

EUROPEAN UNIVERSITY INSTITUTE, FLORENCE

DEPARTMENT OF LAW

EUI Working Paper LAW No. 2003/16

**“The Last Thing He Wanted”:
Realism and Utopia in *The Law of Peoples* by John Rawls**

WOJCIECH SADURSKI

BADIA FIESOLANA, SAN DOMENICO (FI)

All rights reserved.
No part of this paper may be reproduced in any form
without permission of the author(s)

© 2003 Wojciech Sadurski
Printed in Italy in December 2003
European University Institute
Badia Fiesolana
I – 50016 San Domenico (FI)
Italy

“The Last Thing He Wanted”:¹
Realism and Utopia in The Law of Peoples by John Rawls

Wojciech Sadurski*

There is a clear line that connects three main works by John Rawls: *A Theory of Justice* (henceforth referred to as *TJ*), *Political Liberalism* (*PL*) and *The Law of Peoples* (*LP*);² and one illuminating (even if somewhat perverse) way of tracing this line is through the responses of Rawls’s critics to each subsequent book. The admirers of the “original” Rawls, the Rawls of pre-*TJ* articles and the *TJ* itself, were rather surprised and disappointed by *PL* (and, earlier, by a series of articles that preceded *PL*), discerning in it a surprising “relativist” turn in Rawls’s thinking.³ The universalist aspiration of the liberal theory of *TJ*, with the benefits of liberal rights to be conferred on everyone, gave way to (what the critics saw as) a broad deference to non-liberal conceptions and ways of life, restricted only by a vague proviso that they all should be “reasonable”. This disenchantment was later echoed in the responses of many liberals – fans of the *TJ* Rawls – to *LP*, which extended the limits of toleration for non-liberal ideals and political structures prevailing in other societies.⁴ Just as *PL* tried to make sense of the fact that not all people within a society will be (or should be) liberals convinced of the rightness of the principles elaborated in *TJ*, and therefore that liberals should have a theory for coexisting with these others in a stable and morally justified way that would go beyond a mere *modus vivendi*, so *LP* tried to take seriously the fact that not all societies will look like a society governed by the *PL*-based rules

¹ With apologies to Joan Didion, *The Last Thing He Wanted* (Alfred A. Knopf, New York 1996).

* Professor in the Department of Law, European University Institute in Florence, and in the Faculty of Law, University of Sydney.

² John Rawls, *A Theory of Justice* (Harvard University Press, Cambridge Mass, 1971); *Political Liberalism* (Columbia University Press, New York 1993); *The Law of Peoples* (Harvard University Press, Cambridge Mass. 1999).

³ The critiques of *Political Liberalism* are legion. My favorite critique of political liberalism (not just Rawls’s, hence the lower case) from the perspective of integral liberalism is Stephen Gardbaum, “Liberalism, Autonomy, and Moral Conflict”, *Stanford Law Review* 48 (1996): 385-417. For an interesting recent reflection on Rawls’s critics, see Anthony Simon Laden, “The House that Jack Built: Thirty Years of Reading Rawls”, *Ethics* 113 (2003): 367-90, especially at 373-79.

⁴ For two representative examples of such criticism, see Fernando Téson, *A Philosophy of International Law* (Westview: Boulder 1998), chap. 4; Kok-Chor Tan, “Liberal Toleration in Rawls’s Law of Peoples”, *Ethics* 108 (1998): 276-95. Note that Téson’s and Tan’s critiques were directly addressed against an earlier statement of Rawls’s theory, in an essay he had published in 1993; John Rawls, “The Law of Peoples”, in Stephen Shute & Susan Hurley, eds, *On Human Rights: The Oxford Amnesty Lectures* (Basic Books: New York 1993): 41-82 Other criticisms along these lines will be referred to throughout this article.

that prescribe liberal-democratic institutions for the coexistence of liberal and non-liberal individuals. But each new step – the one taken from *TJ* to *PL* and the one from *PL* to *LP* – was fraught with dangers to the essence of liberalism itself, and was received by Rawls’s aficionados with a degree of consternation. Think of a priest who just managed to convert you to a new faith, and who gradually retreats from his beliefs: that is how many “integral liberals” reacted to Rawls’s developments.

Now that, alas, Rawls’s oeuvre is a closed book, we can take stock of this development and raise the question of whether, in taking the path from *TJ* to *PL* and then on to *LP*, Rawls has successfully moved the limits of liberalism beyond their conventionally recognised point, or rather dangerously exceeded them. And I am deliberately using the concept of “limits of liberalism” here because Rawls refers to it explicitly at least once in *LP*: he states that “enlightenment about *the limits of liberalism* recommends trying to conceive a reasonably just Law of Peoples that liberal and nonliberal peoples could together endorse” (78, emphasis added).⁵ I will argue that Rawls’s last book – his last statement on liberalism – fails to convince us as to where the proper limits of liberalism should lie.

1.

TJ was a subtle, elegant and sophisticated exposition of liberal principles that, taken together, attempt to reconcile two fundamental concerns: the first for the maximization of individual liberty, constrained only for the purposes of ensuring the equivalent liberty of others, and the second for solidarity-based distributive justice aimed at the advancement of the worst-off, especially if they are suffering for reasons over which they had no control. The great appeal of Rawls’s book consisted in a masterly combination of some classical (one might even have thought obsolete) instruments of political philosophy, in particular the idea of a hypothetical social contract, within a modern, twentieth-century analytical framework, borrowing from such fields as political economy, social choice theory, etc. In combining these classical and contemporary conceptual instruments, Rawls opened up for liberal political philosophy a third way that could transcend the rather unpalatable alternatives that faced political philosophy in the mid-1950s (when the first pre-*TJ* articles began appearing): either a sterile conceptual analysis of the language used to describe political ideals, or thoroughly ideological expressions of faith unsupported by any coherent analysis. Rawls showed how a normative political theory can be made analytically sound and coherent, and how political philosophy can be aligned with clear, unambiguous and ultimately controversial values espoused by a theorist.

The “disenchantment” felt by many readers of the second major work by Rawls – *Political Liberalism* – was based not so much on the perception of a change of

⁵ All the numbers in brackets in the main text refer to the page numbers of *The Law of Peoples*, supra note 2.

general approach to political philosophy as revealed by *TJ*, but rather on the substance of the “new” theory (first rehearsed in a series of articles published after *TJ* that then culminated in the book-length *PL*). *TJ* had an unambiguously universalistic message built into it: all humans have certain rights that can be derived from the two principles of justice articulated in the book. They are all, regardless of their place in any given society or in the world in general, entitled to the most extensive liberty compatible with the same degree of liberty for other people,⁶ and also to the distribution of material benefits in accordance with the difference principle (which postulates that socio-economic inequalities are only justified if they are to the benefit of the least advantaged) and fair equality of opportunity (understood not in a meritocratic fashion, but in a way that allows, indeed necessitates, affirmative measures by states aimed at the equalisation of access to scarce and sought-after social positions). These benefits were not contextualised or made contingent upon any particular circumstances or moral conceptions of the individuals concerned.⁷ *PL* seemed to put this universalistic ethic in question: individual liberty as the ultimate ideal (understood in a robust way, demanding that it be sufficiently effective) gave way to the ideal of tolerance for diverse moral conceptions, as long as they are “reasonable” (including the non-liberal ones), as the overarching ideal. The missionary zeal written into the very fabric of the liberal universal of *TJ* seemed to have been replaced by a concern for those who reject liberalism itself; hence the deep disenchantment of many liberal readers of *PL*.

One way of defending *PL* available to liberals was to argue that it was not the answer but rather the question that changed in the transition from *TJ* to *PL*. Rather than undermining his own “ideal” theory of liberal justice (and, as will become clear, the word “ideal” is, in this context, a term of art), Rawls removed one important assumption behind the argument in *TJ*: that everyone in a *real* society would in the end become persuaded to the principles reached by the rational members of the original position. *TJ* was about a society designed by liberals only – but such societies do not, will not and (perhaps) should not exist: a

⁶ For a cosmopolitan reading of Rawls’s first principle of justice in *TJ*, see Thomas W. Pogge, “Cosmopolitanism and Sovereignty”, *Ethics* 103 (1992): 48-75 at 55 n. 12.

⁷ There is a minor and highly qualified exception to this universalism in *TJ*. Rawls admits that there may be societies in which “social conditions” do not allow the establishment of rights conforming to the two principles of justice combined with the priority rule giving precedence to the principle of equal maximum liberty; in such circumstances, a “general” conception of justice should rule, which allows trade-offs between increases in (and equalisation of) liberty and in economic gains, *TJ* at 303. But these are, according to Rawls, only temporary circumstances and they tend to diminish in time: “The lexical ordering of the two principles is the long-run tendency of the general conception of justice consistently pursued under reasonably favorable conditions. Eventually there comes a time in the history of a well-ordered society beyond which the special form of the two principles takes over and holds from then on”, *TJ* at 542.

fundamental moral disagreement is a chronic, stable and irreducible property of real societies. This is what the “political” in “political liberalism” really means: it is about the political principles governing the coexistence of liberals and non-liberals in real societies – a coexistence that is in itself morally meaningful rather than being only a grudgingly adopted *modus vivendi*. *PL* was therefore – the admirers of *TJ* could console themselves – not a different theory, but instead resulted from a different optic adopted by the liberal philosopher: while *TJ* was an account of first-order liberalism (addressing the question of a good liberal society, as designed by liberals), *PL* was an articulation of a second-order liberalism (addressing the question of a society in which liberals and non-liberals could find decent room for cooperation and mutual respect).

