



Title	Unsavoury implications of a theory of justice and the law of peoples: The denial of human rights and the justification of slavery
Author(s)	Steinhoff, U
Citation	Philosophical Forum, 2012, v. 43 n. 2, p. 175-196
Issued Date	2012
URL	http://hdl.handle.net/10722/159808
Rights	This is the pre-peer reviewed version of the following article: 'Unsavoury Implications of A Theory of Justice and The Law of Peoples: The Denial of Human Rights and the Justification of Slavery', The Philosophical Forum 43(2) (2012), pp. 175-196, which has been published in final form at http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9191.2012.00416.x/pdf.

Uwe Steinhoff

Unsavoury Implications of *A Theory of Justice* and *The Law of Peoples*: The Denial of Human Rights and the Justification of Slavery^{*}

Many philosophers have criticized John Rawls's *Law of Peoples*. However, often these criticisms take it for granted that the moral conclusions drawn in *A Theory of Justice* are superior to those in the former book. In my view, however, Rawls comes to many of his 'conclusions' without too many actual inferences. More precisely, my argument here is that if one takes Rawls's premises and the assumptions made about the original position(s) seriously and does in fact think them through to their logical conclusions, both *A Theory of Justice* and *The Law of Peoples* have abysmally counterintuitive and immoral implications. These implications comprise, among other things, the justifiability of slavery, the denial of human rights and the permissibility of genocide.

I. The original position

In *Political Liberalism* and in *The Law of Peoples*, Rawls seems not to have much use for principles for individuals. He is exclusively concerned there with the basic structure. (However, why 'principles for individuals' should not themselves be part of the 'basic structure' is anything but clear. However, this need not concern us here.) In contrast, in *A Theory of Justice* Rawls was still concerned with 'a complete theory of right', and such a theory 'includes principles for individuals as well'.¹ Among such principles he includes what he calls 'natural duties':

The following are examples of natural duties: the duty of helping another when he is in need or jeopardy, provided that one can do so without excessive risk or loss to oneself; the duty not to harm or injure another; and the duty not to cause unnecessary suffering.²

And he informs the reader:

A further feature of natural duties is that they hold between persons irrespective of their institutional relationships; they obtain between all as equal moral persons. In this sense the natural duties are owed not only to definite individuals, say to those cooperating together in a particular social arrangement, but to persons generally. This feature in particular suggests the propriety of the adjective "natural."³

^{*} This is the pre-peer reviewed version of the following article: 'Unsavoury Implications of *A Theory of Justice* and *The Law of Peoples*: The Denial of Human Rights and the Justification of Slavery', *The Philosophical Forum* 43(2) (2012), pp. 175-196, which has been published in final form at <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9191.2012.00416.x/pdf>. For very useful comments on earlier drafts of this article I thank Allen Buchanan, Joseph Chan, Robert E. Goodin, Jan-Christoph Heilinger, Luciano Venezia, and Héctor Wittwer as well as the audience of a talk I gave at the Department of Philosophy, University of Hong Kong, in particular Jiwei Ci, Alexandra Cook, Max Deutsch, Chris Fraser, Chad Hansen, Patrick Hawley, William Haines, Kelly Inglis, Joe Lau, and Timothy O'Leary.

¹ John Rawls, *A Theory of Justice* (Cambridge: The Belknap Press of Harvard University Press, 1999), p. 93.

² *Ibid.*, p. 98.

³ *Ibid.*, p. 99.

Yet, the big question one should ask here – interestingly enough, as far as I can see it never has been asked – is how Rawls’s theory of justice is supposed to generate duties with such a scope. After all, in *A Theory of Justice* he envisions an original position in which the members of a closed society decide behind a veil of ignorance rationally and *guided by their own interests* what rules should be established for *their* society. But even if they were to choose what rules (for example ‘principles for individuals’) should be valid *universally*: why should they accept *any* duties towards *outsiders*? How is that supposed to be in their own interest?

Rawls does not address this problem at all. He only states, in the paragraph following the above quote, that ‘even though the principles of natural duty are derived from a contractarian point of view, they do not presuppose an act of consent, express or tacit, or indeed any voluntary act, in order to apply’.⁴ And he assures the reader:

There is nothing inconsistent, or even surprising, in the fact that justice as fairness allows unconditional principles. It suffices to show that the parties in the original position would agree to principles defining the natural duties which as formulated hold unconditionally.⁵

Be that as it may, the question posed at the moment is why the parties to the original position should agree to the principles Rawls ‘formulates’ in the first place. Even if he were right in claiming that it would suffice to show that they would, the fact is that he never does show that. His claims in § 51 (‘The Arguments for the Principles of Natural Duty’) certainly provide no answer to the question. Indeed, the question seems to never have occurred to Rawls; he seems to simply assume that his procedure can justify the natural duties. But it cannot.

Let me explain why this is so: The original position does not invite everyone to take part in the decision procedure, but only those people who will later make up the society whose basic structure is to be chosen. In fact, the principles are conceived of as some kind of agreement between parties to a contract (Rawls considers his theory to be contractarian). Thus, Rawls puts the members of a society, which is conceived of as self-sufficient and ‘closed’, into an ‘original position’ to decide on the basic principles for the basic structure – the basic constitutional and economic makeup – of their society. In addition, they also decide on principles for individuals in the form of so-called ‘natural duties’. They have to make their decision behind a ‘veil of ignorance’, that is, they know the general laws of physics, psychology and economics (if there are such laws), but they do not know much about themselves – whether they are rich in that society or rather poor – and they know nothing about their natural endowments (such as intelligence, talents etc.). So the idea is that parties cannot unfairly influence the outcome to favour themselves. Justice is blind. However, the parties are guided not by *moral* concerns,⁶ but by their own rational self-interest. They are ‘mutually disinterested’,⁷ that is, they really do not care whether or not others (apart perhaps from their descendants) suffer. If they support certain principles for the basic structure of their society that will also benefit or protect other members of that society, then this is only because they do not know, behind the veil of ignorance, what their own position in that society will be. They act according to the motto ‘Better safe than sorry’. Thus, moral duties are not already presupposed by the parties to the original position, but the procedure is designed to *construct* them, as it were, in the first place. The parties to the original position will choose that set of duties that they consider the best choice in the light of a utility-risk calculation (we can leave certain details aside here).

