since this enables them to promote their conception of the good most effectively whatever it turns out to be (Rawls 1999, p. 124, 125).

His second attempt defines rationality in a wider sense, which means including the answer to the question - what goods it is rational to want. The answer relies on notions of "desire of certain things as prerequisites for carrying out ... plans of life", preference for "wider share of liberty and opportunity and of wealth and income", and, most importantly, "self-respect and a sure confidence in the sense of one own's worth" (Rawls 1999, p. 348).

As rationality in broader sense seems to be the ability to form and maintain a "good" plan of life, its definition is due. Defining a **life plan**, Rawls writes, that "... a person may be regarded as a human life lived according to a plan. [...] [A]n individual says who he is by describing his purposes and causes, what he intends to do in his life" (Rawls 1999, p. 358). Also,

«[...] a rational plan of life establishes the basic point of view from which all judgments of value relating to a particular person are to be made and finally rendered consistent. Indeed, with certain qualifications (§83) we can think of a person as being happy when he is in the way of a successful execution (more or less) of a rational plan of life drawn up under (more or less) favorable conditions, and he is reasonably confident that his plan can be carried through (Rawls 1999, p. 359)».

An **end** can be explained as something that serves as a primary good for the fulfillment of a rational life plan (such are "for example, life, liberty and one's own welfare" from §83), an **interest** as something that leads to an end. Importantly, none of these is fixed:

«as free persons they [members of a well-ordered society] do not think of themselves as inevitably bound to, or as identical with, the pursuit of any particular array of fundamental interests that they may have at any given time; instead, they conceive of themselves as

capable of revising and altering these final ends and they give priority to preserving their liberty in this regard (Rawls 1975, p. 95)».

Now let us see how these notions work together in Rawls' formulations of **categorical imperative** (CI). He writes that CI is "a principle of conduct that applies to a person in virtue of his nature as a free and equal rational being" (Rawls 1999, p. 222), that "the principles of justice manifest in the basic structure of society men's desire to treat one another not as means only but as ends in themselves" (Rawls 1999, p. 156), and that "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" (Rawls 1999, p. 226). Rawls also writes at length in "Kantian interpretation":

«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 in virtue of his nature as a free and equal rational being. The validity of the principle does not presuppose that one has a particular desire or aim. Whereas a hypothetical imperative by contrast does assume this: it directs us to take certain steps as effective means to achieve a specific end. Whether the desire is for a particular thing, or whether it is for something more general, such as certain kinds of agreeable feelings or pleasures, the corresponding imperative is hypothetical. Its applicability depends upon one's having an aim which one need not have as a condition of being a rational human individual. The argument for the two principles of justice does not assume that the parties have particular ends, but only that they desire certain primary goods. These are things that it is rational to want whatever else one wants. Thus given human nature, wanting them is part of being rational; and while each is presumed to have some conception of the good, nothing is known about his final ends. The preference for primary goods is derived, then, from only the most general assumptions about rationality and the conditions of human life. To act from the principles of justice is to act from categorical imperatives in the sense that they apply to us whatever in particular our aims are. This simply reflects the fact that no such contingencies appear as premises in their derivation (Rawls 1999, p. 222-223)».

Thus, Rawls develops the following chain of definitions: autonomy is the combination of freedom and rationality; rationality is the urge to win for oneself the highest index of primary social goods, necessary to maintain a freely chosen plan of life;

plan of life forms the essence of a person, and its execution results in happiness and self-esteem. A plan of life includes ends, such as life, liberty and welfare⁴, as well as interests that are instrumental to these ends. Rawls, on one hand, admits that people are free to choose their ends, which should mean that transcendent ends are legitimate; on the other hand, even the "veil of ignorance" does not provide space for public discussion or bargaining about such ends.

2. The Critics

Some of the definitions in this chain came under criticism. In his 1974 article Andrew Levine argues, that Rawls' "Kantian interpretation rests 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). Analyzing Rawls' original position, which is intended to free our choice of basic principles of justice from what Kant would call "empirical" or heteronomous interests and inclinations, Levine shows, that the kind of considerations we do take in account in original position are not what Kant would call "pure". Levine states that Rawls tries to frame Hobbesian egoistic rationality in Kantian universalist terms, which leads to incoherence. Levine then argues that instrumental rationality, effectively employed by Rawls, is empirical and therefore heteronomous in Kant's sense, and suggests that, in order to account for autonomy, a different notion of reasonableness has to be introduced, which would fall in the domain of Kantian "Vernunft". He concludes that in Rawlsian 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). In his subsequent works Rawls attempts to elaborate the distinction between rationality and reasonableness, which, as we'll see, again falls short of complying with Kantian standards.