The Law of Peoples can be seen as a continuation along the same path: as the suspension of yet another idealising assumption; this time, the assumption that all that matters is a single society, or that a liberal-democratic society has a planetary range. This assumption is self-evidently unrealistic, and if a liberal philosopher has the burden of explaining how liberal principles will be implemented in a pluralistic, non-homogeneous society composed of liberals and non-liberals alike, then he or she also has the burden of showing how these principles can be implemented in a pluralistic, non-homogeneous *world*, composed of liberal and non-liberal states. This subsequent change of optic in Rawls’s work can be as recognition of a factor initially left deliberately outside the concerns of the theory: the existence of the borders dividing different states, nations and peoples in the contemporary world. But the way in which this factor was introduced into the theory caused deep anguish to many liberal fans of the “original” Rawls, of the *TJ* era. The universalist theme of *TJ*, already qualified and diluted by *PL*, seemed to be further weakened by the arguments in *LP*. A liberal, initially convinced by *TJ*, must now abandon all of her missionary zeal and dispense with the universalistic ideal of certain robust liberal rights that should be conferred upon everyone. A deep set of individual benefits deducible from the hypothetical social contract has been watered down; not just because of the fact of moral pluralism – as in *PL* – but also due to factors usually regarded as much more arbitrary and contingent from a moral point of view, such as the geographical location at which a person happened to be born. The significance of this anti-universalistic turn, undertaken already in *PL* but completed in *LP*, can account for the strength of the liberal disenchantment with Rawls’s last book.

2.

In *LP* Rawls puts forward eight principles of the Law of Peoples that, taken together, constitute the principles of international justice. They are meant to be, on the one hand, a reflection of “familiar and traditional principles of justice among free and democratic peoples” (37, footnote omitted), which would suggest that they are merely a compilation of the actually existing rules of proper international behaviour; but, on the other hand, Rawls precedes their recital with

the declaration that, in announcing them, he “[p]roceed[s] in a way analogous to the procedure in *A Theory of Justice*” (36, footnote omitted) and further explains, in a footnote attached to this last sentence, that the relevant passages in *TJ* include, *inter alia*, “chapter 3 [which] gives the reasoning from the original position concerning the selection of principles” (36 n. 41). Hence, right at the outset, we should take note of a certain dichotomy between the positive and the critical status of the eight principles: if they are “familiar and traditional” then they are more akin to a compilation of international custom; if they are derived from the hypothetical social contract *à la TJ*, they are more like normative standards in terms of which international custom should be judged and criticised. As we will see, this tension permeates much of Rawls’s book.

The eight principles of the Law of Peoples are:

1. “Peoples are free and independent, and their freedom and independence are to be respected by other peoples.
2. Peoples are to observe treaties and undertakings.
3. Peoples are equal and are parties to the agreements that bind them.
4. Peoples are to observe a duty of non-intervention.
5. Peoples have the right of self-defense but no right to instigate war for reasons other than self-defense.
6. Peoples are to honor human rights.
7. Peoples are to observe certain specified restrictions in the conduct of war.
8. Peoples have a duty to assist other peoples living under unfavorable conditions that prevent their having a just or decent political and social regime”. (37, footnote omitted).

Even a very cursory reflection upon these eight principles indicates why so many liberals could view a “Law of Peoples” constituted in such a manner as almost a betrayal of the liberal creed.⁸ Not only the very impressive, thick individual entitlements corresponding to the two principles of justice articulated in *TJ*, but also most political rights to democratic and liberal procedures (in particular, the freedom of speech, of association etc) that stem from *PL*, are outside the scope of the Law of Peoples. They are certainly not captured by the sixth principle: the respect for human rights. Rawls makes it abundantly clear that the sixth principle covers only very thin, rudimentary rights, which are a pale shadow of the liberal-democratic rights, defended in his two earlier books. This clarification comes in two ways. Firstly, when Rawls announces in general that the sixth principle rights are the right to life, to liberty (but narrowly understood; for instance, liberty of

⁸ See, e.g., John Tasioulas, “From Utopia to Kazanistan: John Rawls and the Law of Peoples”, *Oxford Journal of Legal Studies* 22 (2002): 367-96 at 380-90.

conscience does not necessarily presuppose “equal liberty”, to property and to formal equality understood as treating equal cases equally (65). Secondly, and more tellingly, the thinness of the sixth principle rights is made evident in the description of “decent hierarchical societies”, which fall far short of liberal-democratic standards and yet which pass the muster of the Law of Peoples, and thus meet the criteria of the sixth principle. This thinness is further underlined by his statement that “Human rights set a necessary, though not a sufficient, standard for the decency of domestic political and social institutions” (80). If these *extra* factors that stand between the necessary and the sufficient in the set of criteria for “decent” societies are non-human rights related, and if, in addition, “decent” hierarchical societies fall far short of the standards of liberal democracy, then it is obvious that what is to count as “human rights” in terms of the sixth principle is very rudimentary indeed.

As a result, an individual belonging to a decent but hierarchical society will find precious little grounds in the Law of Peoples to claim his or her rights: if the ethical criteria with which the authorities must comply are exhausted by the eight principles, then even very inegalitarian societies, for instance those denying equal educational and employment opportunities for women, will pass the test.

When it comes to distributive justice, *LP* is even thinner than with respect to liberal-democratic rights: Rawls emphatically declined the invitation by many of his commentators to extend the two principles of justice to the world stage, and to construct a model of global justice that would be an extrapolation of the principles of domestic justice as articulated in *TJ*.⁹ His early followers attempted to describe themselves the international distribution of goods demanded by the global original position: what the world would look like if the idealising Rawls’s idealising assumption limiting himself just to one society were lifted, but the deep argumentative structure of the theory remained unchanged? Small wonder that they came up with quite radical proposals for the transformation of the international redistribution of goods, as a result of the extension of the second *TJ* principle, and for a robust international regime of human rights protection, as a result of the extension of the first *TJ* principle. Both of these proposals are resoundingly rejected by Rawls in *LP*.

In reviewing the reasons given for such a rejection, one must be careful not to ask questions that are out of place in Rawls’s methodology. One should therefore resist the temptation to ask the simple-minded question: (1) “Are these eight principles, taken together, constitutive of a morally satisfactory and sufficient basis for a just international system?” Rather, one should insist on the question:

⁹ See, e.g., Thomas Pogge, “An Egalitarian Law of Peoples”, *Philosophy & Public Affairs* 23 (1994): 195-224; Christian Beitz, *Political Theory and International Relations* (Princeton University Press 1979); Charles R. Beitz, “Rawls’s Law of Peoples”, *Ethics* 110 (2000): 669-96.

(2) “Is it plausible that, in the original position, parties to the hypothetical social contract reasoning rationally in the specified circumstances of relative ignorance would have adopted these eight principles as a satisfactory and sufficient set of rules governing the international justice?” Question (2) is, of course, not fully reducible to question (1); otherwise, the social contract methodology would be fully redundant, as this is precisely what stands between questions (1) and (2). And perhaps it *is* redundant, in *LP*, but this must be shown rather than merely asserted at the outset. This indicates how important it is to explore Rawls’s social contract methodology in *LP*: obviously, the design of the conditions of the contract is crucial for its conclusions, and any modification of the conditions will produce a change in the conclusions. In moving from *TJ* to *LP*, Rawls effected some very significant changes in the description of the conditions of social contract, and these changes are responsible for the startling (to the fans of *TJ*) thinness of the Law of Peoples as a *critical* theory of international justice. The main change has been in the composition of the “participants” in the deliberative process in the original position: those operating in *TJ* as spokespersons for the *individuals* representative of different social groups and strata have been replaced in *LP* by “peoples” as the relevant parties to the social contract.