Now, to my knowledge Rawls nowhere explicitly says that the parties of the original position ‘know’ or assume that their society is a closed one. However, since he wants ‘to formulate a

⁴ Ibid.

⁵ Ibid., p. 100.

⁶ Nor do they set out to develop a liberal constitution – they do not know yet whether the principles they will agree on will be liberal, communist, fascist or whatever.

⁷ Rawls, *A Theory of Justice*, p. 125.

reasonable conception of justice for the basic structure of society conceived for the time being as a closed system isolated from other societies',⁸ I do not see how this restriction on the scope of his enquiry can become operative in the original position except through the assumption of the parties that their society is indeed closed and isolated. Perhaps there *might* be other possibilities, but Rawls certainly does not mention any. In any case, the way suggested here is clearly the most natural one. But then the question arises: What rational advantage does it confer to the parties in the original position if they accept *any* universal duties? Indeed, why should they use any formulation of principles of justice that uses the term 'person' instead of only 'citizens'? As I said before, they will choose certain principles that also protect other members of their society and not only themselves because they cannot know which precise position in the society they will have. In contrast, they do know that they are *inside* the society and not outside. They also know that outsiders do not interact with insiders (that is what 'closed and isolated' means). Thus, there is no reason for them to choose principles that are universal in scope. For example, there is no reason why they should accept a duty not to harm or injure another (under the assumptions of the original position such a duty has no relevance for their interests at all); a duty merely not to harm or injure another *member of one's society* would do just fine. (Incidentally, as I will show at the end of this section, this is true even if, or especially if the parties to the original position do think that they will interact with outsiders. Thus, this deadlock for Rawls's theory arises either way.)

Obviously, this is not only a problem for the so-called 'natural duties' but also for Rawls's two principles of justice for institutions:

FIRST PRINCIPLE: 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.

SECOND PRINCIPLE: Social and economic inequalities are to be arranged so that they are both:

(a) to the greatest benefit of the least advantaged, consistent with the just savings principle, and

(b) attached to offices and positions open to all under conditions of fair equality of opportunity.⁹

Again, there is no reason why the parties in the original position should talk about 'persons' instead about 'citizens of our society', or about 'the least advantaged' instead of 'the least advantaged of our society'.¹⁰ (Accordingly, the 'equal basic liberties' would be formulated as *civic* liberties, as liberties of citizens, not of persons as such.)

One might object here – rather strangely, to be sure – that if the citizens know that they are living in a closed society, they might just as well talk about persons generally, for they know that it will not make any difference in practice. Since there is no interaction (we have to take the term 'isolation' very seriously), duties towards outsiders or duties outsiders might have towards insiders are simply irrelevant for practical purposes. This objection is strange because one could equally well say: 'Because there is no interaction between insiders and outside domestic animals,

⁸ Ibid., p. 7.

⁹ Ibid., p. 266.

¹⁰ I encountered the objection that while Rawls's use of the term 'persons' instead of 'citizens' might be misleading, it makes no real, logical difference because in the situation Rawls asks us to envision the two terms are extensionally equivalent. This objection fails. First, even if the two terms were extensionally equivalent, intensional non-equivalence would be quite sufficient for my argument. As I argue below, a society that *allows* the enslavement of outsiders would still be unjust even if it factually cannot lay hand on outsiders because the society is closed or because, perhaps unbeknownst to that society, outsiders simply do not exist. Second, the two terms are not extensionally equivalent: Rawls asks us to envision a society that is closed, not a world in which outside of that society no other persons exist.

the citizens could well render their principle still more universal and inclusive by replacing “persons” with “persons and outside domestic animals”.’ Indeed, they ‘could’, but I see no reason why they should or would. If they want to be accurate and mention the intended scope of the principles in the principles themselves (which is a good idea), they will use the term ‘citizens’, not the term ‘persons’. At the very least, nothing in Rawls’s description of the original position can exclude such a more accurate choice of terms.

It might be objected that the revised principle (the one referring to citizens instead of persons) is not on the list Rawls wants to present to the parties in the original position. But even if that were true, intuitionistic ‘prima facie principles (as appropriate)’¹¹ actually are on his list, and if appropriateness with regard to the list is measured in terms of the historical importance of the items on the list, then one might note that ethnocentric and nationalist principles have had an enormous appeal in the history of philosophy and in human history in general. They are definitely ‘traditional conceptions of justice’.¹²

Besides, even if the revised principle or some equivalent set of principles were not on the list Rawls presents in *A Theory of Justice*, this would certainly not undermine my point. For one thing, the principle should be on the list. As he correctly states elsewhere:

If it is objected that certain principles are not on the list ... those principles must be added to it.¹³

For another thing, Rawls states that he makes ‘no attempt ... to deal with the general problem of the best solution’. Instead, he wants, mostly for practical reasons, to ‘limit the argument throughout to the weaker contention that the two principles would be chosen from the conceptions of justice’¹⁴ on a certain list that he presents. However, he does claim that ‘the two principles would be shown to be preferable once all agree that they are to be chosen over each of the other alternatives’.¹⁵

But then, in the same vein, I can and do claim that the revised principle is preferable in the light of Rawls’s own theory of justice if it would be chosen over the original principle. I have already demonstrated that indeed it would be chosen over the original principle, at least if people in the original position are interested in clarity. If they are not interested in clarity, then there is still no reason to *prefer* the original principle. Thus, Rawls’s theory of justice cannot exclude that the revised principle would be chosen.

However, it might be pointed out that Rawls also espouses ‘certain formal conditions that it seems reasonable to impose on the conceptions of justice that are to be allowed on the list presented to the parties’.¹⁶ Two such conditions that may appear relevant in the present context are the generality condition and the universality condition.