Oliver Johnson takes similar stance in his 1974 paper. He argues that Rawls and Kant advance different and irreconcilable models of human being, which make key notions and principles of Kantian moral philosophy - i.e. CI, autonomy, rationality - unusable in Rawls' theory. Johnson notes that individuals under "veil of ignorance" are still motivated

⁴ It is worth noting that "life, liberty and estate" is Locke's formula for property (cf. *Second Treatise*, §§87, 123), so, perhaps, one way to sum up the difference between Lockean classical liberalism and Rawlsian egalitarian liberalism is to compare "estate" to "welfare".

by what Kant would call heteronomous interests, that Rawls' notion of CI is in fact equivalent to Kant's notion of "counsels of prudence", which is a specific kind of hypothetical imperative, and that, he thinks most importantly, that Rawls' notion of rationality is opposite to Kant's.

Johnson's claims are contended by Stephen Darwall (Darwall 1976), who writes that the "Kantian interpretation" successfully develops moral foundations for Rawls' political theory, providing deep explanation and justification for the choice of principles in the original position. His main argument is that, although decisions in the original position are made in view of individual interests so could be considered heteronomous, later decisions to adhere to principles of justice in ordinary life are autonomous in Kantian sense. Still, as Johnson responds (1976), his argument regarding the original position stands.

3. Criticism revisited

The issues raised by the critics are essentially anthropological, they touch upon Kant's ultimate question of human nature, of what it means to be autonomous, to be rational and reasonable, to pursue interests and ends, etc. While Rawls centers his interpretation on autonomy, Levine, Johnson and Darwall turn towards rational agency. There's a good reason for this, because, as we've seen, Rawls defines autonomy through freedom and rationality, and of these two the latter receives the most attention from him. But the analysis of Rawls' definition of rationality done above reveals that a life plan is an even more basic notion, and it is in turn defined with mentioning of happiness. From §83 "Happiness and Dominant Ends" we learn that happiness is self-contained and self-sufficient, it comes as the result of one's success (or of belief into such success - this possibility was later addressed by Nozick's "experience machine" argument) in implementing one's life plan. So the ultimate good becomes for Rawls, at least in this important line of arguments, the same as for Aristotle, not Kant.

On this basis one might state that Rawls' understanding of human nature has more in common with the tradition of virtue ethics than with Kantian anthropology of autonomous agents capable of moral self-legislation and striving for the "kingdom of ends". There's nothing transcendental in a Rawlsian person, even in the original position all its substance

boils down to "what he intends to do in his life", meaning empirical life, finite and imperfect. Grand ideals and great aspirations fit neither "original position" nor later public life. Rawls is very clear about his reasons for the exclusion of anything grand, ultimate and universal from his political anthropology:

«[Reformation] introduces into people's conceptions of their good a transcendent element not admitting of compromise. This element forces either mortal conflict moderated only by circumstance and exhaustion, or equal liberty of conscience and freedom of thought. Except on the basis of these last, firmly founded and publicly recognized, no reasonable political conception of justice is possible. Political liberalism starts by taking to heart the absolute depth of that irreconcilable latent conflict» ("Political Liberalism", 1993, p. xxviii).

But for Kant being a person means above all operating ideas and aiming at ends that are beyond the empirical, that, while ever remaining problematic, have precisely those universal pretensions, which Rawls' liberalism finds dangerous. And it ought to be done not only privately, but also publically, politically. Rawls views grand ideas and final ends as the source of trouble that has to be contained; Kant views them as the main source of hope for humankind's future (the other source being "Nature" that indirectly supports our moral striving). Perhaps both are right, but one view has to prevail.

To illustrate Kant's universalism I shall not quote him at length, but only rely on his formulas of categorical imperative. The first "Formula of the Universal Law" contains the word "universal" ("allgemeines"), also present in its acquired name (Universalisierungsformel), it obliges us to evaluate our maxims from the point of view of a general reasonable being, which is not subject to personal empirical interests, but only to duty of promoting reasonability as such. In view of this Rawls' notion of "mutual disinterestedness", experienced by agents in original position, appears awkward - Kantian agents are profoundly, transcendently interested in each other's essential feature, i.e. reason as the end in itself, which happens to be common to all of them (and not empty, as Rawls says about final ends in the original position). Of course, precisely this is emphasized by the second "Humanity Formula". This formula is naturally the most popular in liberal philosophy, because on the surface it seems to prescribe treating individuals as ends. Robert Nozick gives a notable example of such reading of Kant: in *Anarchy, State and Utopia* he writes that "side constraints upon action reflect the underlying Kantian principle that individuals are ends and not merely means; they may not

