

Francesca Pasquali

RAWLS'S REALISTIC UTOPIANISM: A CRITICAL DISCUSSION

working paper

 Centro Einaudi

N5 2009 ISSN 2036-1246



Laboratorio di Politica Comparata
e Filosofia Pubblica

FRANCESCA PASQUALI

**RAWLS'S REALISTIC UTOPIANISM:
A CRITICAL DISCUSSION**

Centro Einaudi • Laboratorio di Politica Comparata e Filosofia Pubblica
con il sostegno della Compagnia di San Paolo

Working Paper-LPF n. 5 • 2009

© 2009 Francesca Pasquali e LPF • Centro Einaudi

Francesca Pasquali si è laureata in Filosofia e ha conseguito il dottorato di ricerca in Studi Politici presso l'Università degli Studi di Milano, con una tesi dal titolo "Precarious Equilibrium. Political Philosophy between Desirability and Feasibility". I suoi interessi di ricerca vertono su questioni di carattere metodologico e metateorico nel campo della filosofia politica, concentrandosi in particolare sui compiti pratici della filosofia politica, sul confronto tra teorie realistiche e utopistiche e sul rapporto tra fatti e principi normativi.

francesca.pasquali@unimi.it

Il Laboratorio di Politica Comparata e Filosofia Pubblica promuove attività di studio, documentazione e dibattito sulle principali trasformazioni della sfera politica nelle democrazie contemporanee, adottando sia una prospettiva descrittivo-esplicativa che una prospettiva normativa, e mirando in tal modo a creare collegamenti significativi fra le due.

L'attività del Laboratorio, sostenuta dalla Compagnia di San Paolo, si concentra in particolare sul rapporto fra le scelte di policy e le cornici valoriali all'interno delle quali tali decisioni sono, o dovrebbero essere, effettuate.

L'idea alla base di questo approccio è che sia non solo desiderabile ma istituzionalmente possibile muovere verso forme di politica «civile», informate a quel «pluralismo ragionevole» che Rawls ha indicato come tratto caratterizzante del liberalismo politico. Identificare i contorni di questa nuova «politica civile» è particolarmente urgente e importante per il sistema politico italiano, che appare ancora scarsamente preparato ad affrontare le sfide emergenti in molti settori di policy, dalla riforma del welfare al governo dell'immigrazione, dai criteri di selezione nella scuola e nella pubblica amministrazione alla definizione di regole per le questioni eticamente sensibili.

LPF • Centro Einaudi
Via Ponza 4 • 10121 Torino
telefono +39 011 5591611 • fax +39 011 5591691
e-mail: segreteria@centroeinaudi.it
www.centroeinaudi.it

INDICE

RAWLS'S REALISTIC UTOPIANISM: A CRITICAL DISCUSSION	5
Introduction	5
1. Desirability and feasibility in <i>A Theory Of Justice</i>	7
2. From <i>A Theory Of Justice</i> to <i>Political Liberalism</i> : Rawls's new focus on stability	13
3. The ambiguities of Rawls's political approach	20
4. The shortfalls of Rawls's realistic utopianism	25
5. Realistic utopianism: an assessment	33
References	36

KEYWORDS

political philosophy, John Rawls, desirability, feasibility, utopianism, realism

ABSTRACT

RAWLS'S REALISTIC UTOPIANISM: A CRITICAL DISCUSSION

The paper investigates Rawls's realistic utopianism with reference to two orders of methodological criteria, namely desirability and feasibility. The paper shows how the roles attributed to the two criteria gradually change throughout Rawls's path from *A Theory of Justice*, to *Political Liberalism* and, finally, to *The Law of Peoples*. The paper focuses, in particular, on the implications connected to Rawls's attempt at enhancing the practical significance of his principles, an attempt that leads him to ascribe an increasing centrality to feasibility considerations and to downgrade his concerns with the desirability and theoretical adequacy of the proposed principles. The paper questions Rawls's methodological approach and it suggests that realistic utopianism tends to distort the character of both desirability and feasibility. More precisely, the paper argues that, due to its attempted synthesis between desirability and feasibility and to its search of a middle-ground between realism and utopianism, Rawls's realistic utopianism ends up trapped into ambiguities that render it unable to appropriately vindicate its principles and to fulfil the normative aspirations of political philosophy.

RAWLS'S REALISTIC UTOPIANISM: A CRITICAL DISCUSSION

INTRODUCTION

John Rawls's works have been at the very centre of debate for almost forty years and his way of understanding and practicing political philosophy has become paradigmatic, at least within the Anglo-Saxon world. This is hardly surprising considering that the publication in 1971 of *A Theory of Justice*¹ was perceived as a turning point for political philosophy. In *A Theory of Justice*, Rawls traces the path for rehabilitating the normative concerns of political philosophy vis-à-vis the predominant non-cognitivist attitude alimanted by logical positivism. Indeed, Rawls provides an argumentative style enabling political philosophy to combine analytical rigour with commitment to addressing substantive public questions. In addition, Rawls's conception of justice – *justice as fairness* – affirms the inviolability of individuals and individual rights, thus constituting a plausible alternative to utilitarianism, which was then the leading ethical doctrine, and it proposes a synthesis between liberty and equality that offered a reliable foundation for liberal commitments.

The influence of *A Theory of Justice* exceeds the disciplinary boundaries of political philosophy. Not only Rawls's arguments avail themselves of insights from the social sciences, but they have a direct bearing on such disciplines. Accordingly, *A Theory of Justice* has engaged, not only political philosophers, but also social scientists as well and it has been the object of both attentive study and punctual criticism. Significant critiques against Rawls's approach and his principles of justice have been raised by thinkers of diverse orientations: for instance, by libertarians, contesting the redistributive implications connected to justice as fairness; by communitarians, challenging Rawls's allegedly individualistic assumptions; and by feminist theorists, criticising his neglect for injustices taking place in private life. Rawls has been an attentive reader of his critics and he has revised and better clarified his substantive and meta-theoretical positions in the light of the received criticisms.

Published in 1993, *Political Liberalism* collects and systematises Rawls's writings that had appeared during the 1980s. In *Political Liberalism*, Rawls is primarily concerned

¹ Quotations from Rawls's works will be signalled as follow: references to *A Theory of Justice* (Rawls 1971) will be indicated by *TJ*, to *Political Liberalism* (Rawls 1993a) by *PL*, and to *The Law of Peoples* (Rawls 1999) by *LP*; references to Rawls's other works will be indicated, instead, by the year of the original publication.

with providing principles able to legitimise the practices and institutions of liberal democracies against the background of a plurality of ethical doctrines endorsed by their citizens. Like *A Theory of Justice*, Rawls's *Political Liberalism* has prompted intensive debate. Particular attention has been devoted to investigating and discussing the understanding of political philosophy that underlies Rawls's later reflection. As it is well known, in *Political Liberalism*, Rawls opts for a *political* conception of justice and he urges political philosophy to disentangle itself from philosophical or metaphysical controversies in order to endorse a more immediately practical function. This paper follows up such a discussion by focusing on Rawls's project for a realistically utopian political philosophy. The idea of political philosophy as realistically utopian is developed in great details in *The Law of Peoples*, where Rawls addresses questions of international justice and aims at proposing principles that are, at once, practically viable and normatively appealing.