Before inquiring into the justification for such a change it is important to appreciate the momentous importance of its implications. If individuals were to be represented in the original social contract, the confining of the rules of international justice to the eight principles cited above would be indefensible. Under the method of social contract behind the veil of ignorance, principles of international justice would have to consist of the rules that would have been plausibly agreed to by the representative individuals, each unaware of the specific society to which they belong. *This* type of ignorance would correspond best to the *TJ*-style veil of ignorance, i.e. the ignorance as to (among other things) one’s place in a real society in which one will find oneself once the social contract has been concluded. And if the parties *qua* individuals, unaware as to their geographical location in the world, were to write the rules of international justice, they would surely go for something much thicker than the minimalist conception described in *LP*. The risk of harm that would be incurred upon finding oneself in a deeply inegalitarian and undemocratic society (even although still “decent” by *LP* standards) would be too catastrophic for the parties to agree to the principles of *LP*. They would surely rather opt for much more rigorous and demanding requirements as to the internal social institutions of the various “peoples” making up the world population, and for much stronger rules for international economic redistribution.¹⁰ If, in contrast, it is the “peoples” who are represented in the original position, understood as unitary corporate bodies, the Law of Peoples as a fair system of international justice becomes much more convincing. That is why

¹⁰ Similarly Thomas W. Pogge, “John Rawls on International Justice”, lecture delivered at the European University Institute, Florence, on 18 March 2002.

it is crucial to explore the reasons for the replacement of individuals by “peoples” in the hypothetical social contract of *LP*.

Moreover, a major change in the transition from *TJ* to *LP* in the design of the conditions of the social contract can be found in the selectiveness of the composition of the parties in the latter. In *TJ*, no representative position in a society, however defined, was *a priori* excluded from a role in formulating the hypothetical social contract. In *LP*, in contrast, only some peoples are included in the contract-making: namely the “well-ordered” peoples (the category that, in Rawls’s taxonomy, occupies the two highest positions in the classifications of peoples on the planet: the liberal and “decent” (although non-liberal) peoples). All of the remaining peoples (which Rawls divides into “outlaw states”, “societies burdened by unfavorable conditions” and “benevolent absolutisms”, 62-63) remain beyond the pale, and they do not participate in the contract. This, of course, is a consequence of the “peoples-based” rather than the “individuals-based” contract: if the rules of international justice were to be contracted by the representatives of individuals, surely the unfortunate inhabitants of the last three categories of societies would be in greatest need of the benefits deriving from the critical potential of international justice.

All this shows, once again, how significant the peoples-based construction is, and that reflection on the arguments for such a construction (in contrast to the individuals-based one) is the main challenge for a reader and a critic of Rawls. I will take up this challenge here, but first we must look more closely at the characteristics of the “decent societies” that are to be included in the international community to be governed by the Law of Peoples.

3.

The distinction between liberal peoples and that other category of “well-ordered” societies that are also “worthy of membership in a Society of Peoples” is of central importance to Rawls’s book and its reception. What worried many of Rawls’s critics most was his excessive tolerance (*un mot juste*, as will become clear later) for societies that are governed in a way that would not be tolerated by Rawls and his liberal followers in their own countries.¹¹ To be characterised as *liberal*, a state must meet certain relatively stringent requirements for liberal-democratic institutions, including those described in Part III of *PL*. This much is obvious. The problem begins with the characterisation of those societies that are not liberal, and yet deserve to be recognised as “equal participating members in good standing of the Society of Peoples” (59). The main aim of such “recognition” is the ascertainment of “how far liberal peoples are to tolerate nonliberal peoples” (59) – and “toleration”, in this context, goes far beyond simple non-intervention: it includes, on the part of the liberal peoples, “a duty of

¹¹ E.g., Pogge, id.

civility requiring that they offer other peoples public reasons appropriate to the Society of Peoples for their actions” (59). The “good standing” of the nonliberal “decent” peoples therefore goes a long way in earning them particular entitlements; thus, the relevant features in deciding which nonliberal peoples are to be elevated to this position are significant.

Without going into the details of Rawls’s own description of decent nonliberal peoples, there are three main ways in which those peoples depart from the standards of liberal peoples.¹² Firstly, those societies incorporate certain “basic inequalities among their members” (70) to the extent to which these inequalities would clearly be considered unjust in liberal societies. We are never told how far inequality is still compatible with the “decency” of a hierarchical society: we have some negative criteria that disqualify a society from being “decent” (obviously, a slave society cannot be so considered, 65) but we are not given any more specific positive criteria. Secondly, there is *some* liberty of conscience but not “*equal* liberty of conscience” (70, emphasis added), and this may include, for instance, an established, state religion, as long as other religions are not persecuted (76). Lastly, these peoples lack what we would consider normal political, representative democratic institutions, but instead have a “decent consultation hierarchy” (71). As one can see, if societies marked by these three defects (“defects”, that is, from a liberal point of view) can *still* be held to comply with the “human rights” condition contained within the Law of Peoples, human rights must be interpreted in a very narrow and feeble way indeed.

Embracing societies tainted by such important defects (from a liberal-democratic perspective) into the community of peoples raises, for a liberal, very serious dilemmas. If we refuse, in our own societies, respect to those moral-political conceptions that would postulate redesigning our societies along these illiberal lines, why should we respect such societies abroad? If something is reprehensible within our own society, why should it gain respectability if it occurs outside our borders? In this question, naturally, the issue of the moral status of “peoples” as the units of international justice re-emerges. The moral costs, for a liberal, of deference to the illiberal “peoples” are obvious as, in the process, the concerns of the individuals who make up these “peoples” are sacrificed; and this is a sacrifice that liberals abhor. A liberal who accepts the position of normative individualism assesses the value of social institutions, rules and practices though the prism of their effect upon the legitimate interests of the individuals affected, and, in particular, those interests linked to their liberty and equality. Within a Rawlsian conception of individual justice, the losses in terms of those interests are (from a liberal perspective) quite obvious, while the advantages are not.

One way of ascertaining those losses is to consider Rawls’s own description of an imaginary “decent hierarchical people”, dubbed “Kazanistan” (75-78). There is

¹² In this brief summary, I follow a useful observation by Thomas Pogge, id.

no separation of church and state, with the establishment of a favoured religion and the understanding that only the adherents to that religion “can hold the upper positions of authority and influence the government’s main decisions” (75); on the other hand, other religions “are tolerated” (76); their members are not persecuted, nor are they “subjected to arbitrary discrimination” (76). No democratic institutions exist in Kazanistan, but rather “a decent consultation strategy” (77) based on collectivist grounds: that is, “all groups must be consulted” while “each member of a people must belong to a group” (77). A degree of dissent is “respected”, but only in the sense that “a reply is due [from the government to the dissenters] that spells out how the government thinks it can both reasonably interpret its policies in line with its common good idea of justice and impose duties and obligations on all members of society” (78).

It would be admittedly an easy exercise to pinpoint various weaknesses, defects and shortcomings of “Kazanistan”. Easy, but not particularly effective against Rawls’s theory because it is not his aim to praise a social ideal embodied by his description; he himself admits: “I do not hold that Kazanistan is perfectly just...” (78). The purpose of the account of “Kazanistan” is not an articulation of an attractive vision of a good society that would rival, or be equal to, Rawls’s own ideals articulated in *TJ* and *PL*. Rather, the purpose is to identify the outer parameters with which societies other than our (admittedly, imperfect) societies must comply in order to be respected as members of the society of peoples, and thus merit all of the rights and duties derived from the Law of Peoples. This immediately puts us on notice that, owing to its minimalistic aims, the critical potential of the theory is severely reduced. As Rawls observes: “it seems to me that something like Kazanistan is the best we can realistically . . . hope for”, and he adds an extremely important observation: “I think enlightenment about the limits of liberalism recommends trying to conceive a reasonably just Law of Peoples that liberal and nonliberal people could together endorse” (78). The importance of this remark can hardly be overstated: considering the role of Rawls in the contemporary restatement of liberal philosophy, here we have an authoritative declaration concerning the outer “limits of liberalism”, and a look at the world “outside” those limits. Can we really accept this declaration without a sense of intolerable damage done to the value-laden presuppositions that triggered our endorsement of the liberal principles in the first place? Damage done, for example, to the presupposition articulated simply by Brian Barry in his recent book, that “[n]obody, anywhere in the world, should be denied liberal protections against injustice and oppression”?¹³

Posing this question is all the more important as the conception of the Law of Peoples should not be seen as aimed at formulating a set of policy guidelines for liberal-democratic governments about how to deal with various non-liberal states

¹³ Brian Barry, *Culture and Equality* (Polity: Cambridge, 2001) at 138.

and governments in the world. It does this, but only indirectly and secondarily. It is important to reflect upon the general purposes of a theory such as the one contained in *LP*. It would be an error to read it as merely a theoretically elaborated statement of *Realpolitik* with the set of prescriptions regarding how liberal governments should treat different categories of states in their foreign policy. Such an approach to the statement of the theory's aims should be resisted. Rather, the theory has philosophical purposes; that is, it aims at articulating the principles that we, as individual citizens in liberal-democratic states, should endorse when reflecting upon our attitude towards illiberal states and that, consequently, should inform our conduct *vis-à-vis* those states.¹⁴ However, governmental conduct is only a part of the picture that we, as individuals, must construct in our own minds; at an earlier, more fundamental level, we must ascertain our own position about how to relate to the fact that there exist other states that are organised in a way that departs drastically from our own ideals. Whether a prescription of governmental intervention follows or not is something secondary and under-determined by this more primary question of our own attitude towards those other states and our assessment of an international system that embraces different types of states.¹⁵ A simple point that I want to emphasise at this stage is that a theory such as Rawls's in *LP* should be seen as critical and normative rather than pragmatic and prudential: it contains a certain ideal of an international system rather than a set of *Realpolitik* prescriptions, and should be judged itself under these standards.