First of all, principles should be general. That is, it must be possible to formulate them without the use of what would be intuitively recognized as proper names, or rigged definite descriptions. Thus the predicates used in their statement should express general properties and relations.¹⁷

Next, principles are to be universal in application. They must hold for everyone in virtue of their being moral persons. Thus I assume that each can understand these principles and use them in his deliberations. ... Moreover, a principle is ruled out if it would be self-contradictory, or self-defeating, for everyone to act upon it.¹⁸

¹¹ Ibid., p. 107.

¹² Ibid., p. 106.

¹³ John Rawls, *Justice as Fairness: A Restatement* (Cambridge: The Belknap Press of Harvard University Press, 2003), p. 83.

¹⁴ Rawls, *A Theory of Justice*, p. 107.

¹⁵ Ibid., p. 106.

¹⁶ Ibid., p. 112.

¹⁷ Ibid., p. 113.

¹⁸ Ibid., p. 114.

Do the revised two principles violate these conditions? There is no difficulty with the universality condition, but there might be one with the generality condition. In particular, if ‘citizens’ is to be understood as ‘citizens of *our* society’, then the principles are not general any more. It is biased towards a certain group. Why should only citizens of our society enjoy such rights, why not also those of other societies?

It should be noted, however, that the original principles, depending on their interpretation, might run into the very same difficulties. To wit, contrary to the actual wording of the difference principle, Rawls does not want the basic structure of a society to arrange social and economic inequalities in such a way that the least advantaged are benefited, but in such a way that the least advantaged *of that society* are benefited. After all, Rawls does not support the extension of the difference principle to the global level. (The actual wording of the principles, taken literally, of course does imply such an extension.)

Yet, this in itself does not yet lead Rawls’s two principles into a contradiction with the condition of generality. They could be interpreted as a somewhat misleading and confused way of trying to say the following (RP):

(1) Each citizen of a society is to have an equal right against his society to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all members of the society.

(2) Social and economic inequalities within a society are to be arranged so that they are both:

(a) to the greatest benefit of the least advantaged members within the society, consistent with the just savings principle, and

(b) attached to offices and positions open to all members of the society under conditions of fair equality of opportunity for all members of the society.

These principles do satisfy the generality condition. And *these* principle are the ones Rawls actually has in mind (and that so far all interpreters have implicitly attributed to him); he only expresses them badly with his original formulation.

Thus, we have the following situation: Rawls’s original formulation of his two principles of justice certainly excludes slavery – and implies the application of the difference principle to the global level. However, there is no reason why the parties to the original position should agree to those principles. They would rather prefer to agree to principles that give *them* rights and not other persons. However, such principles would not satisfy the universality condition. The revised principles (RP) do satisfy that condition. In fact, these principles are not so much ‘revised’ as more accurate restatements of what Rawls had in mind. However, these principles do *not* exclude slavery. Yet, they would in the original position be preferred to the initial formulation of Rawls’s principles, for these revised ones clearly and explicitly refer to both the subjects and the subject of the original position, namely to citizens and their society. The principles Rawls formulates do not.

Accordingly, there is also no reason for people in the original position to agree to duties like ‘Do not torture others for fun’. ‘Do not torture other members of your society for fun’ would be much more agreeable, for the reasons stated. There would also be no agreement on principles like ‘Do not enslave people’ or ‘Do not commit genocide’ if the alternative more restricted principles ‘Do not enslave members of your own society’ and ‘Do not commit genocide against members of your own society’ are also on the list. The latter two principles (which are, of course, both general) would be preferred to the former.

There is a further reason why this is the case. My argument so far would even work under the assumption that the parties in the original position are, as it were, *fanatically convinced, with not the slightest doubt, not even a ‘paper doubt’*, that their society is isolated and closed. However, ascribing this state of mind to the parties in the original position contradicts the rationality assumption: the parties in the original position are supposed to be *rational* individuals. Moreover, when Rawls takes up the question as to why people even in a well-ordered society would want to have a penal system and a government with coercive powers, although citizens in a well-ordered society ‘know that they share a common sense of justice and that each wants to adhere to the existing arrangements’, his explanation is that ‘they may nevertheless lack full

confidence in one another'.¹⁹ Thus, even their *knowledge* that others will behave so and so is compatible with some residual doubts whether others might not still perhaps act otherwise in the end. Therefore it is rational for them to install some safeguards in the form of sanctions – after all, you never know *for sure*.

But then the assumption of people in the original position that their society is closed and isolated is, if we are talking about rational individuals, also susceptible to residual doubts, and hence parties in the original situation will again act on the motto 'Better safe than sorry'. Thus, the (rather daring) argument referred to above that if the parties to the original position know that they will not interact with outsiders anyway, they might as well happily attribute all kinds of rights to those outsiders without having to fear being taken on their word collapses. For now they will say, rational and self-interested as they are: 'Well, yes, it seems that we will not interact with outsiders. But we do not know that *for certain*. So why take the risk? We'd better *not* burden ourselves with duties towards them.' Wisely spoken.

But could this argument not be turned around? If the parties to the original position take into consideration that there might, in the end, be some interaction with outsiders, then a duty like 'Do not enslave anybody' would also benefit them, wouldn't it? Not only they themselves would be restrained by duties to outsiders, but outsiders would also be restrained by those very same duties towards them. And that can benefit the insiders.

This is a misunderstanding of the workings of the original position (and in explaining why, I will make good on my above promise to show that even in the case that the parties to the original position assume that they will interact with outsiders they would still not subscribe to principles that rule out maltreatment of those outsiders). As Rawls explains, in the original position the parties can rely on each other to understand and to act in accordance with whatever principles are finally agreed to. Once principles are acknowledged the parties can depend on one another to conform to them.²⁰

On one another. Not on outsiders. In other words, in the original position the parties would assume that choosing a principle like 'Do not enslave anybody' protects them and outsiders from being enslaved by the parties to the contract, that is, by the insiders (who can be relied on to follow the chosen principles); but they would *not* assume that outsiders – who are, after all, not parties to the contract anyway – will also abide by this principle. Thus, such a general norm would give outsiders more protection than insiders. But then choosing it would be irrational if a more advantageous norm is at hand. And of course there is. 'Do not enslave members of your own society' is precisely such a norm (of course, in the light of the two principles of justice, choosing such an additional rule would be superfluous anyhow). It gives them all the protection that the first principle affords, without giving outsiders any possibly unreciprocated protection. From the standpoint of self-interested individuals in the original position it is the superior principle. Similar considerations hold for the above-mentioned alternative principles regarding recreational torture and genocide. There is absolutely no reason for the parties to the original position to choose principles that would prohibit the torturing or slaughtering of outsiders.