This paper proposes to address Rawls's project for a realistically utopian political philosophy in order to enlighten its merits and its limits. To this end, the paper endorses a meta-theoretical perspective, thus bypassing discussion about the substantive principles developed by Rawls and focusing, instead, on the methodological strategy he employs. In particular, the paper will explore Rawls's approach with reference to two orders of methodological criteria, namely *desirability* and *feasibility*. This conceptual pair will be used as an analytical tool for clarifying the structure of Rawls's realistic utopianism and for providing insights about its appropriateness with respect to the normative commitments of political philosophy. As stated, desirability and feasibility are methodological criteria: they operate within political theories and orient them in the elaboration and justification of their principles and models. More precisely, desirability concerns the adequacy of principles and models, while feasibility regards the possibility of realising a given principle or model at the practical level. Desirability judgments are normative assertions: by stating that a certain principle is desirable, political philosophy does not intend to affirm that it *is* de facto desired, but that it *ought* to be desired. Feasibility judgments, instead, are not normative in character and they focus on the notion of constraint: a principle is feasible if its practical realisation does not imply the violation of certain constraints. Political theories endorse and defend a wide variety of desirability criteria – that is, what is defined as desirable varies widely from theory to theory – and political theories acknowledge different classes of feasibility constraints – connected to individuals' motivations and attitudes or to institutional arrangements and social or political practices. The internal articulation of desirability and feasibility will be better spelled out in the following pages, but it is necessary to emphasise that political theories combine desirability and feasibility criteria in different ways. The most influential models for political philosophy – realism and utopianism – assign primacy to one order of criteria over the other: realistic approaches ascribe priority to feasibility, while utopian perspectives attribute primacy to desirability. Rawls's realistic utopianism stands out for its sharp refusal to rank desirability and feasibility.

Rawls's project for a realistically utopian political philosophy relies on the commitment to situate itself midway between realism and utopianism and to find a

synthesis between desirability and feasibility. Rawls's realistic utopianism is the outcome of a revision process he has undertaken starting from the early 1980s, a process that has led him to modify the scope and the status of *justice as fairness* as well as its justification. The changes Rawls has introduced parallel the development and refinement of his meta-theoretical view about the role and function of political philosophy. The paper will examine how and why the roles attributed to desirability and feasibility and the relationship between them progressively changes throughout Rawls's path from *A Theory of Justice* to *Political Liberalism* and, finally, to *The Law of Peoples*. Nonetheless, the analysis here proposed has definitely no historiographic character and a detailed and exhaustive reconstruction of Rawls's thought and its development exceeds my objectives. The purpose consists, as already suggested, in investigating how desirability and feasibility criteria can be traced down in Rawls's reflection and in understanding how he argues in favour of the desirability of its principles and how feasibility considerations enter his argument².

1. DESIRABILITY AND FEASIBILITY IN *A THEORY OF JUSTICE*

In *A Theory of Justice*, Rawls's starting point for developing his principles and models is represented by what he calls *considered judgments*. In Rawls's definition, considered judgements are «those judgements in which our moral capacities are most likely to be displayed without distortion» (TJ, 47). That is, considered judgements are moral evaluations that, on due reflection, seem particularly sound and firm and appear unaffected by personal interests. Considered judgements are characterised by Rawls as «provisional fixed points which we presume any conception of justice must satisfy» (TJ, 20). Considered judgments are accounted for and systematised through the *original position*, the hypothetical situation of choice Rawls envisages as appropriate for devising principles of justice. More precisely, the original position models considered judgments through the *veil of ignorance* and through the *formal constraints on the concept of right*³. According to Rawls, the original position serves heuristic purposes: it can be interpreted as «a natural guide to intuition» (TJ, 139). As Rawls better explains, the «idea [...] is simply to make vivid to ourselves the restrictions that it seems reasonable to impose on arguments for principles of justice, and therefore on principles themselves» (TJ, 18). The outcome of the original position is a conception of justice – *justice as fairness* – that, Rawls argues, the *parties*, described as fairly located, equally rational and free, would choose.

² An account of Rawls's methodology developed with reference to desirability and feasibility criteria has been offered by Kukathas and Pettit in *Rawls. A Theory of Justice and Its Critics* (Kukathas and Pettit 1990). The analysis here developed will focus more specifically than they could on the later Rawls, on the Rawls of *Political Liberalism* and of *The Law of Peoples*.

³ On the veil of ignorance see TJ pp. 136-142. On the formal constraints on the concept of rights see, TJ, pp. 130-136.

It is plain to identify *justice as fairness*, and the two principles that make it up, as the *substantive content* of Rawls's desirability criteria, as what Rawls proposes as desirable⁴. Justice as fairness qualifies as desirable because it is the outcome of an «*appropriate* initial status quo which insures that fundamental agreements reached in it are fair» (TJ, 17, emphasis added). The original position is then the *appropriate perspective* from which to express desirability judgements. It interestingly combines partial and impartial standpoints: the parties are described as endowed with instrumental rationality and they are moved by self-interest⁵, but they are forced to act out of impartial reasons since their reasoning is constrained by the veil of ignorance and, accordingly, «no one knows his situation in society nor his natural assets, and therefore no one is in a position to tailor principles to his advantage» (TJ, 139). In turn, a similar constraint is warranted because it models our considered judgement according to which, for instance, «no one should be advantaged or disadvantaged by natural fortune or social circumstances in the choice of principles» (TJ, 18). *Reflective equilibrium*⁶ is the test Rawls employs in order to vindicate the appropriateness of the original position and of its constraints: it requires to verify whether and to what extent the principles chosen by the parties in the original position «match our considered convictions of justice or extend them in an acceptable way» (TJ, 19).

Correspondence between principles of justice and considered judgements is a necessary but not sufficient condition for vindicating the desirability of justice as fairness. As Daniel Weinstock points out, according to Rawls, a conception of justice is warranted only if it satisfies two conditions. First, it must be «appropriately related to certain shared considered moral convictions», as required by the *desirability* condition. Second, it must contribute «better than its principal competitors to the resolution of the conflicts and disagreements which have given rise to the need for a *theory* of justice» (Weinstock 1994, 166), thus satisfying the *reconciliation* condition⁷. Weinstock is right: on Rawls's account, correspondence with considered judgements is not enough for underpinning the desirability of justice as fairness. It only represents a *prima facie* justification. For its desirability to be fully vindicated, justice as fairness must be preferable, it must be more desirable than other conceptions of justice. Yet, since they both focus on the *adequacy* of normative principles, the two conditions identified by Weinstock can be reunited under the single label of desirability. However, Weinstock's observations suggest to emphasise that Rawls distinguishes between *absolute* and *relative* desirability, the former setting minimal requisites of desirability and the latter displaying criteria for comparing and ordering desirable options. In addition, Rawls

⁴ This is not intended to deny that the interpretation of the two principles is highly controversial, or that it is debatable whether the apparatus of the original position actually leads to Rawls's two principles. Yet, for the purpose of the analysis here proposed, it is not necessary to enter into such controversies.

⁵ See TJ, 142-150.

⁶ See TJ, 17-22 and 46-53. On the concept and potentialities of reflective equilibrium as a method of justification, see Daniels 1979 and 1996.