It seems that, at times at least, Rawls makes it clear that this is precisely how he sees the purpose of his theory. Right at the outset, indeed in the first sentence of the book, he declares that by Law of Peoples he understands “a particular political conception of right and justice that applies to the principles and norms of international law and practice” (3); he also announces further that the Law of Peoples is “an extension of a liberal conception of justice for a domestic regime to a Society of Peoples” (9). The equivalence of the *LP* conception to that of *TJ* is therefore expressly stated – and so it should be seen, just as the former book, as having a critical, normative function. From this point of view, it is important to try to ascertain whether the critical-normative function of the theory extends to the domestic arrangements of the non-liberal peoples. A negative answer would push the theory much closer towards the *Realpolitik* end of the spectrum: it would basically mean that, once we are satisfied that a non-liberal society can be

¹⁴ A similar point is made, in a somewhat different way, by Allen Buchanan, “Rawls's Law of Peoples: Rules for a Vanished Westphalian World”, *Ethics* 110 (2000): 697-721 at 713.

¹⁵ For an argument, made in the context of a critique of *LP*, that there is a gap between what is morally justifiable and what can be morally implemented on international scale, see Martha Nussbaum, “Women and the Law of Peoples”, *Politics, Philosophy & Economics* 1 (2002): 283-307 at 297-302. See also Tan *supra* note 4 at 285 and Beitz, “Rawls's Law of Peoples”, *supra* note 9 at 687.

properly characterised as “decent” (rather than belonging to the other, more troublesome categories of illiberal peoples), all that we (as individual citizens of liberal societies) can do is deal with them within the Law of Peoples, without trying to affect their institutions and practices or, perhaps, even justifiably criticise them.¹⁶ A certification of a society as “Kazanistan”-like would, under this interpretation, leave an inter-governmental track consistent with the eight principles of the Law of Peoples as the only morally justified way of dealing with that society.

What makes such an understanding of the purpose of the theory of international justice (and I will consider in a moment whether it is the only acceptable interpretation of Rawls’s theory) particularly startling is that it is presented as part of an *ideal* theory. One could see the reasons why, as part of a *non-ideal* theory we could accept that it is unjustified to try to affect the evolution of the illiberal (even if decent) societies in a liberal direction. After all, an ideal/non-ideal theory distinction is explicit in the *LP* and, just as in *TJ*, the difference between both is based fundamentally on partial compliance: a non-ideal theory provides the rules of conduct for a world in which the ideal theory is only partially respected. Now it would be quite palatable to many liberals to have a “non-ideal” theory of international relations in which the norms, *justified under an ideal theory*, for influencing the illiberal societies to change would be renounced. This could be for all sorts of reasons related to the fact of partial compliance, but we need not enter here into this hypothetical argument because, most emphatically, the admission of nonliberal though decent peoples into the Society of Peoples (with all its moral implications, including non-intervention) forms part of the *ideal* theory of *LP*. Thus, it is not seen as a regrettable concession to imperfect reality, but rather as part of the moral ideal itself: the existence of liberal democratic and non-liberal and non-democratic (though “decent” and “well-ordered”) peoples is viewed as part of the morally satisfactory *status quo*. We need to search, within the resources provided by Rawls’s theory, for some good reasons that justify such a position.

4.

One trope that may help clarify Rawls’s reasons for embracing non-liberal peoples in such a way is by considering his self-characterisation of his own theory as a “realistic utopia” – his favourite and frequently repeated¹⁷ formula of

¹⁶ For the latter point, see Tasioulas *supra* note 8 at 387-88; see similarly Fernando Teson, “The Rawlsian Theory of International Law”, *Ethics and International Affairs* 9 (1995): 75-99 at 88-89 (note that Teson’s criticism is addressed against Rawls’s essay, *supra* note 4, and not the book) I am not entirely convinced that Rawls’s precept of non-interference imposes upon us, citizens of liberal states, a duty to abstain from criticism of nonliberal states, but I do not propose to pursue this point here.

¹⁷ The Index to *LP* lists at least ten places in the book where the notion of “realistic utopia” to describe Rawls’s conception appears, *LP* 194.

the conception of the Law of Peoples. The first extended introduction of the notion of “realistic utopia” already indicates the nature of the problem: “political philosophy is realistically utopian when it extends what are ordinarily thought to be the limits of practicable political possibility and, in so doing, reconciles us to our political and social condition” (11). There is an inevitable tension between the aspiration of “extending the limits of political possibility” and “reconciling us to our conditions”, or between “the realistic” and “the utopian” in the “realistic utopia”. This need not be a contradiction in terms, of course: it is a normal consequence of the natural aspiration of a political philosophy to be critical towards reality while at the same time observing the outer parameters of reality that cannot, for all we know here and now, be transcended. But in *LP* the tension is the source of a real problem, and it is hard to avoid the observation that Rawls risks coming too close to favouring “the realistic” in the “realistic utopia” for the conception to perform a critical function towards the prevailing political reality.

This danger is clear when we consider the moral *status* of the inclusion of non-liberal (though “decent”) peoples into the Society of Peoples: there is certain ambiguity about it in Rawls that suggests that the tension in the “realistic utopia” is severe indeed. On the one hand, there are some hints in the book that suggest that the presence of those societies is not something morally positive *per se* but rather that they are embraced grudgingly, as a concession to political reality, in the manner of non-ideal theory, and that they need not be seen as something stable and persistent in the Society of Peoples. This interpretation is supported by those remarks that indicate that, for Rawls himself, liberal-democratic societies are superior to their non-liberal counterparts (for example, when he states: “I am not saying that a decent hierarchical society is as reasonable and just as a liberal society”, 83), and that he identifies *moral* defects in non-liberal decent peoples (“A decent hierarchical society . . . does not treat its own members reasonably or justly as free and equal citizens, since it lacks the liberal idea of citizenship”, 83). Even more clearly, the possibility of liberal universalism comes through a startling (considering the overall message of the book) suggestion that one of the *functions* of liberal tolerance towards non-liberal decent peoples may be its positive impact upon those societies in terms of encouraging their development towards liberalisation and democratisation: “With confidence in the ideals of constitutional liberal democratic thought, [the Law of Peoples] respects decent peoples by allowing them to find their own way to honor *those ideals*” (122, emphasis added). So there is, after all, a legitimate hope inherent in the Law of Peoples, to the effect that the decent nonliberal peoples will find the light at the end of the tunnel; that they will embrace “the ideals of constitutional liberal democratic thought” in their political practice. As those ideals are certainly not implemented in the political life of Kazanistan and the similar peoples, the strategic function of the Law of Peoples (with its principles of non-intervention, etc.) is viewed as helping Kazanistan *et al.* to become liberal and democratic.

Alas, there is much in the book that supports an opposite interpretation of the status of the inclusion of various “Kazanistans” into the Law of Peoples: the interpretation that supports the idea that the presence of liberal *and* non-liberal peoples is morally significant, stable, accepted in a non-grudging fashion, and need not be overcome. At the very end of the book, Rawls observes: “not all peoples can reasonably be required to be liberal” (122), and there is no sense of regret in this statement and in its context. The strongest indication that the inclusion of nonliberal decent societies is itself of moral significance is provided by the (already noted) recognition of the existence of liberal and nonliberal peoples alike in the Society of Peoples *as part of the ideal theory* of international justice. However, the moral reasons for this significance are unclear – particularly if one considers the observations, quoted above, concerning the moral defects of hierarchical societies, which imply that liberal-democratic norms constitute a yardstick for the moral evaluation of hierarchical, decent societies.

Incidentally, by acknowledging this moral inferiority of hierarchical societies and the elevated status of liberal-democratic norms, Rawls surrenders the possibility of defending the moral status of decent hierarchical societies on the basis of any *individual* interests of their members that those societies may satisfy better than liberal-democratic institutions would. One could imagine such a defence along the well-known lines of showing that, due to various cultural and historical specificities of some peoples, their members yearn for social harmony, non-antagonistic political arrangements, subjection to benevolent but powerful authorities etc., and that they abhor the cacophony of conflicting voices that a democratically organised society necessarily produces, along with the chaos, conflict and primacy of the individual over a collective that it entails. (No need to add that such conceptions must not be taken at face value and that, more often than not, they constitute disingenuous rationalisations for despotic rules rather than represent the genuine preferences of citizens of authoritarian states as we know them today – a point to which I will return by the end of this Working Paper). But theoretically one could follow this path to try to defend the moral standing of decent hierarchical societies by connecting them with the genuine individual interests of their members. That the author of *TJ* and *PL* does not follow this path is obvious, and not surprising. But if he *had*, the main alteration of the design of the social contract in *LP* would be pointless: there would be no reason to model it in terms of representatives of *peoples* rather than individuals. A global contract composed of representatives of individuals might well be designed in such a way that the individuals, placed in the position of uncertainty about which people they will belong to, would opt for a design that would be more sympathetic towards non-democratic social design, precisely in order to promote those individual interests that, presumably, are badly catered for by liberal-democratic societies. But, again, Rawls emphatically does *not* follow this path, and there is no reason for us to contemplate any further this possible strategy of argument.