At this point I have encountered the following objection: Since Rawls limits his theory of justice to 'the basic structure of society conceived for the time being as a closed system isolated from other societies',²¹ his theory is simply not applicable to relations between insiders and outsiders; it was never meant by Rawls to apply to them. Even if it were true, so the objection goes, that the parties to the original position would choose principles that do not exclude the enslavement of outsiders (and it definitely is true, as my above arguments show), it would not follow from this that these principles *allow* the enslavement of outsiders (according to the liberal principle 'What is not prohibited is permitted'). This would only follow if the principles were applicable to the relation between insiders and outsiders, which, again, they are allegedly not.

¹⁹ Ibid., p. 211.

²⁰ Ibid., p. 125.

²¹ Ibid., p. 7.

I think that this objection overlooks several things. First of all, in *A Theory of Justice* Rawls is, contrary to some of his pronouncements, *not* only concerned with the basic structure of a society, but also with natural duties. Ignoring this fact – as seems to happen quite often – does not make it go away. As already quoted, these duties are according to Rawls not only owed ‘to those cooperating together in a particular social arrangement’ (like a basic structure, for instance), ‘but to persons generally’. He goes on to explain:

One aim of the law of nations is to assure the recognition of these duties in the conduct of states. This is especially important in constraining the means used in war, assuming that, in certain circumstances anyway, wars of self-defense are justified ...²²

Now, the means used in wars of self-defence are means used against *outsiders*, namely against the individuals of the attacking state. Thus, if constraining the means used in war is a way of recognizing these natural duties, these natural duties necessarily also refer to relations between insiders and outsiders as well as between outsiders and outsiders. In short, the premise upon which the objection rests is simply wrong.

Second, if *A Theory of Justice* and the principles of justice it tries to justify were really only applicable to closed and isolated societies, the theory – and this fact should then better be taken seriously, instead of glossing over it – *would have no practical significance whatsoever*. The reason is simple: There are no closed and isolated societies. But of course Rawls actually does attribute a very high significance to the two principles of justice, to wit, ‘the two principles of justice provide an Archimedean point for appraising *existing* institutions, as well as the desires and aspirations which they generate’.²³ Obviously, the principles of justice could hardly offer an Archimedean point for appraising existing institutions if they were not applicable to existing institutions. Existing institutions, however, involve relations between insiders and outsiders.

Third, if one insisted that the ‘logical’ way to apply ‘Rawls’s’ theory to outsiders were to have the outsiders already participate in the domestic original position, then this can hardly be an objection against my criticism of Rawls’s theory, but should rather be directed against Rawls himself. I do not criticize the theory Rawls should have developed in the opinion of some people; I criticize the theory he did develop. His refusal to include (representatives of) all who are affected by certain duties or institutional arrangements in the original position (in which, after all, these duties and the principles guiding the design of the institutional arrangements are to be chosen) did not first come up in *The Law of Peoples*, it is already there in *A Theory of Justice*. After all, as just demonstrated, the natural duties are also to apply to outsiders, but Rawls nevertheless derives those duties from an original position only including insiders. In the same vein, the law of nations does not only affect states, it also affects individuals. However, when in *A Theory of Justice* Rawls describes the international original position, he says that the contracting parties are ‘representatives of states’ – thus, he makes no reference to representatives of individuals whatsoever.²⁴

Besides, it should be noted that the whole idea of dealing with the relations between insiders and outsiders (and hence with the problem that outsiders are affected by the basic structures of societies that are not their own) by the inclusion of outsiders in the domestic original position completely contradicts the division of labour between the domestic and the international original position already laid down in *A Theory of Justice*. For if the domestic original position included outsiders, either the law of peoples would become completely superfluous, or the principles arrived at on the first (once domestic) level would contradict the principles arrived at by the representatives of states.²⁵ In any case, if Rawls thought that the relations between insiders and

²² Ibid., p. 99.

²³ Ibid. p. 456, my emphasis. Compare also pp. 215, 216 and 231.

²⁴ Ibid., p. 331f.

²⁵ Besides, if Rawls were right in claiming that in the original position people would agree on the two principles of justice, then of course in an original position including outsiders principles would be chosen that apply these principles, including the difference principle, to the whole

outsiders should be dealt with by including the latter in the domestic original position, he could easily have said so instead of merely proposing a domestic original position exclusively inhabited by insiders and two international original positions exclusively inhabited by ‘peoples’.

I have also encountered the objection that since the principles of justice say *nothing* about slavery, they cannot be construed to allow it. The reasoning now goes like this: When I sign a contract to rent an apartment, I do not sign a contract that says that slavery is not wrong – but that does evidently not imply anything about my attitude towards slavery, in particular, it does not imply that I have nothing against it. In this objection, signing the contract about renting an apartment is supposed to be somehow analogous to the agreement reached in the original position.

However, it clearly is not. The parties in the original position do not sign a contract to rent an apartment. They sign a contract that specifies the principles which shall be used to evaluate the justice of the basic structure of their society, that shall, as Rawls says, be used ‘to regulate all subsequent criticism and reform of institutions’.²⁶ They agree on this contract. They thus agree: ‘The principles we have just chosen, these two principles of justice, shall be used in evaluating the justice of the basic structure of our society. A basic structure being in line with these principles is just. A basic structure not being in line with these principles is unjust.’ While these principles might not *explicitly* say anything about slavery or outsiders in general, there is nevertheless no way that a basic structure that allows citizens to enslave other citizens could be in line with the two revised principles of justice. Thus, according to those principles, enslaving other citizens cannot be just. There is, however, on the other hand nothing in those principles that would imply that a basic structure that allows freedom of the press, democratic elections or (as I show) the enslavement of non-citizens is unjust. Thus, according to those principles of justice, enslaving non-citizens can be as just as freedom of the press or democratic elections. In short, when I sign a contract stating the conditions under which the basic structure of our society is just and my having signed this contract is considered as a proof of its validity, then the fact that nothing in these conditions implies that a basic structure allowing rape or slavery is unjust does definitely imply something about rape or slavery; it implies precisely that a basic structure allowing rape or slavery can be just. *This is* analogous to the case of the original position and the principles of justice chosen in it.