⁷ Weinstock adds a third condition, the *pragmatic* condition, according to which a conception of justice must be «feasible (in that it is possible to realize, given what we know about human society and psychology) and stable, in that it generates the basis of its own support over time» (Weinstock 1994, 166). The following pages address such a requirement.

presents the combination of the two dimensions of desirability as necessary in order for political philosophy to advance stringent and conclusive normative proposals. Indeed, while absolute desirability plays a preliminary function in delimiting the set of desirable options, considerations of relative desirability allow to select the most desirable option, the one to be normatively recommended.

Therefore, in order to properly vindicate the desirability of justice as fairness, Rawls must show that it is more relatively desirable compared to other conceptions of justice. This explains why the parties in the original position are presented with a list of alternative conceptions and they are asked to compare and adjudicate among them. The criteria the parties adopt for such a comparison are feasibility criteria. Feasibility is interpreted by Rawls in terms of stability: as he writes, «we checked to see [...] if justice as fairness is a *feasible* conception» and, as he affirms, «this forced us to raise the question of *stability*» (*TJ*, 580, emphasis added). Feasibility considerations, in fact, may require not simply to consider whether a certain principle or model can be realised at the practical level, but also to take into account whether a certain principle or model is able to endure over time and to reach stability. It is precisely feasibility intended in terms of stability that Rawls is concerned with. Nonetheless, it is worth stressing that thus intended and on Rawls's account, «rather than a characteristic of political systems, stability is an attribute of moral theories» (Klosko 1994, 1883): stability is a feature of conceptions of justice and a conception of justice is stable if it is able to «generate its own support» (*TJ*, 177). That is, as Rawls writes, «a conception of justice is stable when the public recognition of its realization by the social system tends to bring about the corresponding sense of justice» that induces the citizens of a society regulated by a set of principles «to develop a desire to act in accordance with these principles and to do their part in institutions that exemplify them» (*TJ*, 177).

In *A Theory of Justice*, Rawls's argument from stability comes into two stages⁸. At the first stage, stability considerations are introduced into the parties' reasoning within the original position: Rawls asserts that «for an agreement to be valid, the parties must be able to honor it under the relevant and foreseeable circumstances» (*TJ*, 175). As he writes, the parties «cannot enter into agreements that may have consequences they cannot accept» and they «will avoid those that they can adhere to only with great difficulty» (*TJ*, 176). The parties are led to take into account similar considerations by their general knowledge about human psychology and, more precisely, by their awareness about the *strains of commitment*, that is by their awareness that, although plausible to be subscribed to by the parties, certain principles may turn out to be particularly exigent for empirical individuals to comply with⁹. Rawls argues that, these facts acknowledged, the parties choose the principle of equal freedom and the difference principle. Thus, Rawls's two principles are preferred because they are more relatively feasible with respect to other principles: if compared to utilitarian principles, for instance, Rawls's

⁸ See Copp 1996.

⁹ See *TJ*, 145 and 176.

proposed principles are less demanding towards individuals since they «give a greater support to citizens' sense of justice and self-respect» (Freeman 2003a, 22).

The second stage of Rawls's argument from stability is sketched in the third part of *A Theory of Justice*. Here, Rawls argues, first, that justice as fairness is able to give rise to an adequate sense of justice and, second, that that «sense of justice is compatible, and can even constitute part of a person's good» (Freeman 2003b, 280). The former step of Rawls's argument relies on a psychological account of individuals' moral development showing that individuals who are brought up and live in a society regulated by justice as fairness tend to acquire a sense of justice and to develop an «effective desire to act upon the principles of justice» (*TJ*, 505). Therefore, justice as fairness is a stable conception of justice, since once it is publicly affirmed and developed into just institutions it is able to generate its own support by favouring the emergence of an appropriate sense of justice. Nonetheless, appealing to the sense of justice is not sufficient for vindicating the stability of justice as fairness. As Rawls asserts, the «stability of a conception depends on a balance of motives: the sense of justice that it cultivates and the aims that it encourages must normally win out against propensities toward injustice» (*TJ*, 454).

The latter part of the argument from stability is precisely meant to illustrate how the sense of justice is able to overrule «impulses and temptations to act unjustly» (*TJ*, 454). Accordingly, Rawls's argument focuses on the tension between the principles of justice and individuals' plans of life, on the tension between the right and the good. As he writes:

Just institutions are collectively rational and to everyone's advantage from a suitably general perspective. It is also rational for each to urge others to support these arrangements and to fulfil their duties and obligations. The problem is whether the regulative desire to adopt the standpoint of justice belongs to a person's own good. (*TJ*, 567)

It seems clear that Rawls «assumes a gap to exist between accepting the principles of justice and being motivated to act on them» (Barry 1995, 885). For bridging such a gap, Rawls appeals to the *Kantian interpretation* of justice as fairness¹⁰. According to this interpretation, allegiance and compliance with institutions shaped by principles which were reached in the original position are expression of individuals' moral nature. In Rawls's reading, Kant claims that «a person is acting autonomously when the principles of his action are chosen by him as the most adequate possible expression of his nature as a free and equal rational being» (*TJ*, 252). Consequently, acting from the principles chosen by the parties, precisely described as rational, free and equal moral persons, is tantamount of acting autonomously and expressing one's moral nature. As Rawls states, «the desire to act justly and the desire to express our nature as free moral persons turn out to specify what is practically speaking the same desire» (*TJ*, 572). Thus, starting from

¹⁰ See *TJ*, 251-257. Compare with Kant 1785.

the Kantian interpretation and assuming that «human beings have a desire to express their nature as free and equal moral persons» (TJ, 528), Rawls concludes that individuals are led to select and pursue plans of life consistent with both their moral nature and the principles of justice. Then, justice as fairness is stable in that it does not imply conflicts between principles of justice and individuals' conceptions of the good: a society shaped by justice as fairness is characterised by the *congruence* between the right and the good¹¹.

The justificatory apparatus deployed by Rawls in *A Theory of Justice* is complex, much more complex than this brief exposition might suggest. Nonetheless, this rough presentation allows to schematically outline the method employed by Rawls with reference to desirability and feasibility criteria. In *A Theory of Justice*, Rawls starts from desirability criteria, modelled through the original position and, on their basis, he specifies the set of acceptable and desirable conceptions of justice. The elements of such a set are then ranked according to their relative feasibility. Once the most feasible conception of justice among the desirable ones has been identified, it is tested against further feasibility constraints that are to be intended as further checks on the plausibility of the argument developed in the original position¹². At first, this might seem a reliable interpretation of Rawls's methodological approach, but it is still imprecise and in need of clarification. In particular, in order to adequately understand Rawls's appeal to stability considerations, it is necessary to point out that, despite appearances, Rawls is not concerned with properly *independent* feasibility constraints: Rawls's argument is elaborated within an *idealized* theoretical space and feasibility considerations are *immanent*¹³ to the boundaries of such a dimension.

In *A Theory of Justice*, it is possible to identify three different elements of idealisation. First, the original position is an «*idealized* situation» (TJ, 147, emphasis added): the parties are required to deliberate respecting formal constraints, without information they would normally dispose of and as if they were free from certain attitudes they are likely to develop under ordinary circumstances. These idealized conditions are meant to avoid contingent and morally irrelevant biases in the deliberation about justice. Second, individuals' deliberations concerning their good and their plans of life take place with «fully deliberative rationality, that is with full

¹¹ Actually, Rawls's argument for the congruence between the right and the good also appeals to the idea of *social union* and to the *Aristotelian Principle*. Taken together, these two ideas explain why just institutions may be seen as intrinsically valuable. Rawls upholds that a well-ordered society shaped by justice as fairness can be depicted as a social union, that is, as a polity characterized by the citizens having «shared final ends» and by their valuing «common institutions as good in themselves» (TJ, 522). Just institutions can be appreciated as intrinsically good since they allow each citizen to enjoy «the greater richness and diversity of the collective activity», thus permitting the «Aristotelian Principle to have its wider effect» (TJ, 571). The Aristotelian Principle stating, in fact, that «other things equal, human beings enjoy the exercise of their realized capacities (their innate or trained abilities), and this enjoyment increases the more the capacity is realized or the more its complexity» (TJ, 426).