5.

As I have already noted, in the substantive content of the Law of Peoples, Rawls' "realistic utopia" comes much closer to the "realistic" rather than the "utopian" pole. The same observation can be made about the *method* applied by Rawls. Despite an ostensible adoption of the social-contract method, which had been used with such originality and invention in *TJ*, in *LP* Rawls devotes surprisingly little attention to the derivation of the substance of the deal from the contractarian procedure. To be sure, we have a general description of the parameters of the contract: we know, for example, what the "veil of ignorance" involves in this case, i.e. what "the peoples" represented in the original position know (for instance, whether they are liberal or hierarchical) and what they do not know (for instance, what is their level of economic development and what their natural resources are). We also have the list of participants. What we do *not* know is the nature of the precise chain of reasoning that would lead from those conditions of the original position to the substance of the eight principles of the Law of Peoples.

All in all, the social contract – as a model of representation of moral reasoning about the principles of justice – takes, in *LP*, a very pale form. "Reflective equilibrium", which played such an important role in shaping the conditions of the original position in *TJ*, barely exists here. It appears only twice: once explicitly but with a marginal role in the reasoning process, in a footnote (86 n. 32), and once implicitly, in a way recognisable to Rawls aficionados but with hardly any operative function attached to it: "The social contract conception of [the Law of Peoples], more than any other conception known to us, should tie together, into one coherent view, our considered political convictions and political (moral) judgments at all levels of generality" (58). This last sentence appears almost at the end of the chapter and is not followed through upon. As a result, it remains an assertion but not an argument: we are not told how the reconciliation between the general principles and the specific intuitive convictions (which would serve as "provisional fixed points which we presume any conception of justice must fit")¹⁸ is used, *à la TJ*, to settle "the most favored description"¹⁹ of the original position. In *TJ*, "reflective equilibrium" was an effective and helpful device precisely because it suggested a procedure for bringing our general moral principles (about which we feel reasonably confident but still are prepared to modify them, within limits) and our more specific judgements of justice (again, considered as *provisional* fixed points of our moral argument) into line with each other; we are ready to "work from both ends" up to the point at which the principles will match the convictions. But if Rawls ended there, it would be just another form of a coherentist model of moral argument: the

¹⁸ *TJ* at 20

¹⁹ *Id.*

forcefulness of Rawls's method was that this "reflective equilibrium" was just the *beginning* of a moral argument: it served to describe the original position that "yields principles which match our considered judgments duly pruned and adjusted".²⁰ So, even if we reached the point at which "at last our principles and judgments coincide",²¹ this only paves the way to the subsequent, real work to be done by a philosopher: the outcome was certainly not self-evident, and this particular task occupied much of Part One of *TJ*. But in *LP*, where reflective equilibrium, as we have just seen, exists only in a declaratory and pale fashion, with no implications (similar to those in *TJ*) drawn from it, the subsequent process of reasoning is left in a moral vacuum. No wonder that this process has only the most cursory account given of it by Rawls in *LP*; it is basically reduced to the following statement: "the representatives of well-ordered peoples [both liberal and non-liberal but decent] simply reflect on the advantages of these principles of equality among peoples and see no reason to depart from them or to propose alternatives" (41). This is a very clear impoverishment of the nature of the hypothetical "deliberation" leading to a social contract in *TJ*. Rawls makes it even more explicit: "The parties [to the social contract in *LP*] are not given a menu of alternative principles and ideals from which to select, as they are in *Political Liberalism*, or in *A Theory of Justice*" (41). What can justify such a change?

It is important to be careful about the criticism of Rawls at this point. It is not sufficient simply to state that the whole process of the social contract has been "impoverished" (as it certainly has been), and then infer that this is a weakness of the new theory. Social contract stands for *something*, and the "impoverishment" thereof reflects a change in the model of reasoning. From *PL* onwards Rawls insisted that the "social contract" is a device of representation of our moral reasoning on justice: it models "restrictions on reasons in such a way that it becomes perfectly evident which agreement would be made by the parties as citizens' representatives".²² He further urged us to see this in the light of political constructivism in which "the principles of political justice (content) may be represented as the outcome of a procedure construction (structure)" – a procedure in which "rational agents, as representatives of citizens and subject to reasonable conditions, select the principles to regulate the basic structure of society".²³ In *LP*, he declares, accordingly: "Since the original position includes the veil of ignorance, it also models what we regard as appropriate *restrictions on reasons* for adopting a political conception of justice..." (30, emphasis added).²⁴ Hence, what we called a moment ago an "impoverishment" of the social contract can be

²⁰ *Id.* at 21.

²¹ *Id.* at 20.

²² *PL* at 26.

²³ *Id.* at 93.

²⁴ Here Rawls echoes his own identical formula from *PL* at 26.

more generously seen as appropriate “restrictions on reasons” to be taken into account when making decisions about international justice. Viewed in this way, the realism-utopianism dilemma returns: by disallowing the parties to the original position a choice from a reasonably broad “menu” of possible principles, and by confining their choice to those principles that can be culled “from the history and usages of international law and practice” (41), Rawls clearly identifies very rigorous restrictions upon the reasons available to the parties in the original position when the rules of international justice are to be decided. These reasons are largely confined to the already existing (although properly pruned) principles of customary law and good international practice.

This is in stark contrast to the *TJ* social contract, in which the parties had a reasonably broad “menu of alternative principles and ideals from which to select” available to them. By discarding some of these “alternative principles and ideals” (in particular, utilitarianism and intuitionism understood as “the doctrine that there is an irreducible family of first principles which have to be weighed against one another”)²⁵ Rawls had, in *TJ*, discharged the burden of showing why some reasons for adopting a conception of justice are better than others. It should be recalled that, in the deduction of a plausible conception of justice from the original position (the position, it should be emphasised, that merely modelled the right reasons to be taken into account, namely disinterestedness, reciprocity, universalisability, etc.), a particularly important role was played by the strategy of “maximin”, which was supposed to be the most reasonable strategy to be adopted by rational individuals acting under conditions of uncertainty as to the actual scenario of events that will actually occur. This strategy recommended that “we are to adopt the alternative the worst outcome of which is superior to the worst outcome of the others”.²⁶ It seems, however that there is no maximin strategy at work in *LP*. It neither emerges in an explicitly articulated fashion nor is it *de facto* at work in the reasoning leading to the eight principles of the Law of Peoples. For if it was, it would be extremely implausible for the maximin *not* to dictate something like a principle of global distributive justice (even if the participants were restricted to the “well-ordered” peoples: after all, under the *LP*-style veil of ignorance, the parties to the original position do not know their level of wealth) and a robust international human rights regime. It still may be a controversial proposition: it could be argued that the parties would reject these two principles. However, the point is that, in *LP*, Rawls does not even “allow” the parties to the original position to contemplate these conceptions as potentially available options: these conceptions do not even make it into the “menu of alternative principles and ideals from which to select” (41).

²⁵ *TJ* at 34.

²⁶ *Id.* at 153.

Now we can see that the rather complex model of moral reasoning on justice employed in *TJ* has undergone a truly fundamental transformation in *LP*. The *TJ* model consisted, to put it schematically, of three ingredients. Firstly, the device of reflective equilibrium as a coherence-aimed framework oriented towards mutual adjustment and readjustment between our pre-reflective intuitive specific convictions of justice and general, abstract principles of justice. Secondly, the original position, modelling the type of reasons that are acceptable when proposing the principles of justice (or, put negatively, aimed at discarding those reasons that may be used self-servingly by the individuals trying to tailor the general principles of justice to serve their own individual needs). As a result, only universalisable, disinterested reasons can be filtered through into the process of reasoning about justice. Thirdly, and most substantively, the strategy of maximin as the most rational strategy of decision-making under the conditions of relative uncertainty: an extrapolation of a real decision making strategy into the field of moral reasoning, with of all the important (though, no doubt controversial) presuppositions such as risk-aversion etc. that this implies. None of these three essential ingredients are present in *LP*, or, if they are, it plays nothing like as important a role in representing a model of moral reasoning as it did in *TJ*. Reflective equilibrium is present, if at all, only in the palest of forms, and has effectively no role in shaping the description of the original position. The original position has been eroded by restricting the types of reasoning about justice that are available to the parties. And the principle of maximin is absent, or at least, is not detectable in the deliberative process leading to the eight principles of international justice.