be sacrificed or used for the achieving of other ends without their consent. Individuals are inviolable” (Nozick 1974, p. 30-31). Although later in the book Nozick does quote the “Humanity Formula” in its entirety, the word “humanity” is lost in this important paraphrase. It is humanity in individuals that Kant literally proclaims an end, not individuals *per se* (although, of course, humanity consists of individuals). The danger here is to replace the humans as bearers of reason by humans as persons having empirical interests and inclinations mixed with some moral sentiments, emotions, which we would call “humanity” as an attitude. This would deeply distort the spirit of Kantian philosophy.

It seems that Rawls also treats the second formula of CI rather loosely, when he writes that “the principles of justice manifest in the basic structure of society men’s desire to treat one another not as means only but as ends in themselves” (Rawls 1999, p. 156). When Rawls in *A Theory of Justice* speaks of “humanity”, he uses the word in the sense of certain attitude, moral character (see Rawls 1999, p. 428-9), not in the sense of humankind (which is discussed at p. 459). But it is worth noting that “Menschheit” of the second formula is frequently used by Kant throughout his writings in contexts where it could mean either humankind or the moral attitude of humanity - or, likely, is intended to mean both as being inseparable (for instance, extensively throughout *Anthropology* and *The Conflict of the Faculties*), - whereas the moral attitude of humanity is defined in *Metaphysics of Morals* as “Menschlichkeit” (MS, AA 6: 456-7). Kant also uses the word “Humanität” when he speaks specifically of the attitude (cf. A, AA 7: 282), and “Menschenrace” and “Menschengattung” when he speaks of humankind (as in *Bestimmung des Begriffs einer Menschenrace*)⁵.

This is not to say that treatments of the second formula by Rawls and Nozick are incorrect, it is only to note that they are biased towards individualism that is not quite Kantian in spirit. Allen Wood reminds that “Kantian ethics is grounded on the dignity of rational nature. It requires not only respect for individual rights and the equal worth of

⁵ The “problem of humanity” is addressed in detail in (Dean 2006). Although Richard Dean is defending the view that “humanity” is in fact primarily a moral attitude, he does admit that there’s a problem with this term. Indeed, the use of the same word “humanity” to translate two different Kantian terms gives rise to danger of misunderstandings akin to those well familiar in case of “Objekt” and “Gegenstand”. The “problem of humanity” is aggravated by the fact that it is rooted in “Problem der Menschheit”, to which there also seems to be no clear solution: for example, Karl Vorländer makes accent on “Menschheit” as “community of millions of reasonable beings” (Vorländer 1924, 298), whereas Rudolf Eisler opposes “Menschheit” to “Tierheit” as mode of conduct (Eisler 1930, 352). Having neither capabilities nor intentions to propose a solution, I can only point at the fact that the term “Menschheit” of the second formula cannot be treated casually without risk of distorting Kant’s meaning (whatever it might be).

human beings, but also the idea of a cosmopolitan community in which the ends of all rational beings must form a unity to be pursued collectively" (Wood, 1999). These observations lead to the conclusion that the deep breach between universalism clearly admitted by Kant's philosophy and liberal particularism of Rawls' theory seems to exist despite Rawls' attempts to downplay it. Of course, this is far from saying that Kant, the theorist of personal intellectual liberty and of perpetual peace, with his universalism paved the way to *polemos*, to what Rawls called "mortal conflict" and Carl Schmitt dwelled upon as political enmity. No doubt, there's plenty of individualism in Kant's philosophy to counterbalance universalism, but this balancing does make it harder to pigeonhole him as a liberal philosopher.

Here an objection can be and has been⁶ raised: can we conceive of a "pure" person, devoid of all "empirical" (and hence individual) features? Isn't having empirical interests part of being human, or even a rational being in general? Is there anything to discuss behind the "veil of ignorance" for Kantian transcendental subjects? Can we even use the plural "subjects" to denote this abstraction? While contemplating on this question, we could mix behind the "veil of ignorance" soon-to-be human beings and non-human rational beings and make them discuss the principles of justice being unaware of their future species. We might expect that in order for this discussion to happen at all, these species have to share some minimal conceptual features, among which has to be something like "belonging to the single unified spatio-temporal world" (in P.F. Strawson's terms⁷), and in this world, as Rawls notes, they have to experience scarcity of resources, i.e. be under condition where justice is necessary. Now these features both of the subject and of the world would qualify as empirical in Kant - it is contingent that our reason is limited (and often misguided and subjugated) by sensibility, and it is contingent that we do not inhabit the "kingdom of ends". If we get rid of these features, the very idea of original position seems to become empty as well.