¹² See Copp 1996, 193, fn. 15, and Hill 1994, 337.

¹³ I borrow this term from Habermas who, in *Reconciliation Through the Public Use of Reason* (Habermas 1995), employs the adjective *immanent* in order to describe the approach to stability displayed in *A Theory of Justice* and in order to contrast it with Rawls's later approach focused on the idea of an overlapping consensus.

awareness of the relevant facts and after a careful consideration of the consequences» (TJ, 408) and it is «assumed that there are no errors of calculation or reasoning, and that the facts are correctly assessed» (TJ, 417). It is evident that fully deliberative rationality too is a form of idealization, namely it represents an *ideal* standard of deliberation. Finally, Rawls develops his principles of justice, as well as the illustration of «a basic structure that satisfies these principles» (TJ, 195), under the assumption of *strict compliance*. That is, in arguing in favour of justice as fairness, Rawls assumes that everyone «act[s] justly and [...] do[es] his part in upholding just institutions» (TJ, 8). This is apparently unrealistic and, in fact, Rawls's reflection develops within *ideal theory*. In *A Theory of Justice*, Rawls appeals to ideal theory in order to offer a depiction of «the nature and aims of a *perfectly* just society», since this is, as he states, «the fundamental part of the theory of justice» (TJ, 9, emphasis added).

With the different levels of idealization in mind, let us get back to the relationship that ties desirability and feasibility. First, it is perfectly true that Rawls starts from desirability. Desirability considerations are embedded in the original position that constitutes the adequate perspective from which to elaborate and assess principles of justice. The formal constraints on the concept of right together with the veil of ignorance lay down the limits of the set of desirable options and they jointly constitute a test of acceptability for conceptions of justice. More precisely, they represent constraints of *absolute desirability*: only conceptions that conform to the requirements of generality, universality, publicity and finality, and that do not discriminate among individuals on the basis of contingent and morally irrelevant facts can be deemed desirable. It is true as well that considerations of feasibility (or of stability) play a fundamental role within the original position: they allow to rank different conceptions of justice and to adjudicate among them. In particular, considerations of stability are appealed to for defending the preferability of the two principles of justice as fairness over the principles of utilitarianism, to vindicate their *relative desirability*. Nonetheless, stability is defined as «a *desirable* feature of moral conceptions» (TJ, 455, emphasis added): stability is presented as a dimension of desirability. Therefore, on the one hand, desirability is prioritarian and it frames feasibility: feasibility criteria are applied once the set of desirable options has been identified. On the other hand, the relative desirability of a conception of justice depends on its stability, on its foreseeable capacity to be complied with by the parties. Yet, such a capacity is not assessed on the basis of individuals' *actual* motivations or dispositions. Justice as fairness's stability is initially evaluated from within the original positions in which references to empirical motivations are completely absent and in which the parties are unrealistically described as free from certain sentiments they would develop in ordinary circumstances.

Second, in verifying the stability of justice as fairness on the basis of a psychological account of individuals' moral development, Rawls assumes that just institutions are already in place and that they are generally complied with. In fact, although he defines the psychological laws underlying his account of moral development as «true, or approximately so» (TJ, 502), Rawls recognizes they are partially shaped by the particular kind of society they are associated with.

Therefore, the psychological laws Rawls envisages are not *fixed* and *independent* constraints of feasibility any conception of justice should confront. Rather, they constitute a plausible and not empirically falsified description of individuals' moral development, a description that is able to account for the rise of the sense of justice supporting justice as fairness. Accordingly, Rawls's psychological account is not only compatible with justice as fairness, but it is immanent in and dependent on justice as fairness itself: it is only under the circumstances brought about by the realisation of justice as fairness that Rawls's psychological account is significant.

Finally, since the reasoning concerning the congruence between the right and the good is framed in the idealized dimension of fully deliberative rationality, Rawls «does not require that moral principles be compatible with *given* preferences and conceptions of the good» (Freeman 2003b, 288, emphasis added). The argument from congruence, indeed, is not aimed at demonstrating that justice as fairness is realizable here and now. Rather, the congruence between the right and the good is intended to show that «principles of justice, derived on grounds independent of given preferences», can «be within the reach of human capacities and be compatible with a human good that affirms our nature» (Freeman 2003b). That is, the congruence between the right and the good is not barely implausible and unachievable insofar as it can be attained under the circumstances brought about by justice as fairness's public affirmation.

In *A Theory of Justice*, the argument from stability shows that justice as fairness is not completely impossible to be realized: it is compatible with a reliable understanding of human psychology and, consequently, it does not require attitudes that are out of human reach. Then, stability is to be interpreted as a confirmation of justice as fairness's internal robustness, as a value added to the plausibility and to the desirability of the two principles Rawls proposes: stability is a dimension of desirability. It seems possible to conclude that Rawls's approach in *A Theory of Justice* is utopian: desirability is assigned a prioritarian role. Rawls himself has come to consider his original approach as utopian. More precisely, according to Rawls, the approach put forward in *A Theory of Justice* is to be regarded as barrenly and uselessly utopian in that it relies on an idealized, and therefore practically unserviceable, account of individuals' motivations. This, together with a growing conviction about the necessity of attributing political philosophy with a proper and more immediate practical function has led Rawls to revise his argument supporting justice as fairness.

2. FROM A THEORY OF JUSTICE TO POLITICAL LIBERALISM: RAWLS'S NEW FOCUS ON STABILITY

With respect to desirability and feasibility, *Political Liberalism*'s internal structure seems to reflect the one of *A Theory of Justice*: the substantive content of desirability criteria is identified through the original position and, only once it is defined, its feasibility is assessed. Nonetheless, the status of the premises supporting the

original position is different: they are no more considered judgements, but *shared ideas* implicit in the public culture of democratic societies. Again differently from *A Theory of Justice*, in *Political Liberalism* justice as fairness's stability is evaluated not only considering its capacity to generate an appropriate sense of justice, but also with reference to its appropriateness in constituting the focus of an *overlapping consensus* among different and conflicting comprehensive doctrines endorsed by democratic citizens. In addition, in Rawls's later approach, justice as fairness assumes a contextual profile: it is no more intended for «a loosely specified 'us' perhaps for anybody» (O'Neill 2003, 349) – as it appears in *A Theory of Justice* – it is presented instead as a conception of justice suitable for a specific kind of societies and for a specific kind of persons – namely for liberal democratic societies and their citizens. Indeed, focusing on the essays that have prepared the ground for *Political Liberalism*¹⁴ and reading through the sparse notes you find here and there expressing Rawls's meta-theoretical position, it is clear that political philosophy is intended by Rawls to serve *contextual* aims.