Again, the point is not to “deplore” the impoverishment of the model of social contract. It is rather to conclude that “social contract”, as a particular, persuasive model of representation of our actual, plausible moral reasoning about the principles of justice, has all but evaporated from *LP*, despite the lip service paid to it in the vocabulary employed by Rawls. The social-contract model is interesting to us in the context of theorising about justice (and was so plausible in *TJ*) only insofar as it helpfully encapsulates some plausible, attractive structural devices of reasoning, such as reflective equilibrium, the reasons for the specific design of the original position, and maximin. Once these devices are gone, the attractiveness and effectiveness of social contract *qua* model of representation for moral reasoning on justice is almost completely eroded.

6.

Not only the (seeming) rejection of the maximin, but also the adoption of the current division of the world into independent and (more or less) sovereign states is a very important factor in reducing the range of options available to the parties (whoever they might be) to the social contract in *LP*. It is never particularly clear whether the current state-based structure of the world is treated by Rawls as a fact that is morally neutral, and therefore that may, but also may *not* (depending on

the degree to which we wish to incorporate a set of empirical, contingent facts into the theory-building process) be a factual constraint upon the choices available to the parties of the contract *or* whether it is an ethically meaningful factor with a positive value attached to it. Of course, which interpretation is privileged is of considerable importance. The former interpretation would be largely consistent with *TJ*, which tried to remain neutral regarding the fact of borders between nation states. Under this interpretation, all that matters is that the justice-constituency (the set of people who owe the duties of justice to each other) is a bounded one, but the nature and delimitation of the borders of the justice-community is something external to the theory itself: it may be sub-state, state, or planetary. In that sense, the existence of borders was a non-moral issue, and was not built into the conception itself, apart from an understanding that there must be *some* borders because we need to know the scope of the community – of the justice-constituency, if you will – within which the mutual ties and interrelations produce the duties of justice. Both “cosmopolitans” and “nationalists” could take up Rawls’s theory and use it for their own purposes: the borders exist, but they were considered to be morally insignificant. (Which, incidentally, is not the same as to say that they were morally arbitrary: if something is “morally arbitrary”, in Rawls’s parlance at least, then we have a duty to offset its impact upon the social distribution of primary goods: this is the case of the “natural distribution of talents”, which is said to be morally arbitrary, and as such demands compensatory measures.²⁷ However, the existence of borders between nation-states was not “morally arbitrary” in this sense; rather, it was a non-moral, or morally insignificant, variable.) In *LP*, while we are told that societies’ boundaries “may appear [arbitrary] from a historical point of view” (38) nevertheless we are warned that “to fix on their arbitrariness is to fix on the wrong thing” (39). Why? Because, we are told, “[i]n the absence of a world-state, there *must* be boundaries of some kind...” (39). But this is an obviously question-begging argument: whether there should be a “world-state” (as a code word for some form of cosmopolitan international justice) is precisely what is at issue here.²⁸ What is crucial is the moral status of the non-existence of a “world-state”: is it just a fact of life, which is extrinsic to the theory and which therefore need not be taken into consideration in the original social contract, or is it a morally good thing in itself?

There is much, however, that suggests that Rawls opted for the latter interpretation, that is, that the division of the world into nation-states is a morally meaningful fact, and as such should constitute one of the parameters of the original social contract that cannot itself be challenged or transcended. The

²⁷ See *id.* at 100-108.

²⁸ For a criticism of Rawls’s dismissal of a world state see Tasioulas *supra* note 8 at 369 n. 2. For a critique of “state-centredness” in Rawls, see Seyla Benhabib, *The Claims of Culture* (Princeton University Press: Princeton 2002) at 168-70.

strength with which Rawls rejects the idea of cosmopolitan international order (82) and of “world government” (36) implies that the alternatives to the state-based structure of the world are not merely rejected as too “utopian”, but also as not particularly attractive. “The Law of Peoples proceeds from the international political world as we see it...” (83), Rawls tells us, and since the Law of Peoples is part of the *ideal* theory (as noted above), the implication is that it is something more than a grudgingly adopted concession to the prevailing international reality. However, the normative grounds for such a world structure, understood in positive terms, are missing, and they are a source of understandable concern for those who derived – fairly and correctly, in my view – universalistic conclusions from *TJ*. It is also disconcerting that Rawls does not consider as a serious alternative a world structure in which the importance of national borders will gradually diminish, in which state sovereignty will be progressively replaced by supra- and sub-national entities, and in which the post-Westphalian world structure of self-sufficient, sovereign units called states will become obsolete.²⁹

To be sure, Rawls is at pains to emphasise that he is deliberately employing the concept of “peoples” rather than “states” (23-30); but this substitution is not significant enough to dispel the impression that Rawls adopts the division of the world into states as a normatively meaningful presupposition of his theory.³⁰ “Peoples” are not defined by Rawls in an ethnic-national way (which is understandable and positive, as international justice based on a social contract between the peoples ethnically defined would be a decidedly unattractive idea!). The consequence, however, is that his “peoples” are largely identical with the societies bounded by state frontiers. It is significant that we will not find in Rawls any strong support for the idea that stateless “peoples” are equal partners of all other peoples in the international community of peoples;³¹ on the contrary, we can find several hints that only those “peoples” that are territorially defined and represented by a government register in Rawls’s conception. Right at the beginning we are told that “an important role of government, however arbitrary a society’s boundaries may appear from a historical point of view, is to be the effective agent of a people as they take responsibility for their territory and the size of their population...” (8). On the other hand, however, when the “Society of Peoples” is defined in the abstract and most general terms, we are told that it comprises of “all those peoples who follow the ideals and principles of the Law of Peoples in their mutual relations” (3). There is nothing that implies, *at this initial stage*, that only the peoples organised in states may be members of the community of peoples in good standing, and thus be subject to the rights and obligations stemming from the Law of Peoples. At the level of the definition of

²⁹ For a general critique of Rawls along these lines, see Buchanan, *supra* note 14.

³⁰ As John Tasioulas puts it, “the technical notion of a ‘people’ may obscure its deep affinities with that of a state”, Tasioulas, *supra* note 8 at 374.

³¹ See similarly Buchanan *supra* note 14 at 699.

“people” there is nothing that necessitates state-boundedness; and yet, at the level of the development of the theory, this state-boundedness emerges as the obvious template within which the peoples are viewed. After all, most of the eight principles of the Law of People are capable of being complied with only by state (or quasi-state) entities: peoples are therefore modelled on nation-states. The upshot is that the distinction between “peoples” and “states” is in fact less significant in the logic of Rawls’s theory than Rawls himself declares.

7.

One other symptom of the surprising concessions to “reality” in Rawls’s “realistic utopia”, hard to reconcile with status of the Law of Peoples as an ideal theory, concerns the role of *governments* in the social contract leading to the eight principles. In a long essay, on which *LP* was eventually based, and which carried the same title, Rawls made the following observation as a response to the proponents of a “global original position”: “Historically speaking, all principles and standards proposed for the law of peoples must, *to be feasible*, prove acceptable to the considered and reflective public opinion of peoples *and their governments*”.³²

The status of the “feasibility” proviso in this proposition is unclear. Why must the principles be acceptable to the *governments* in addition to their acceptability to the *peoples* in order to pass the constructivist test of justification? After all, the Law of Peoples is determined in the same constructivist way as principles of justice in the conception of justice-as-fairness; hence, only the “appropriate reasons” guiding the specification of the Law of Peoples under “fair conditions” count (32). It is true that principles *unacceptable* to governments (while acceptable to their peoples) have little chance of being universally followed; but then we face the issue of non-compliance, and hence of non-ideal theory. The principles of the Law of Peoples, on the other hand, belong to ideal theory, which aims to describe the world “in which all peoples *accept and follow* the (ideal of the) Law of Peoples” (89, emphasis added). As I have already observed several times, Rawls explicitly states that the inclusion of hierarchical societies into the Law of Peoples belongs to ideal theory; it is therefore not a step triggered by non-compliance, unfavourable conditions, etc, and, as such, is subject to the same justification procedure as within the liberal societies. The feasibility test demanding an additional acceptance of principles by government, over and above that of their people, presupposes that they are not the accurate spokespersons for

³² See Rawls *supra* note 4 at 50, both emphases added. I have not found an equivalent statement in the book but neither have I found any clear, or even implicit, repudiation of the view expressed in this statement. In fact, there are several implicit reiterations of this point; for instance, in the context of his rejection of an idea of “global original positions” in which all persons (as opposed to peoples or their governments) participated, Rawls adds: “The Law of Peoples proceeds from the international political world *as we see it...*”, *LP* at 83, emphasis added.

their peoples' preferences; but this seems to put them outwith the category of societies that are "well-ordered and just". They are therefore worse than being merely non-liberal, in that they are "hierarchical" in the Rawlsian sense: not perfectly democratic, not separating the church and state, etc. Those societies in which the governments routinely fail to track the preferences of its peoples must surely fall below the level of "hierarchical but decent". It is significant that Rawls explicitly contrasts a "consultation hierarchy" with a "paternalistic regime" (72), with the implication that the latter would not pass the test of a well-ordered society. It follows that a decent hierarchical regime must track the avowed preferences of its people and form a conception of the common good through consultative means, rather than through a paternalistic announcement of the "true" common good of the citizens regardless of what they themselves think about it. The only difference between the common good test (displayed by decent hierarchical societies in Rawls's theory) and paternalism (disavowed by Rawls) is that in the former but not in the latter we have a government that tracks the *avowed* preferences of the people, even though it does it through instruments other than representative and liberal democracy.