Does this mean that Rawls' thought experiment only works for heteronomous beings, so there's nothing Kantian in his theory? Maybe not. We can probably further generalize it and imagine rational beings that would decide upon protecting their rationality, morality

⁶ Cf. Nozick 1974, p. 228.

⁷ See his "Bounds of Sense" (1966) and, more generally, "Individuals" (1959) for detailed arguments, which seem to become highly relevant here. Also relevant is Strawson's "Social Morality and Individual Ideal" (1961), where the notion of a life plan plays a central part and pluralism of ends is defended.

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 nature(s). This would certainly mean treating not only humanity, but also any other form of reasonable being, as an end in itself.

However, this would also mean severely modifying Rawls' theory in important respects, which he probably wouldn't have accepted. One of the ways to introduce this modification would be to expand the meaning of rationality, or, more precisely, subjugate rationality to a higher faculty. This "going transcendental" would bring back into the theory universalism Rawls was careful to avoid. If we expand the notion of rationality, people would not be at liberty while choosing ends (as in Rawls 1975), because the ends would be imposed by universal faculty of reason⁸. Autonomy would mean adhering to these ends, it would ensure freedom in a peculiar sense of being determined by ends, thus making freedom very different and at times even conflicting with what is commonly understood as liberty. Liberty, both negative and positive, would be relegated to the empirical choice of means to pursue ends in individual life, and also to choice of needs and inclinations one embraces and structures as interests. A reasonable (in Kantian, not Rawlsian, sense) as well as rational plan of life would then be if not strictly centered around, then at least loosely attracted towards ends that are transcendental, so might have universal pretensions of the character that political liberalism tries to avoid. Finally, categorical imperative, or a "CI-procedure", would work behind the "veil of ignorance" not to establish positive rules for constructing a just society, but to merely set something like negative, conservative "side constraints" protecting the "pure" from the "empirical", reason from instrumental rationality, universal from the particular, freedom from liberty. This is a sketch of an interpretation quite different from the one suggested in *A Theory of Justice*.

However, there are hints at the possibility of such interpretation in Rawls. For example, it does not suffice for the notion of "rational life plan" to mean a contingent assortment of personal empirical interests and aims (resembling Plato's "democratic man"), for "even rational plans of life which determine what things are good for human beings, the values of human life so to speak, are themselves constrained by the principles of justice" (Rawls, 1999, p. 348). So Rawls is not certain whether there's something

⁸ Here one has to stress that these ends cannot be enforced politically – "paternalism is the greatest despotism imaginable" (TP, AA 8: 291).

beyond what instrumental rationality wants, which counters its claims and serves as constraints ("justice" in broader sense), or these constraints are part of rational wants because, on deliberation, we notice that they serve our other empirical interests. Should Rawls choose the latter and embrace reason as mere "slave of the passions", he would justly be called utilitarian, Humean or Hobbesian; should he pursue the former, seemingly more Kantian, option, then his interpretation would have to be modified significantly.

It seems that Rawls is at least inclined towards the Kantian strategy, for in his later works (1980, 1993) he continuously turns to explaining the difference between being rational and being reasonable:

«Fair terms of cooperation articulate an idea of reciprocity and mutuality: all who cooperate must benefit, or share in common burdens, in some appropriate fashion as judged by a suitable benchmark of comparison. This element in social cooperation I call the *Reasonable*. The other element corresponds to the *Rational*: it expresses a conception of each participant's rational advantage what, as individuals, they are trying to advance» (Rawls 1980, 528)

«We tend to use "reasonable" to mean being fair-minded, judicious, and able to see other points of view, and so forth; while "rational" has more the sense of being logical, or acting for one's own good, or one's interests. In my own work, and in this discussion, the reasonable involves fair terms of cooperation; while the rational involves furthering the good or advantage of oneself, or of each person cooperating» (Rawls, 1993, p. 54).