Rawls asserts that «the aims of political philosophy depend on the society it addresses» (1987, 421). As he explains, «we are not trying to find a conception of justice suitable for all societies regardless of their particular social or historical circumstances» (1980, 305). Rather, as he continues, «we want to settle a fundamental disagreement over the just form of basic institutions within a democratic society under modern conditions» (1980, 305-306). Being context-dependent, political philosophy must be sensitive to the specific circumstances shaping the society it addresses. Rawls draws attention to five *general facts* that characterise liberal democratic societies and that political philosophy must acknowledge. First, the fact of *reasonable pluralism*¹⁵, that is, the fact that, under «the political and social conditions secured by basic rights and liberties of free institutions, a diversity of conflicting and irreconcilable – and what's more reasonable – comprehensive doctrines will come about and persist» (PL, 36). As Rawls specifies, «this is not a mere historical condition that may soon pass away; it is a permanent feature of the public culture of democracy» (PL, 36). Second, the fact of oppression, the fact that «a continuing shared understanding on one comprehensive religious, philosophical or moral doctrine can be maintained only by the oppressive use of state power» (PL, 37). The third fact is that «an enduring and secure democratic regime, one not divided into contending doctrinal confessions and hostile social classes, must be willingly and freely supported by at least a substantial majority of its politically active citizens» (PL, 38). Fourth fact, «the political culture of a reasonably stable democratic society normally contains, at least implicitly, certain fundamental intuitive ideas from which it is possible to work up a political conception of justice» (1989, 475). Rawls also signals a fifth fact that accounts for the emergence and endurance of reasonable pluralism: the burdens of judgement. He affirms that, due to the burdens of judgements, «we make many of our most important judgements subject to conditions which render

¹⁴ See in particular Rawls 1985, 1987, 1989.

¹⁵ Certainly influenced by Joshua Cohen (see Cohen 1993), Rawls distinguishes between the fact of pluralism and the fact of *reasonable* pluralism: see PL, 36-37.

it extremely unlikely that conscientious and fully reasonable persons, even after free discussion, can exercise their powers of reason so that all arrive at the same conclusion» (1989, 478).

The fact of reasonable pluralism is assigned a particular relevance. According to Rawls, the argument concerning stability as conceived in *A Theory of Justice* is based on an unrealistic, «indeed utopian» (PL, 39), account: a well-ordered society shaped by justice as fairness implies that «all its citizens endorse this conception on the basis of [...] a comprehensive philosophical doctrine» (PL, xvi) and this is not plausible given the fact of reasonable pluralism. More precisely, Rawls states that, since it requires the endorsement of a single comprehensive doctrine, the «account of stability in part III of *Theory* is not consistent with the view as a whole» (PL, xvi-xvii). In the introduction to the Paperback Edition of *Political Liberalism*, Rawls better clarifies the source of such an inconsistency:

Since the principles of justice as fairness in *Theory* require a constitutional democratic regime, and since the fact of reasonable pluralism is the long-term outcome of a society's culture in the context of these free institutions, the argument in *Theory* relies on a premise the realization of which its principles of justice rule out. (PL, xl)

Consequently, in *Political Liberalism*, the question of stability is reconceived in order to take into account the fact of reasonable pluralism: a conception of justice qualifies as stable if it can be the focus of an overlapping consensus among reasonable comprehensive doctrines, that is if it can be considered as acceptable by individuals endorsing different reasonable doctrines.

Beyond his new contextual focus, Rawls also emphasises that political philosophy must undertake a properly practical and political function. According to Rawls, the *practical* task of political philosophy arises «from divisive political conflict and the need to settle the problem of order» (2001, 1) and it consists in focusing «on deeply disputed questions and to see whether, despite appearances, some underlying basis of philosophical and moral agreement can be uncovered» (2001, 2). In *Justice as Fairness. A Restatement* (Rawls 2001), Rawls envisages two other roles for political philosophy, namely *orientation* and *reconciliation*, which are both strongly linked to the first practical one¹⁶. The orientative function of political philosophy is based on the idea that «it belongs to reason and reflection (both theoretical and practical) to orient us in the conceptual space, say, of all possible ends, individual and associational, political and social» (2001, 3). Under this respect, the role of philosophy, «as a work of reason», consists in «specifying principles to identify reasonable and rational ends of those various kinds» and «in showing how those ends can cohere within a well-articulated conception of a just and reasonable society» (2001, 3). The role of orientation is related to the practical one in that

¹⁶ The title of the section in question is «Four Roles of Political Philosophy». In this text and in Rawls 2007, Rawls envisages a further task for political philosophy, namely the task of «probing the limits of practicable political possibilities» (Rawls 2001, 4), thus assuming a *realistically utopian* profile. This task will be the focus of section 5.

«such a conception may offer a unified framework within which proposed answers to divisive questions can be made consistent» (2001, 3). The task of reconciliation, instead, requires political philosophy «to calm our frustration and rage against our society and its history by showing us the way in which our institutions, when properly understood from a philosophical point of view are rational, and developed over time as they did to attain their present rational form» (2001, 3). By facilitating reconciliation, political philosophy can contribute to settle the practical problem of order: by showing the rational character of our society, political philosophy leads us to «accept and affirm our social world positively, not merely to be resigned to it» (2001, 3).

Although Rawls distinguishes among these three tasks, they seem to be more than strongly interlinked: they seem to be different specifications for one and single function, that of reconciliation¹⁷. Indeed, the pursuit of reconciliatory aims on the part of political philosophy may come into different forms: political philosophy may aim at reconciling citizens who endorse different comprehensive doctrines, or single individuals with themselves by giving coherence to their different ends, or individuals with their own social world by showing its inner rationality. It seems plausible to state that Rawls's emphasis on reconciliation arises from his acknowledgement that the «course of democratic thought over the past two centuries or so makes plain that there is no agreement on the way basic institutions of a constitutional democracy should be arranged if they are to specify and secure the basic rights and liberties of citizens and answer to the claim of democratic equality» (1985, 391). In order to find a basis of public agreement, as Rawls writes, «we must find a new way of organizing familiar ideas and principles into a conception of political justice so that the claims in conflict, as previously understood, are seen in another light» (1985, 393). And the practical task Rawls ascribes to political philosophy consists precisely in articulating such a conception of justice, a conception of justice that can be the focus of an overlapping consensus.

In the light of a similar purpose, political philosophy must abandon metaphysical or epistemological claims. As Rawls states, «philosophy as the search of truth about an independent metaphysical and moral world cannot [...] provide a workable and shared basis for a political conception of justice in a democratic society» (1985, 395). Indeed, a conception of justice based on metaphysical assumptions or advancing epistemological claims is likely to prove unacceptable

¹⁷ As it appears clear in the *Preface* to his *Philosophy of Right* (Hegel 1821), reconciliation is the aim Hegel assigns to philosophy in general and to political philosophy in particular (see Hardimon 1994; see also Rawls 2000, 329). In discussing the reconciliatory function of political philosophy, Rawls makes explicitly references to Hegel. Yet, the kind of argument Hegel relies on is unavailable to Rawls: since Rawls aims, as specified below, at leaving any philosophical and controversial assumption aside, he cannot resort to any philosophy of history. The question is then whether it makes any sense to endorse a Hegelian perspective without endorsing, at the same time, a certain philosophy of history able to account for why our social world is to be seen as rational and acceptable instead of being deemed irrational and unacceptable. This question – to which I will come back at the end of the paper – is to be combined with a reflection on whether Rawls can endorse Kant's idea of a *reasonable faith*, as he does, without subscribing to certain Kantian claims he seems bound to discard due to their controversiality.