Therefore, if decent peoples are indeed distinguishable from paternalistic ones, there is no reason for their *governments* to be included in the reflective equilibrium on the law of peoples. In the end, we face an alternative: either they are so non-democratic as to place themselves beyond the pale of well-ordered societies, or they do track the preferences of their people, in which case they need not be included because they are treated, in the theory of justification, merely as a mouthpiece for their people: it is redundant to include them in the original position as independent parties to the contract.

8.

We should now cease playing "Hamlet" without the Prince, and introduce the overarching value that, in *LP*, is responsible for all these puzzles regarding the construction of the social contract in this particular way, so deferential towards the "realities" of international relations as we know them: the erosion of the contract by a dramatic reduction of the options available to the parties, the abandonment of the reflective equilibrium and of the maximin strategy, and most crucially, the selection of peoples rather than individuals as the parties to the original contract defining the terms of international justice. That overarching value is the ideal of toleration. There can be very little doubt to the readers of *LP* that, whenever the troubling questions concerning the bias in favour of realism in the "realistic utopia" appear, it is the importance of toleration for other, non-liberal societies that is provided as the response. When explaining why liberals should not attempt to force their ideals of what constitutes a good society upon the nonliberal peoples, Rawls says: "If all societies were required to be liberal, the idea of political liberalism would fail to express *due toleration* for other acceptable ways (if such there are, as I assume) of ordering society" (59,

emphasis added). In the same context Rawls clarifies that his meaning of toleration goes beyond the precept of non-interference: “to tolerate means not only to refrain from exercising political sanctions – military, economic, or diplomatic – to make a people change its ways. To tolerate also means to recognise these nonliberal societies as equal participating members in good standing of the Society of Peoples, with certain rights and obligations...” (59). Much later in the book, when he elucidates the reasons for his acceptance of the non-liberal but decent societies, as “deserv[ing] respect”, (84) he provides the following, telling analogy: “Liberal societies may differ widely in many ways: for example, some are far more egalitarian than others, yet these differences are *tolerated* in the society of liberal peoples. Might not the institutions of some kinds of hierarchical societies also be similarly *tolerable*? I believe this to be so” (84, emphases added, footnote omitted). One can supply many more citations from Rawls’s book that, whenever the set of problems as identified earlier in this article arise, appeal to the ideal of toleration as an answer.

But if toleration is an answer, what was exactly the question? Before we go on inquiring into whether the ideal of toleration can perform such an onerous job as it is assigned in Rawls’s theory, it is important to be clear about the *function* that the ideal of toleration *may* perform (if we share the assessment of its worth) in a conception of international justice such as Rawls’s. To put it in negative terms, it is important to clarify what the ideal of toleration is *not* about in this context. It is not about the permissibility and the potential scope of governmental interference by one state (say, a liberal one) in the institutional and legal arrangements of another state (say, a nonliberal one). This is consistent with what we have observed earlier (in section 3 of this article): an ideal theory of international justice does not provide directly any prescriptions for governmental policy *vis-à-vis* other states; rather, it is about what good theoretical reasons we, as individual citizens, have to criticise and “interfere”, through the means available to us, in societies whose domestic institutions, distribution patterns, human rights record etc. we dislike. The forms of “interference” available to us, as individuals, are usually related to the advocacy of change in a different society: we do not have the means of coercive interference. This sounds banal, but it needs to be clearly stated because Rawls himself, at times, seems to be oscillating between the two orders: the order of an ideal theory containing prescriptions for individual liberal citizens, and the order of pragmatic policy proposals with prescriptions for a governmental action. He says, for example: “A decent hierarchical society meets moral and legal requirements sufficient to override the political reasons we might have for *imposing sanctions* on, or *forcibly interfering* with, its people and their institutions and culture” (83, emphases added). However the principle of toleration that is relevant to the context of international justice (in order to play the role assigned to it, that is, explaining why we should embrace some nonliberal societies within the Law of Peoples) does not reach the question of the acceptability of governmental sanctions or coercive interference. A prior

question, preceding that of sanctions and interference, is whether we, as individuals, have good reasons to try to effect change in the societies that we dislike: the principle of toleration may pre-empt, and render improper, such an aspiration *at this stage*. If we decide that, indeed, the principle toleration overrides the moral rightness of advocating change, the argument ends there. If we conclude, however, that toleration is incapable of trumping the aspiration to effect the change, only then can we move to the secondary issue of what means of such interference are acceptable for the citizens to advocate and for the governments to undertake. And it may be the case that even if we had good moral reasons to interfere in abstract, “zero-cost” circumstances, the actual costs of interference may render it unjustified; this, however, would not be due to toleration, but rather to pragmatic considerations related to the costs of interference in the real world, the risk of error and miscalculation, etc.

The upshot is: the principle of toleration should be treated in the context of international justice as an ethical norm rather than a pragmatic policy recommendation. In this perspective, is this ethical norm capable of supporting the inclusion of the hierarchical, non-liberal but decent peoples into the Society of Peoples with the consequences described by Rawls? Or, going backwards in the chain of reasoning, is it capable of persuading us to modify the rules of the social contract, with the constraints on the options available, the abandonment of the maximin strategy, and – most importantly – the replacement of individuals by “peoples” in the original position? I do not think so. I will present my reasons for holding this view in a succinct way, and I hope that I can be excused for this brevity by the fact that I have, elsewhere, explained in more detail why the principle of toleration is not a good reason to object to the universality of human rights (which is a parallel argument to the one considered here).³³

Toleration is an important ethical value: there are good moral reasons why we should often abstain from interference with other people’s sincerely held values (and, consequently, with their behaviour based on these values) even if we reject them. Toleration is not (and, in a weaker version of the argument, need not be) based on some form of moral agnosticism; it can be (and, in a stronger, and more attractive version of the argument, must be) based on strong substantive moral reasons that deny us the moral right of imposing our other (other than toleration, that is) moral values on those who do not share them. Liberals who believe in the value of toleration must, within limits, abstain from trying to impose liberal conceptions of life, such as the conception of an autonomous life, shaped by the individual herself, subjected only or predominantly to one’s own choices, upon other people. Liberal toleration requires, consequently, that we do not interfere with non-liberal forms of social life, insofar as they are necessary to support

³³ Wojciech Sadurski, “‘It All Depends’: The Universal and the Contingent in Human Rights”, EUI Working Paper Law 2002/7, European University Institute, Florence 2002.

genuinely shared preferences for non-autonomous conceptions of life. We must accept, and draw conclusions from the fact, that there are some people whose personal self-fulfilment best occurs on non-autonomous forms of life: in religious orders, for example, or in the military, or in a highly authoritarian family. If we, the liberals, believe in toleration, we must respect (that is, abstain from interfering with) lives that are non-autonomous, insofar as they fulfil the genuine needs and desires of the individuals who lead these lives. This abstention is in itself morally meaningful: respect for non-autonomous lives is based on the recognition of human dignity, of the importance of self-fulfilment, and of the fact that there are different paths for different people for reaching self-fulfilment; some of these paths are non-autonomous.

Such a morally meaningful principle of toleration is, of course, a controversial ideal, and it raises all sorts of different problems that have absorbed, and divided, liberals for decades. The most important and difficult of these problems is encapsulated in the question: should non-autonomous ways of life adopted in a non-autonomous (or less than fully autonomous) way also be respected, or is the respect for non-autonomous ways of life due only when there has been an autonomous choice to live a non-autonomous life? We all know how difficult and fraught with dangers each option is. If we choose the former answer (extending respect for non-autonomy also to non-autonomous processes of deciding about non-autonomous life), our “toleration” turns into an insensitivity to oppression, exploitation, discrimination and coercion. We know that non-autonomous ways of life are often imposed upon people, or are adopted (or accepted) by those people as the last resort – to reduce the suffering related to the lack of freedom – and insensitivity to this fact erodes the moral worth toleration has in the first place. But if we choose the latter answer (respect for non-autonomous ways of life but only if adopted freely, autonomously), then we run immediately into seemingly insurmountable problems of distinguishing between the second-order decisions *about* the choice of the way of life, and the first-order decisions that *constitute* the way of life. But human life, as we know, does not divide so easily into first-order and second-order decisions, or into our initial choices and their consequences. It is not as if, at certain point in our life, we come across a doorway leading to a cell, and the non-autonomous way of life enclosed within the walls will be certified by a liberal as valuable (hence, deserving of toleration) if the decision to enter was taken freely. Such imagery is schematic in the extreme, and to base the liberal acceptance of non-autonomous lives on a distinction between the *procedure* of the individual decision-taking and the *substance* of one’s decisions is a somewhat questionable foundation upon which to erect the principle of liberal toleration.