But is this notion of "reasonability" Kantian enough? Being cooperative and fair-minded is a viable strategy in non-zero-sum games to advance one's interests that are again heteronomous (it is often rational to be reasonable). To establish autonomy, "reasonable" would have to mean more than that. Indeed, Rawls himself admits that Kant's use of the word "vernünftig" is "worlds away from "rational" in the narrow sense. It's a deep question (which I leave aside here) whether Kant's conception of reason includes far more than reason" (Rawls 2000, 164-65).

Another question can be raised: doesn't adherence to the demands of practical reason even in Kantian sense serve, upon yet another round of deliberation, to advance one's empirical interests? Don't we stick to them out of hope to someday find ourselves happy amidst the "kingdom of ends", following the line of reasoning (or, rather, "rationing") famously expressed in Pascal's Wager? This is a difficult question. However, there's one important respect, in which these interests seem to differ from the ones stemming directly from inclinations: they are not agent-specific, personal differences are irrelevant to them,

they are transcendental in Kantian sense, they are not subject to justice in Rawlsian sense because they're not subject to conflict. So, perhaps, we could call these interests transcendental ends, which we have to treat as essential in every reasonable being, including ourselves.

All this brings to the conclusion that Rawls himself formulates regarding Kantianism of *A Theory of Justice*: "... the adjective 'Kantian' expresses analogy and not identity; it means roughly that a doctrine sufficiently resembles Kant's in enough fundamental respects so that it is far closer to his view than to the other traditional moral conceptions that are appropriate for use as benchmarks of comparison" (Rawls, 1980, p. 517). The suggestion of this paper is basically to omit the word "fundamental" – also because it doesn't fit liberal vocabulary well – and end the sentence with "respects".

References

- Chaly, V. (2014), Kant mezhdru liberalismom i conservatismom (Kant between Liberalism and Conservatism). *Kantovsky Sbornik* 50 (4), 54-60
- Darwall, S. (1976), A Defense of the Kantian Interpretation. *Ethics* 86 (2), 164-170
- Dean, R. (2006), *The Value of Humanity in Kant's Moral Theory*. Oxford: Clarendon Press
- DeLue, S. Aristotle, Kant and Rawls on Moral Motivation in a Just Society. *The American Political Science Review* 74 (2), 385-393
- Eisler, R. (1930), *Kant-Lexikon. Nachschlagewerk zu Immanuel Kant*. Berlin: Georg Olms
- Johnson, O. (1974), The Kantian Interpretation. *Ethics* 85 (1), 58-66
- (1977), Autonomy in Kant and Rawls: A Reply. *Ethics* 87 (3), 251-254
- Levine, A. (1974), Rawls' Kantianism. *Social Theory and Practice* 3 (1), 47-63
- Louden, R. (2000), *Kant's Impure Ethics. From Rational Beings to Human Beings*. New York: Oxford University Press
- (2011), *Kant's Human Being. Essays on His Theory of Human Nature*. Oxford: Oxford University Press

- Nagel, T. (1973), Rawls on Justice. *The Philosophical Review* 82 (2), 220-234
- Nozick, R. (1974), *Anarchy, State and Utopia*. New York: Basic Books
- Pogge, T. (1981), The Kantian Interpretation of Justice as Fairness. *Zeitschrift für philosophische Forschung* 35 (1), 47-65
- Rawls, J. (1975), A Kantian Conception of Equality. *Cambridge Review* 96 (2225), 94-99
- (1980), Kantian Constructivism in Moral Theory. *The Journal of Philosophy* 77 (9), 515-572
- (1993), *Political Liberalism*. New York: Columbia University Press
- (1999), *A Theory of Justice*. rev. edn. Cambridge, MA: The Belknap Press of Harvard University Press
- (2000), *Lectures on the History of Moral Philosophy*. Cambridge, MA: Harvard University Press
- Strawson, P.F. (1959), *Individuals. An Essay in Descriptive Metaphysics*. London: Methuen
- Tampio, M. (2007), Rawls and the Kantian Ethos. *Polity* 39 (1), 79-102
- Taylor, R. (2011), *Reconstructing Rawls: The Kantian Foundations of Justice As Fairness*. University Park, PA: Pennsylvania State University Press
- Vorländer, K. (1924), *Immanuel Kant. Der Mann und das Werk*. Leipzig, Felix Meiner
- Williams, H. (1983), *Kant's Political Philosophy*. Oxford: Blackwell
- (1992), Kant's Optimism in his Social and Political Theory. In *Essays on Kant's Political Philosophy*. ed. H. Williams. Cardiff: University of Wales Press
- Wood, A. (1999), *Kant's Ethical Thought*. New York: Cambridge University Press