for individuals endorsing different comprehensive doctrines and, therefore, it is likely to qualify as practically irrelevant. Accordingly, political philosophy must apply the «principle of toleration to philosophy itself» (PL, 10) and refrain from claiming truth for its own assertions. As Rawls states, since «the aim of justice as fairness as a political conception is practical, and not metaphysical or epistemological», it «presents itself not as a conception that is true, but one that can serve as a basis of informed and willing political agreement between citizens» (1985, 394). In addition, political philosophy must «stay on the surface, philosophically speaking» and it must adopt what Rawls calls the «method of avoidance» (1985, 395). That is, political philosophy must step aside from «philosophical controversies whenever possible» and it should avoid «philosophy's longstanding problems» (1985, 395). In fact, while articulating a conception of justice «may involve settling theoretical difficulties, the practical social task is primary» (1980, 306). It is only by leaving apart metaphysical and epistemological claims that political philosophy can be relevant on the practical level and aspire to elaborate a workable, a feasible conception of justice able to be accepted by individuals who endorse conflicting comprehensive views. That is, it is only by leaving philosophy aside that such an aspiration can be fulfilled.

In order for a conception of justice addressing democratic societies to be practical, or as Rawls would say to «fall under the art of the possible» (1989, 486), it must be *political*. In turn, a *political* conception of justice must meet three conditions. The first condition is a matter of scope: a political conception of justice is a «moral conception worked out for a specific kind of subject, namely, for political social and economic institutions» (PL, 11). The second condition concerns the mode of presentation: «a political conception is presented as a freestanding view» (PL, 12). That is, a political conception of justice must be «capable of coherent statement that is independent of any particular comprehensive philosophy» (Ackerman 1994, 366). The third condition relates to the appropriate content for a political conception of justice, which must be «expressed in terms of certain fundamental ideas seen as implicit in the public political culture of a democratic society» (PL, 13). Indeed, Rawls maintains that «in a democratic society there is a tradition of democratic thought, the content of which is at least familiar and intelligible to the educated common sense of citizens generally» (PL, 14). Such content is specified by Rawls into three ideas: the idea of society as a fair system of cooperation, the idea of citizens as free and equal persons and the idea of a well-ordered society. These ideas are captured and modelled in the original position: as Rawls writes, «as a device of representation the idea of the original position serves as a means of public reflection and self-clarification» (PL, 26). The role the original position plays in *Political Liberalism* is similar to the one it is assigned with in *A Theory of Justice*, but between Rawls's earlier and latter accounts there are differences deserving to be pointed out.

First, the elements modelled in the original position in *A Theory of Justice* are considered judgements, while, in *Political Liberalism*, they are the shared ideas embedded in the public culture of democratic societies. This is consistent with and justified by Rawls's more contextual focus. Yet, it is worth emphasising that, as a

consequence of Rawls contextual shift, the desirability of the principles of justice chosen in the original position is reduced in scope: justice as fairness's desirability is no more universalistic, but it can be vindicated only with reference to a specific context. Therefore, it comes with no surprise that the parties are no more asked «to choose principles that hold unconditionally whatever the circumstances» (TJ, 125) and that the original position is no more intended to be a perspective that allows us to see «our place in society [...] *sub species aeternitatis*» (TJ, 587). Rawls's new account of the original position, a more contextual account¹⁸, suggests to introduce a distinction between *universal* and *contextual* desirability, a distinction between desirability criteria that are valid independently from the context and desirability criteria that hold only under specific circumstances. And the kind of desirability Rawls pursues is clearly the latter.

Second, while in *A Theory of Justice*, Rawls's focus is mainly on the *appropriateness* of the original position as a perspective from which to express desirability judgements, in *Political Liberalism*, Rawls stresses its suitability as a *device of representation*. It seems that, whereas in *A Theory of Justice* the conditions of fairness modelled by the original position are intrinsically valuable, in *Political Liberalism*, these very conditions are accounted for because they happen to be, as a matter of fact, the circumstances characterising liberal democracies. The first account is, of course, problematic and question-begging: if the original position is the correct point of view in order to express value judgements, what justifies the value attached to its very conditions? The second account is problematic too, but for different reasons. The conditions characterising the original position in *Political Liberalism* are supposed to be derived from factual considerations concerning the public culture of democratic societies, but they are far from being neutral and uncontroversial¹⁹.

Third, while in *A Theory of Justice* the question of stability is preliminarily introduced in the original position and it affects the parties' reasoning, in *Political Liberalism* the parties do not address such a question: stability enters the argument only in a later stage. It is the requirement of freestandingness that leads Rawls to postpone stability considerations:

Justice as fairness is best presented in two stages. In the first stage it is worked out as a freestanding political (but of course moral) conception for the basic structure of society. Only with this done and its content – its principles of justice and ideals – provisionally on hand do we take up, in the second stage, the problem whether justice as fairness is sufficiently stable. (PL, 140-141)

¹⁸ The contextual justification of justice as fairness implies that Rawls cannot address individuals «who are not fellow citizens, who are excluded or marginalised within a polity, or who do not accept democracy and its constraints» (O'Neill 2003, 353). In addition, Rawls's contextual focus leads him to *freeze* and *reify* the context, thus ruling out the very possibility of detecting negative aspects of the context he refers to (see Winghenbach 1999).

¹⁹ I will come back to this point at some length in the following section.

The rationale for presenting justice as fairness into two stages is that, otherwise, justice as fairness would be «political in the wrong way» (1989, 491). As Rawls clarifies, although political, a conception of justice «must not be political in the sense of merely specifying a workable compromise between known and existing interests, nor political in looking to the particular comprehensive doctrines known to exist in society and in then being tailored to gain their allegiance» (1989, 491). If justice as fairness were arrived at in like manner, the charge Jürgen Habermas raises against Rawls – namely the charge of failing to draw a sharp «distinction between acceptability and acceptance» (Habermas 1995, 122) – would be undeniable and the overlapping consensus would simply be a prudential compromise. In order to address such a critique, Rawls distinguishes between two different kinds of consensus:

One idea of consensus comes from everyday politics where the task of the politician is to find agreement. Looking to various existing interests and claims, the politician tries to put together a coalition or policy that all or sufficient number can support to gain majority. This idea of consensus is the idea of an overlap that is already present or latent, and could be articulated by the politician's skill in bringing together existing interests the politician knows intimately. The very different idea of consensus in political liberalism – the idea I call *reasonable overlapping consensus* – is that the political conception of justice is worked out first as a freestanding view that can be justified *pro tanto* without looking to, or trying to fit, or even knowing what are, the existing comprehensive doctrines. (1995, 389)

This characterization, intended to guarantee justice as fairness, is not developed with the aim of being accepted as a political compromise among different existing doctrines. Therefore, an overlapping consensus is not a mere *modus vivendi*: it entails an agreement on a moral conception whose acceptance is grounded on individuals' moral convictions²⁰. With respect to a *modus vivendi*, an overlapping consensus is more stable since it does not depend «on a balance of relative forces» (1987, 433): given that the overlapping consensus is based on moral grounds, «those who affirm the various views supporting the political conception will not withdraw their support of it should the relative strength of their view in society increase and eventually become dominant» (1987, 432). In addition, Rawls clarifies that the idea of an overlapping consensus is different also from that of a constitutional consensus: an overlapping consensus is deeper and broader than a constitutional agreement²¹. On the one hand, constitutional principles are «accepted simply as principles and not as grounded in certain ideas of society and person of a political conception, much less in a shared public conception» and, therefore, a constitutional consensus «is not deep» (*PL*, 158). On the other hand, a constitutional consensus «is narrow in scope, not including the basic structure but only the political procedures of democratic government» (*PL*, 159).