It has to be emphasized, by the way, that the fact that a closed society cannot act on its unjust basic structure against outsiders because *ex hypothesi* it does not interact with outsiders does not make it less unjust. A serial-killing maniac who would kill the minute he is let out of jail does not become a just person through the mere fact that he is kept behind bars and thus does not get the chance to murder someone. Similarly, a society whose basic structure allows the enslavement of outsiders does not become just through the mere fact that it is (fortunately) kept inside its borders and thus cannot act on its unjust inclinations towards outsiders. To use an analogy: A philosopher who develops principles of justice applicable to the character of isolated persons only, that is, to the character of persons who for some reason will not interact with others, has obviously still made a rather grave mistake if he proposes principles for a just character that in all seriousness imply that an isolated person can be just who is absolutely convinced (and inclined to act accordingly if she had the chance) that it would be quite all right to torture or rape or kill or enslave others for fun. To respond, ‘But my principles are only to apply to isolated persons’ does not in the least refute this argument; it completely misses the point. After all, the argument already *grants* that the principles are only applicable to isolated persons – and points out that this changes nothing. The fact is that a person who is so egomaniacally and sadistically convinced or inclined is clearly *not* a just person, whether she can actually interact with and harm other people or not. In the same vein, a basic structure that allows the enslavement and all

planet. This, however, is something Rawls adamantly opposes. Thus, ‘amending’ Rawls’s theory by including the outsiders on the domestic level would lead to a theory that is definitely no longer the one offered by Rawls– neither in his later work, nor in *A Theory of Justice* itself.

²⁶ Ibid., p. 12.

kinds of abuses of outsiders is *not* a just basic structure, whether the society is closed – and thus does not get the chance to harm others – or not.

Finally, let me note an objection that is both ingenious and desperate. The problem with the above objection was, among other things, that a basic structure that allows the enslavement of outsiders is still unjust even if there are no outsiders in reach. Thus the stubborn insistence on the claim that Rawls, allegedly, is only concerned about principles of justice for *closed* societies does not help even if the claim were true, since a basic structure, which includes the constitution and other laws of the society, simply cannot avoid implying something about the relation to outsiders – *whether the society is closed or not*.

But what if, so the new objection goes, Rawls meant his two principles of justice to apply *only to the interactions between the members of the society*? Well, I am quite willing to admit that *then* the two principles of justice do really imply absolutely nothing about the enslavement of outsiders. Therein lies the ingenious part of the objection. The desperation of the objection, on the other hand, lies in the fact that it completely ignores Rawls's insistence that in the first place he wants to find principles of justice *for the basic structure of a closed society*. He nowhere says or suggests or proceeds in a way that would suggest (remember the case of the natural duties if nothing else) that he merely wants to find principles for the interaction between members of the society. Thus, this objection completely misses the point of Rawls's endeavour.

For the reasons given, I conclude that the two principles of justice *are* applicable to the relations towards outsiders. And they do allow the enslavement or genocide of outsiders.

However, if people try to escape my contention that Rawls's theory of justice implies the justifiability of slavery and genocide by continually insisting that Rawls's theory actually implies *nothing* about the treatment of outsiders, then this should be explicitly recognized and the conclusion be drawn: From the perspective of Rawls's *Theory of Justice* the genocide and enslavement of outsiders can neither be endorsed *nor criticized*. And since Rawls's *Law of Peoples explicitly* applies to the relation of states towards outsiders and between each other, and liberal and decent peoples in the two original positions would, as we will see in section III, choose principles allowing the genocide and enslavement at least of persons who do not belong to liberal or decent peoples, Rawls's *overall* theory would still imply the justifiability of slavery and genocide.

II. The Second Stage

The original position is only the first stage in a four stage sequence:

Thus I suppose that after the parties have adopted the principles of justice in the original position, they move to a constitutional convention. Here they are to decide upon the justice of political forms and choose a constitution: they are delegates, so to speak, to such a convention. Subject to the constraints of the principles of justice already chosen, they are to design a system for the constitutional powers of government and the basic rights of citizens. ... Since the appropriate conception of justice has been agreed upon, the veil of ignorance is partially lifted. The persons in the convention have, of course, no information about particular individuals: they do not know their own social position, their place in the distribution of natural attributes, or their conception of the good. But in addition to an understanding of the principles of social theory, they now know the relevant general facts about their society, that is, its natural circumstances and resources, its level of economic advance and political culture and so on.²⁷

There is no need here to go into the remaining two stages. In the present context it is sufficient to point out that in the second stage the 'delegates' know, for example, whether they are living in a slave-holding society. Slaves, however, are not members of the society – there is no 'co-operation' with them; rather, they are forced to work. They are simply not a party to the (hypothetical) social contract. It also has to be noted that the delegates have of course not lost their rational self-interest with the partial lifting of the veil of ignorance. Thus, they will want to

²⁷ Ibid., p. 172f.

arrange for a constitution that uses the lucky circumstance that there are slaves available in such a way as to enhance the gross national product (or to provide other services which I can leave to the fantasy of the readers) – subject of course to the two ‘revised’ principles of justice. Thus, the slavery will be so arranged that the worst-off *members of the society* are benefited by the slave labour. There is no reason in sight to abolish slavery. In short, Rawls’s theory of justice justifies slavery. In fact, under certain circumstances it makes slavery *mandatory*: A constitution outlawing slavery although slavery could be so arranged that it benefits the worst-off members of one’s society more than any alternative arrangement would be grossly unjust in the light of Rawls’s theory.