Fortunately, in the context of discussing *LP*, we do not have to choose between the broader and the more integral liberalism: the one that is more deferential to non-autonomous ways of lives (and is prepared to ignore the defects in terms of autonomous choices about such lives) and the one that considers individual

autonomy to be the paramount value dictating the parameters of, rather than being constrained by, the principle of toleration. We need not enter into this controversy because even if we adopted the broadest conception of toleration, one that would be non-critical of the non-autonomous processes of choosing a way of life, we would nonetheless be doing it for the (alleged) good of the individual concerned. The moral worth of the principle of toleration would stem from the (alleged) benefit to the individual to whom “toleration” is due: otherwise, the toleration itself would be left in a moral vacuum. It would be perverse to demand toleration for someone’s choices if we thought these choices to be at least partly non-autonomous, and there was no particular benefit (of moral significance) that such toleration conferred upon the individual concerned. Ultimately, toleration in terms either conception is based upon normative individualism; it is the good of the individual that endows the ideal of toleration with moral worth. It is morally valuable insofar as it recognises the ultimate value of human dignity and expresses respect for persons.

LP, however, fails to identify any convincing connection between the principle of toleration and the good of the individuals who are to be tolerated: it is peoples, rather than the individuals themselves, who are the avowed beneficiaries of the principle of toleration as announced in his book. If, however, a “people”, through its institutional design and practices, is intolerant towards some of its members, then liberal toleration is due to those individuals rather than to the “people” itself: *their* good is the ultimate source of the moral worth of toleration.³⁴ Rawls never tries to convince us, and we have no reason to believe, that the individual members of the hierarchical but decent peoples are, on balance, the net beneficiaries of the principle of toleration as demanded in *LP*. The central question that would have to be asked is, what is the balance of benefits and costs for the individual members of the decent hierarchical societies produced by the adoption of the principle of toleration of such societies? If Rawls had followed the line of reasoning suggested by this question, he would no doubt have had to have given a rather complex and differentiated answer: he would have had to admit that the benefits and costs divide unequally *within* hierarchical societies, and that some are the gainers while others (perhaps, a majority) are the losers. A principle of “toleration” of gender inequality in a hierarchical society will definitely make some people worse-off (compared to a principle of interference) even though it will also make some people comfortable in sustaining their ways of life based, as they are, on societal discrimination against women. Now both parts of this proposition will have to be made more subtle and nuanced than merely stating that men will benefit from such “toleration” while women will be its victims: there may be many men who would rather see the end of open discrimination but who for complex collective-action problems are unwilling to

³⁴ See, similarly, Michael Blake, “Toleration and Reciprocity: Commentary on Martha Nussbaum and Henry Shue”, *Politics, Philosophy & Economics* 1 (2002): 325-35 at 326.

undertake the steps towards reform on their own, just as there may be some women who may feel such a strong attachment to their local culture that they will genuinely abhor any departure from the traditional gender-defined roles, even if these mean deep inequality. The latter case will call for an investigation into the degree to which such commitments are genuine as opposed to being a form of “adaptive preferences”, or having been adopted under pressure or fear of coercion, etc. I will not follow through on this line of argument³⁵ because one thing is clear: this is *not* the type of question that Rawls entertains. The connection between the principle of toleration and an unequal distribution of the benefits and costs produced by this principle for its alleged beneficiaries is nowhere elucidated in the book; this consideration is simply precluded by identifying “peoples” as the beneficiaries of the principle of toleration. But this identification is profoundly implausible for the reasons I have just listed: if the benefits and costs divide unequally within a people, it is simply insufficient to declare that the “people” itself is the beneficiary. A more complex picture needs to be drawn, and the fact of an on-balance benefit needs to be evidenced rather than merely asserted.

Moreover, a partisan of the principle of toleration *à la* Rawls would have to go a step further, and test not only whether different members of decent hierarchical societies benefited from non-interference by liberal societies: one would, in addition, have to affirm that they *supported* it. “Toleration” towards someone who does not want it loses its moral bearings: it becomes an indifference to someone else’s misfortune. If I desire your interference aimed at changing my situation, then your abstention cannot be presented as “toleration” in any meaningful sense of the word. Now whether all (or even many) members of “hierarchical but decent” societies in fact favour the policy of non-interference in their domestic arrangements is a very difficult thing to verify, due to the fact that, by the very description of those societies provided by Rawls, a reliable account of views, values and public opinion in those societies is not easily available. By the same token, however, we can rarely say with any degree of certainty that all members of such societies reject the idea of interference from the outside: some do and some don’t, and it matters a great deal who actually does and who does not. For it is a very plausible hypothesis that those who do support outside interference find themselves on the lowest rung of the social ladder, and would thus be representative of the “worst-off” people who, from the *TJ* perspective, would need to be taken into account first in designing the “right” system of justice. Unless we adopt a deeply implausible (and alien to the overall philosophical system of Rawls as developed in *TJ* and *PL*) collectivist ontology of a “people”, under which it constitutes a single entity with its own undifferentiated subjectivity, we have absolutely no reason to accept at face value the declarations of the *leaders* of decent hierarchical peoples (and of their

³⁵ But more in Sadurski, *supra* note 33 at 13-16.

ideologues) concerning the true “values” of those peoples. Neither have we any reason to accept at face value the declarations of those leaders as to whether their people accepts or rejects the idea of interference from the outside (even although we know that internal dissent towards the leaders’ declared ideology does not necessarily translate into support for external interference – but this is something we do not need to go into here). We know –on the basis of our empirical experience (and the debate about so-called “Asian values” is helpful in this regard³⁶) – that very often, probably more often than not, the declarations by the leaders of societies that would fit Rawls’s description of decent hierarchical peoples’ are nothing but ideological rationalisations of oppression, privilege and authoritarianism.³⁷ We know that, when we are told that “their” societies dislike the idea of democracy, of individual rights, of freedom of the press etc., the leaders of authoritarian societies are developing a self-serving, deceitful rationalisation in order to exempt themselves from political accountability and critique. We must not take these ideological rationalisations at face value, nor believe that the citizens of societies subjected to this rule actually share them.

How has Rawls reached the point at which the respect for an undifferentiated “people” may make us oblivious to the diverse voices reaching us from non-liberal, non-democratic societies – including the dissident voices, and the voices of the oppressed? How is it that he reached the paradoxical result whereby an attempt to escape parochialism led to the obviously parochial view that the notions of individual dignity and respect for individuality are specifically “our” values, not shared by individuals in non-Western societies? The central motivating engine of the elevation of “toleration” as an overarching value in *LP* is an aspiration to avoid an arrogant view that we – “we” the liberal democrats – are in a position to impose upon others our views on how best to design a society and protect rights. The liberal conceit of giving others lessons in political morality is the main target of Rawls’s critique. In the background, there is an analogy to a powerful liberal government that wishes to impose the best liberal ways of life – the life of autonomy, self-creation and proud independence – upon all individuals and groups in society, including those who reflectively and genuinely reject such conceptions for themselves. Such an ideal of a liberal government that wants to implant liberal values “all the way down” within society was the target of Rawls’s critique in *PL*, and in *LP* he extrapolated this to the relationship between liberal and non-liberal states. The analogy, however, is misleading; liberal states are not like liberal bullies trying to refashion everyone in their own image. A better domestic analogy would be with a liberal-democratic government that tries to extend the protection of fundamental individual liberties,

³⁶ See Mark R. Thompson, “Whatever Happened to ‘Asian Values’?”, *Journal of Democracy* 12 (2001): 154-65.

³⁷ See, similarly, Henry Shue, “Rawls and the Outlaws”, *Politics, Philosophy & Economics* 1 (2002): 307-24 at 312.

and to guarantee non-discriminatory treatment and fair satisfaction of material interests to everyone in the society, regardless of their group membership. Such a liberal-democratic government would need, at times, to take on illiberal groups whose leaders wish to maintain their power over the members of the group, including those most vulnerable, powerless, and oppressed.

A liberal-democratic government resists accepting the declarations of leaders of such illiberal groups, concerning the “true” values of group members, at face value; it insists on some common threshold of protection for everyone in the society it governs. It is *this* model that is analogous to the position of liberal peoples towards the illiberal ones in the world: liberals should insist that everyone, regardless of the state of which they happen to be citizens, should benefit from certain protections and liberties, similar to those that the citizens of liberal-democratic states enjoy in their own countries. There is no reason to believe that “enlightenment about the limits of liberalism” (78) dictates a satisfaction with the confinement of protection of individual liberties to liberal states only, and disregard for the victims of oppression in authoritarian states. Such a disregard would be based on the belief that the ideological rationalisations for authoritarianism in non-liberal societies are genuine reflections of actual consensus there. The readers of Rawls’s *Theory of Justice* and *Political Liberalism* know, however, that such an acceptance of ideological rationalisations for oppression and suffering would have been the last thing he wanted.