²⁰ See Rawls 1987, 432.

²¹ Rawls introduces the distinction between overlapping consensus and constitutional consensus in order to counter Kurt Baier's critiques: see Baier 1989.

Rawls also engages in demonstrating that his idea of an overlapping consensus is not utopian by showing that there are «sufficient political, social and psychological forces either to bring about an overlapping consensus (when one does not exist), or to render one stable (should one exist)» (PL, 158). To this end, Rawls relies on an account of «how an initial acquiescence in a liberal conception of justice as a mere *modus vivendi* could change over time first into a constitutional consensus and then into an overlapping consensus» (PL, 168)²². Similarly to the plausibility of the three psychological laws, the plausibility of such an account is meant to illustrate that achieving and securing an overlapping consensus is not barely impossible. The overlapping consensus is, thus, a feasible and not utopian – in the sense of not practically irrelevant – goal political philosophy can pursue.

This is, once again, a rough and schematic account of Rawls's argument in *Political Liberalism*, but it seems to confirm my initial claim according to which Rawls's argument displays, approximately, the same structure in *A Theory of Justice* and in his later works. Yet, Rawls's later approach is framed in a totally different perspective: political philosophy is meant to endorse a properly practical and political function and, accordingly, the features a desirable conception of justice must possess are identified on the basis of the general facts characterizing liberal democratic societies. That is, the general facts Rawls identifies as relevant in describing liberal democratic societies, the fact of reasonable pluralism in particular, shape the profile for a suitable conception of justice: it is on the basis of these facts that Rawls calls for the necessity of a *political* conception of justice able to secure stability. To be clearer, the standards a satisfying conception of justice must meet are determined by factual considerations that constitute the feasibility constraints political philosophy must accommodate in order to be practically relevant. It seems then possible to conclude that, in Rawls's revised approach, feasibility frames desirability: principles that cannot accommodate the fact of reasonable pluralism, that cannot be the focus of an overlapping consensus are to be rejected, no matter how desirable they are. Admittedly, this characterization does not help in assessing whether Rawls way of combining desirability and feasibility is satisfactory. The following section offers some clues with regard to this by focusing on the benefits and costs of Rawls's later methodology.

3. THE AMBIGUITIES OF RAWLS'S *POLITICAL* APPROACH

According to the conclusions of the previous section, in Rawls's later approach feasibility frames desirability. Since political philosophy should serve practical and contextual aims, the form for an appropriate conception of justice is dictated by the relevant facts characterising the society it addresses. In particular, the fact of reasonable pluralism requires political philosophy to elaborate a conception of justice that can be recognised as acceptable by all reasonable democratic citizens,

²² This account is built up relying on the exemplar case of the European wars of religion: see PL, 145-146.

no matter which reasonable comprehensive doctrine they endorse. Therefore, political philosophy must avoid metaphysical and epistemological claims and it must rely on premises drawn from the public culture of democratic societies.

Some merits should be recognised to Rawls's practical and contextual approach. If assessed from *within*, Rawls's theory seems a *robust* theory. Rawls identifies the shared understandings characterising the public culture of democratic societies and he assumes these understandings as his starting point. He models the original position on the basis of such assumptions and he derives a conception of justice that, conceding his derivation is sound, cannot but be accepted by those who share the premises. Furthermore, the content of the premises as well as their source is made explicit, more explicit than in *A Theory of Justice*. Yet, shifting the perspective from within Rawls's argument and adopting an external point of view, Rawls's approach proves less satisfying: although internally robust – or exactly because internally robust – Rawls's political liberalism seems to lack an overall justification. More precisely, as the following pages will argue, Rawls's project suffers from this lack because of the *character* of the premises he relies on.

Rawls insists that his premises – the idea of individuals as free and equal persons and that of society as a fair system of cooperation – are *political* and not metaphysical. By a similar distinction, Rawls seems to suggest, first, that his premises are contextual – they are suitable *only* for democratic societies – and restricted in scope – they apply *only* to the domain of the political; second, that they do not rest upon theses drawn from comprehensive doctrines or from other parts of philosophy; and, third, that they are not presented as true, but simply as intuitive ideas embedded in the public culture of democratic societies²³. Combining these features, it is safe to conclude that the organizing ideas Rawls relies on are political and not metaphysical in that they are not controversial. In particular, Rawls seems to attach to his premises a *factual* character: they are ideas that every reasonable democratic citizen would recognize, on due reflection, as the key elements characterising, as a matter of fact, democratic societies. As Joseph Raz notices, Rawls's justification «starts not with general moral truths but with the *givens* of our common culture, which it takes as *facts*, irrespective of their validity or truth» (Raz 1990, 8, emphasis added). As Raz continues, Rawls's «justification starts with the fact that certain beliefs form the common currency of our public culture» and he is not concerned «with their justification or with its absence» (Raz 1990, 8). Therefore, as Raz states, Rawls's project relies on *shallow foundations*.

Rawls's intention to rest on factual and, therefore, uncontroversial premises implies a twofold difficulty. First, the existence of uncontroversial facts and the possibility of ascertaining them are controversial. Since it is difficult to uphold a sharp dichotomy between facts and values, it seems implausible to rely on assumptions that are *barely* factual: they are likely to be informed, at least partially, by value considerations. The problem is that Rawls's method would result undermined should it be impossible to find out widely accepted and

²³ For interesting insights on Rawls's distinction between *political* and *metaphysical*, see Neal 1994.

uncontroversial premises. In fact, it seems that Rawls's shared ideas are to be interpreted as *points of agreement*: they can be depicted as being less controversial than principles of justice are. While Rawls assumes as a matter of fact the existence of common understandings concerning the public culture of democracy, he emphasises that such a culture is characterised by a deep disagreement concerning principles of justice, a disagreement originated by the unresolved tension between the so-called liberties of the moderns and liberties of the ancients²⁴. Yet, doubts may be cast against the uncontroversiality of the ideas Rawls refers to and, therefore, against their appropriateness in undertaking the function he envisages for them.

On the one hand, it may be claimed, in liberal democratic culture, it is impossible to identify a *single* set of shared ideas able to uncontroversially capture all the relevant understandings concerning such a culture. On the other hand, even granting there is one and single set of relevant understandings constituted by the ideas Rawls singles out, these very ideas make reference to *thin concepts* that need to be further developed and articulated in order to serve as a basis for elaborating a conception of justice. Yet, it seems safe to assert that freedom and equality, for instance, are *essentially contested concepts* and their use as well as their interpretation is unavoidably controversial. More precisely, freedom and equality are controversial both from a theoretical or philosophical point of view and in the light of their political usage. As Jean Hampton points out, it is «striking implausible to claim not merely that freedom and equality are accepted features of our democratic regimes, but also that the *Rawlsian conceptions* of these ideas are commonly accepted in these regimes» (Hampton 1994, 198). Indeed, Rawls «does not merely embrace the concepts of freedom and equality, he also interprets and develops these concepts into particular conceptions» (Hampton 1994, 198) that are likely to be contested and unlikely to be widely accepted.