Here it might again be insisted that Rawls, at least as regards his ideal theory, worked under the assumption of a society that is closed and isolated from others. Thus, even if, as I argue above, the principles of justice that would be chosen would indeed allow the enslavement of outsiders, there are simply no outsiders available in ideal theory that could be enslaved. Thus, aren’t those principles not unobjectionable as long as one only considers closed and isolated societies?

First of all, one needs to remember here a point already made above: even if one remains completely within the narrow confines of ‘ideal theory’ and argues that Rawls only wanted to develop his principles of justice for closed societies, the principles that would be chosen in the original position and the constitution that would be engendered by them are *still* objectionable. The reason, to repeat, is quite simple. Consider a judge who is determined to give black people harsher punishments for the same crime under the same circumstances than white people. This is an unjust judge *even if in fact he will never have the opportunity to judge black people*. That he will never have this opportunity means that he will, as a judge, never *treat* black people unjustly; it does not mean, however, that he himself is just. Similarly, a constitution that does not on principle rule out the enslavement of outsiders (this must not happen explicitly, it can happen by implication) is still unjust even if it will never in actual fact lead to the enslavement of somebody because the potential slaves are lucky enough to be out of reach of the constitution.

Second, even if the principles were indeed unobjectionable as long as one only considers closed and isolated societies, this still would be small comfort. As already pointed out, the principles of justice are to guide reform efforts in the real world. But this means that they could not offer any standards by which to criticize slavery as such; on the contrary, they would under certain real circumstance call for the institutionalization of slavery. This means that they are complete failures as principles of justice.

To escape these conclusions some might now be tempted to simply claim that the slaves *are* part of the society. But that would be completely *ad hoc*. Why should they be part of the society (society is meant here in a strong sense: as those people who are part to the hypothetical contract entered into in the original position)? Maybe it is because they are within the territorial boundaries of the society? But then Rawls’s theory cannot distinguish anymore between citizens and mere residents or visitors (or soldiers of an invading army, for that matter); yet, nothing in Rawls’s writings suggests that he would be willing to give up this distinction (and I see certainly no reason why we should).

Besides, even if we granted for a moment, rather inanely but for the sake of argument, that all persons within certain territorial boundaries are automatically part of one and the same society (what about tourists?), this still would not solve the problem of slavery, for a society can have its slaves outside of its borders, for example in colonies. This problem could only be solved by claiming, for example, that everybody who lived in or entered a British colony in the 18th Century thereby automatically became British. Perhaps one might think that it could also be solved the other way around, namely, that the British automatically became a member of the society they invaded and made a colony. The problem for both ‘solutions’, however, is that with regard to many colonies it was very difficult to discern where the ‘borders’ of the respective alleged ‘society’ were. Thus, if a society can be a society only if it is enclosed by clear borders, then there were (and are) many territories without societies – and thus without the social contract. By entering these territories the colonialist (or the colonized) would therefore not have become parties to a common (hypothetical) social contract with the other people there. Hence,

the colonialist would have been perfectly free, in light of the implications of Rawls's theory, to enslave those other people.

In addition, even if we granted that the whole Earth is made up of territories confined by clearly defined borders (but what about the oceans and hence about slavery on ships?), this still would not solve the problem. Of course, on the hypothetical concessions we just made for the sake of argument, people within territory A would not be permitted to enslave other people within territory A under the 'revised' principles of justice (and, again, these 'revised' principles are not only what Rawls's theory implies but also what he actually had in mind). Thus, people of country B would not be allowed to invade A and enslave people there, for by invading A they would themselves become members of A and thus be bound to the (hypothetical) social contract. However, it would be perfectly permissible for A to point its missiles equipped with biological or chemical warheads on B, telling the people there: 'If you do not work for us in the mines on your territory and throw the diamonds over the border so that we can have them, we will kill you all, or many of you', thus enslaving them in this way. It would also be perfectly justified – probably even mandatory for a responsible delegate in the second stage of the original position – to set up a scheme by which certain people on B's side are paid for enslaving other people on B's side. To be sure, the former group would violate the civic rights of the latter, but the people on A's side would *not* violate any duties towards the slaves on B's side, for they have no such duties. Their duties are to their own compatriots.

Thus, the *Theory of Justice* implies the justifiability of slavery. In addition, there is of course also nothing in the theory of justice that would prohibit the people in A from exterminating all people in B by firing their genocidal weapons onto B's territory and to then annex it. The *Theory of Justice* is compatible with genocide.

III. The Law of Peoples

The *Theory of Justice* does not recognize any human rights and, in fact, the term 'human rights' is not used there once, as far as I can see. *Political Liberalism* is also somewhat short on human rights; however, the term is used there on occasion. In *The Law of Peoples*, by contrast, the term 'human rights' is used quite often. Of course, many critics have argued – convincingly and correctly – that Rawls's list of human rights is disappointingly short. A few other critics have argued that it is, moreover, rather mysterious how the international original positions envisaged in *The Law of Peoples* should be able to justify *any* human rights.²⁸ These critics, too, are quite correct. However, it is worth emphasizing the implication of this: If Rawls's theory cannot

²⁸ M. Victoria Costa, 'Human Rights and the Global Original Position Argument in *The Law of Peoples*', *Journal of Social Philosophy* 36(1) (2005): 49-61; Peter Jones, 'International Human Rights: Philosophical or Political?', in Simon Caney, David George, Peter Jones, *National Rights, International Obligations* (Boulder, CO: Westview Press, 1996), pp. 183-204; Darrel Moellendorf, *Cosmopolitan Justice* (Boulder, CO: Westview Press, 1998), pp. 9-10. Kok-Chor Tan, *Toleration, Diversity, and Global Justice* (University Park, PA: Pennsylvania State University Press, 2000), pp. 19-45; Fernando Teson, *A Philosophy of International Law* (Boulder, CO: Westview Press, 1998), pp. 109-22. Some of these authors mention, without too much conviction, that Rawls could perhaps argue that the representatives of liberal peoples in the first international original position will subscribe to a principle endorsing human rights because they are already committed to human rights on the domestic level. However, first, the 'because' represents a *non-sequitur*; second, such an argument will cause Rawls other problems (for example: why do they then not endorse egalitarian distributive justice on the international level, if after all they are committed to it on the domestic one?); and third, and most important, the above argument has shown that liberal peoples – if that means, as it does in Rawls's construction, peoples that abide by the principles chosen in the first, domestic original position – are in fact *not* committed to human rights.

justify universal rights against slavery or genocide, then slavery and genocide are in principle allowed by his theory. What is not prohibited is permitted.

Why does the theory developed in *The Law of Peoples* imply a denial of human rights and a rejection of general injunctions against torture and genocide?

Rawls distinguishes between ‘five types of domestic societies’: ‘*reasonable liberal peoples*’, ‘*decent peoples*’, ‘*outlaw states*’, ‘*societies burdened by unfavorable conditions*’, and ‘*benevolent absolutisms*’.²⁹ The first two of these *societies* are also *peoples*. However, it is safe to say that the ‘outlaw state’ does not qualify as a people at all in Rawls’s theory. It is also not clear whether ‘burdened societies’ or ‘benevolent absolutisms’ count as peoples. In any case, certainly none of the three latter societies belong to what Rawls calls ‘well-ordered peoples’.³⁰ However, since in both international original positions only representatives of ‘well-ordered peoples’ take part, there is absolutely no discernible reason why those representatives should grant any right to entities that are not well-ordered peoples, like, for example, to persons or to any of the other types of society.

In fact, while in the domestic case Rawls misleadingly uses the term ‘persons’ in his two principles (although the parties in the original position themselves would rather use the term ‘citizens’), he uses the term ‘peoples’ with regard to the principles that would allegedly be accepted in the two international original positions. However, there are still some errors in his formulations. Parties in the original position would prefer to talk more precisely of well-ordered peoples. Second, consider his principle 4, which states: ‘Peoples are to observe a duty of non-intervention.’ But this principle would not be chosen in one of the international original positions. Rather, the principle chosen would look more like: ‘Well-ordered peoples are to observe a duty of non-intervention into the internal affairs of *other well-ordered peoples*.’³¹

But does such a wording not violate the formal generality condition? This is not quite clear from the little Rawls has to say on the topic. How exactly do you decide if the use of ‘well-ordered’ is in line with the expression of general properties or if it amounts to a rigged definite description? In any case, if need be, the well-ordered peoples in the original position can easily avoid this complication by using ‘societies’ in all their articles of international law and simply agreeing on a preliminary article stating that societies are bound by the other articles only with respect to societies that themselves abide by all of the articles completely. This article *is* general.

The reasons why well-ordered peoples would formulate the articles of international law using the term ‘well-ordered peoples’ in the way described above or, if need be, using the term ‘societies’, but only after having reassured themselves with the preliminary article, are parallel to the reasons given in the domestic case. I explained above that in the domestic original position

²⁹ John Rawls, *The Law of Peoples* (Cambridge, MA: Harvard University Press, 2001), p. 4.

³⁰ Ibid.

³¹ Again, if one tried to object here that this or similar principles are not on the list Rawls provides, the same arguments apply as in the domestic case. Besides, even if the parties in the original position were not free to choose other items than the ones on the list (Rawls gives the term ‘choice’ a new meaning!), they would still be free to *reject* the ones on the list. Precisely this is what they would do. There is absolutely no advantage for them in being bound by principles that might not be reciprocated by outlaw states rather than not being bound by principles at all. In addition, while the representatives of the peoples do not know certain specifics in the original position, they have important general knowledge. Thus they believe, with Rawls, in the democratic peace thesis and hence do not have to worry that not entering into a contract with the other liberal states might endanger them. In the light of this thesis they can also trust that in the real world they can enter into defensive alliances with other democratic states –and even with decent states, which, after all, are partly defined by their willingness to honour contracts and agreements. This then allows them to deal with the rest of the world entirely free from cumbersome obligations towards that rest.

the parties would assume that choosing a principle like ‘Do not enslave anybody’ protects them and outsiders from being enslaved by the parties to the contract, that is, by the insiders; but they would *not* assume that outsiders – who are, after all, not parties to the contract – will also abide by this rule. Therefore, such a general rule would give outsiders more protection than insiders; and this is why the parties to the original position prefer a rule like ‘Do not enslave members of your own society’, which gives them all the protection that the first principle affords, without giving outsiders any possibly unreciprocated protection. Similarly, the representatives of the well-ordered societies in the international original position assume that the other well-ordered societies will abide by the terms of the contract (this trustworthiness is actually a feature of their being well ordered³²), whatever these may be, but they will not assume the same about other societies (let alone about ‘outlaw states’). Thus, it would be irrational for them to prefer ‘Peoples are to observe a duty of non-intervention’ (unless tempered by the preliminary article) to ‘Well-ordered peoples are to observe a duty of non-intervention into the internal affairs of other well-ordered peoples’. Similar considerations hold for principles regarding genocide: principles that rule out genocide against peoples that are not well-ordered (and that are in violation of an article of international law) would not be accepted by well-ordered societies in the original position. Thus, in light of the implications of Rawls’s *Law of Peoples* slavery and genocide is permissible; and in light of his whole theory slavery will, as shown above, even be mandatory in certain not at all unusual circumstances. In this respect at least – as it concerns the denial of human rights, the permission of genocide and the endorsement of slavery – there is no conflict between *A Theory of Justice* and *The Law of Peoples*.

IV. Summary of the Argument

1. The two principles of justice as formulated by Rawls rule out the moral permissibility of genocide or enslavement of persons who are not members of one’s society. However, they also imply the extension of the difference principle to the global level. Thus they cannot be what Rawls actually intended. He probably intended them to refer only to the citizens of the society in question, without realizing that this would imply that the genocide or the enslavement of outsiders would be just.

2. The two principles of justice as formulated (but not as intended) by Rawls would actually *not* be chosen in the original position. Nor would any natural duties be chosen that give any rights to outsiders. The reasons are as follows:

2a. If the members in the original position think, as Rawls suggests, that their society is closed and they will have no interaction with outsiders, and if, furthermore, they are self-interested and concerned with the basic structure of their own society, than there is absolutely no reason for them to use the terms ‘persons’ or ‘least advantaged’ in the formulation of the two principles. Rather, they will use the terms ‘citizens of our society’ and ‘least advantaged of our society’ instead. But thus revised, the principles of justice imply that the genocide or the enslavement of outsiders can be just. In addition, self-interested individuals who assume that their society is closed have no reason to choose natural duties that imply any obligations towards outsiders.

2b. Even if the parties to the original position took into consideration that there might be some interaction with outsiders, they would still not subscribe to principles that rule out maltreatment of those outsiders. While the contracting parties assume, according to Rawls, that the insiders will abide by the terms of the contract, no such assumption is in place with regard to outsiders (who are, after all, not parties to the contract). It would be irrational, however, to bind oneself to duties others would not feel bound by. Thus, only those principles of justice and natural duties that permit the enslavement and genocide of outsiders would be accepted from the self-interested perspective of the original position.

³² Ibid., p. 25.

3. Attempts to block these conclusions do not work.

3a. The objection that Rawls's use of the term 'persons' instead of 'citizens' makes no difference in the situation Rawls asks us to envision, because the two terms are extensionally equivalent in that situation, fails. First, intensional non-equivalence would be quite sufficient for my argument. Second, the two terms are not extensionally equivalent: Rawls asks us to envision a society that is closed, not a world in which outside of that society no other persons exist.

3b. The objection that Rawls limits his theory of justice to the basic structure of a *closed* society that has no interaction with outsiders and that this theory is therefore not applicable to relations between insiders and outsiders does not work either. *First*, Rawls also tries to use the original position to justify natural duties. However, since according to Rawls constraining the means used in international wars is a way of recognizing these natural duties, these natural duties necessarily also refer to relations between insiders and outsiders as well as between outsiders and outsiders. In short, the premise upon which the 'limitation objection' rests is simply wrong. *Second*, even if that premise were correct, the 'therefore' would still be misplaced – that is, even if the theory were limited to societies that do not have any interaction with outsiders, it would still be wrong. To repeat the above analogy: A philosopher who develops principles of justice applicable to the character of only those judges that never come into contact with black people has obviously still made a rather grave mistake if he proposes principles implying that such a judge can be just even if he thinks (and is inclined to act accordingly) that there is absolutely nothing wrong with sending innocent black people to prison at will. To respond, 'But my principles are only to apply to judges who will never interact with black people' does not refute this argument; it completely misses the point because for the sake of argument it has already been accepted that the principles are limited in the described way. The limitation simply does not save the principles. *They are wrong even if applied only to such judges.* The fact that a certain judge will actually never harm anybody due to some lucky external circumstances does not make him just. In the same vein, a basic structure that allows the enslavement and all kinds of other abuse of outsiders is *not* a just basic structure, whether the society is closed – and thus does not get the chance to harm others – or not.

3c. Thus, that a society is closed does not undermine the argument expounded here. A basic structure, which includes the constitution and other laws of the society, simply cannot avoid implying something about the relation to outsiders – whether the society is closed or not. If, however, the two (revised) principles of justice applied *only to the interactions between the members of the society*, this indeed would mean that they would imply absolutely nothing about the enslavement of outsiders. Yet this desperate objection completely ignores Rawls's explicit insistence that in the first place he wants to find principles of justice *for the basic structure* of a closed society. He nowhere says or suggests or proceeds in a way that would suggest that he merely wants to find principles *for the interaction* between members of the society.

3d. The claim that the 'logical' way to apply 'Rawls's' theory to outsiders would be to have the outsiders already participate in the domestic original position can hardly be an objection to my criticism of Rawls's theory, but should rather be directed against Rawls himself. I am criticizing the theory Rawls did develop, not the one he 'should have' developed. If Rawls thought that the relations between insiders and outsiders should be dealt with by including the latter in the domestic original position, he could easily have said so instead of merely proposing a domestic original position exclusively inhabited by insiders and international original positions exclusively inhabited by 'peoples'.

4. In the two international original positions described in the *Law of Peoples* the parties would also not choose any principles that would outlaw enslaving or slaughtering peoples or the members of peoples that are not parties to one of the two international original positions. The reasons parallel those applied to the domestic original position.

5. If people try to escape my contention that Rawls's theory of justice implies the justifiability of slavery and genocide by continually insisting that Rawls's theory actually implies *nothing* about the treatment of outsiders, then their claim would still imply that from the perspective of

Rawls's *Theory of Justice* the genocide and enslavement of outsiders can neither be endorsed *nor criticized*. And since Rawls's *Law of Peoples* explicitly applies to the relation of states towards outsiders and between each other, and liberal and decent peoples in the two original positions would choose principles allowing the genocide and enslavement at least of persons who do not belong to liberal or decent peoples, Rawls's *overall* theory would still imply the justifiability of slavery and genocide.

Conclusion: Rawls theory of justice is a failure and completely unacceptable from a liberal point of view. The same, incidentally, is true of quasi-Rawlsian theories that propagate a global original position involving *all* persons on Earth. Only a theory that propagates a *universal* original position, involving all persons in the Universe, would escape the above objections. Such a theory, however, would not find Rawls's approval since it is literally *cosmopolitan*.³³

³³ For very useful comments on earlier drafts of this article I thank Allen Buchanan, Joseph Chan, Robert E. Goodin, Jan-Christoph Heilinger, Luciano Venezia, and Héctor Wittwer as well as the audience of a talk I gave at the Department of Philosophy, University of Hong Kong, in particular Jiwei Ci, Alexandra Cook, Max Deutsch, Chris Fraser, Chad Hansen, Patrick Hawley, William Haines, Kelly Inglis, Joe Lau, and Timothy O'Leary. **(These acknowledgements are missing in the published version due to a mistake of the publisher.)**