The first difficulty – related to the dubious uncontroversiality of Rawls's premises – is far less troubling than the second one. Even conceding the ideas Rawls selects as starting points display the factual and uncontroversial character they need to possess for assuring the success of his project, it is necessary to consider whether it is possible to derive normative conclusions from exclusively factual assumptions and whether normative principles can be justified solely on the basis of factual premises²⁵. Rawls avoids directly discussing similar problems. Nonetheless, it seems necessary to emphasise that it is at least debatable that political philosophy should start from factual premises and that whether political philosophy should endorse a similar strategy depends on the kind of practical function it pursues. Indeed, political philosophy may rely on premises that are not presented as factual: it may rest on assumptions that can be argued for in a rational way, on assumptions individuals can accept on rational grounds. Certainly, this way of proceeding only insures the rational tenancy of a theory, but it offers little

²⁴ See *PL*, 5.

²⁵ The implausibility of justifying normative principles solely on the basis of facts has been recently reaffirmed by Gerald Cohen, with the precise intent of countering the Rawlsian way of framing the concept of justice. See Cohen 2003 and 2008.

guarantees from the perspective of feasibility: nothing assures that empirical individuals would accept the premises and would be moved by the solutions arrived at in like manner. This explains why Rawls refrains from adopting a similar approach: the practical and reconciliatory function Rawls envisages for political philosophy forces him to rely on premises that are *de facto* accepted by democratic citizens as uncontroversial. Yet, Rawls's methodology turns out to be self-defeating: his programmatic attempt to avoid controversial claims leads him to the impossibility of convincingly justifying his overall project and, what is more, to appropriately vindicate the desirability of the proposals he advances.

The shortfalls just mentioned are due to Rawls's commitment to the method of avoidance and to his *epistemic abstinence*²⁶. According to Rawls, since its aims are eminently practical and not epistemological, political philosophy must not be concerned with advancing a *true* conception of justice: it must simply elaborate a workable conception of justice that can be the focus of an overlapping consensus and that, as such, is able to ensure stability to democratic societies. With regard to this, it is worth asking, as Raz does:

Why should philosophy contribute to these goals rather than to others? Presumably because they are worthwhile goals. So it would appear that while the goal of political philosophy is purely practical – while it is not concerned to establish any evaluative truths – it accepts some such truths as the presuppositions which make its enterprise intelligible. It recognizes that social unity and stability based on a consensus – that is, achieved without excessive resort to force – are valuable goals of sufficient importance to make them and them alone the foundations of a theory of justice for our societies. (1990, 14)

Following Raz's reasoning, if Rawls's theory is to be considered as a satisfactory theory of justice, the goal of achieving stability, of enlightening the possibility for an overlapping consensus should be presented as valuable in itself. The question is, then, whether Rawls can advocate such a value for stability while remaining faithful to his commitment to avoid controversial claims. Unfortunately Rawls cannot. Even if Rawls strongly emphasises that the overlapping consensus he appeals to is different from a mere *modus vivendi* and «even if the ideas in the overlapping consensus are believed by the citizenry to be right as opposed to merely expedient», as Hampton emphasises, «Rawls can offer only Hobbes-style expediency arguments for the generation of the consensus itself» (Hampton 1989, 807). That is, since he is committed to avoid controversial claims, Rawls can justify his own attempt to construct a conception of justice that aims at being the focus of an overlapping consensus only with reference to the goal of promoting stability. Yet – and this is the fallacy – Rawls's search for stability is justified only for stability's own sake. Let us see why.

It may certainly be argued that the stability of liberal and democratic societies is instrumentally desirable: since liberalism and democracy are valuable, their

²⁶ This expression is borrowed from Raz 1990.

endurance ought to be safeguarded. Yet, Rawls's method starts from *within* liberal democratic societies and it leads to principles of justice that are valid only internally to such societies, therefore it does not allow to scrutinize the desirability of liberal democratic societies themselves. Consequently, Rawls's search for stability cannot be justified in the light of elaborating a stable conception of justice aimed at protecting valuable institutions and practices. Nonetheless, the implication of ascribing priority to stability seems to cut much deeper: it seems that the priority of feasibility prevents the desirability of liberalism from being appropriately vindicated. In fact, the desirability of liberalism seems to be asserted solely on the basis that it is the only *feasible* conception for democratic societies: given the fact of reasonable pluralism and the fact of oppression, «any *workable* conception of political justice for a democratic regime must [...] be in an appropriate sense liberal» (1987, 426, emphasis added). It is necessary to stress that the impossibility of properly justifying his search for stability and of convincingly vindicating the desirability of liberalism is connected to Rawls's attempt to cope with the fact of reasonable pluralism. Indeed, reasonable pluralism is to be intended as a feasibility constraint that imposes political philosophy to eschew any controversial claim, thus rejecting any appeal to independent normative principles²⁷.

Although feasibility seems to be the criterion moulding Rawls's overall program, I am inclined to think that Rawls would not be disposed to admit that his project is completely constrained by feasibility considerations. Neither would he be disposed to retreat from his liberal commitments. Rather, Rawls looks for a *middle ground* between a *pragmatic* defence of liberalism, «which abjures reference to moral ideals», and a *principled* assertion of it that «relies on moral ideas which are irremediably controversial» (Brighouse 1994, 319). Accordingly, Rawls's political liberalism «steers a course between the Hobbesian strand in liberalism – liberalism as a *modus vivendi* secured by a convergence of self- and group-interests as coordinated and balanced by well-designed constitutional arrangements – and a liberalism founded on a comprehensive moral doctrine such that of Kant or Mill» (1987, 446). Rawls explains that the rationale form appealing to such a *third way* is that «the former cannot secure an enduring social unity, the latter cannot gain sufficient agreement» (1987, 446): the former would result unstable, the latter would show itself to be inconsistent with the fact of reasonable pluralism. To be clear, Rawls's third way does not stem out as an alternative form of liberalism that

²⁷ Not any form of political liberalism is bound to avoid reference to prioritarian and normative principles. For instance, Charles Larmore asserts that political liberalism «forms a freestanding conception in regard to comprehensive moral visions of the good life, but it cannot coherently claim to be freestanding with respect to morality altogether. In particular, we would be wrong to suppose that the moral principle of respect for persons has the political significance it does because reasonable people share a commitment to it. On the contrary, the idea of respect is what directs us to seek the principles of our political life in the area of reasonable agreement. Respect for persons lies at the heart of political liberalism, not because looking for common ground we find it there, but because it is what impels us to look for common ground at all» (Larmore 1999, 608). Larmore's political liberalism, therefore, is not motivated by instrumental considerations or by the acknowledgement that only a liberal political conception of justice is able to endure and win acceptance in pluralistic and democratic society. Rather, Larmore relies on a prioritarian and intrinsically normative principle – respect for persons – that shapes his overall project. Such a strategy is not available to Rawls.

is different in kind with respect to both *modus-vivendi* and comprehensive liberalism. Rawls's solution is a *middle-ground* solution that relies on a combination of elements proper of the two just mentioned forms of liberalism: it is a combinatory solution, combinatory as his realistic utopianism is.

The attempt to find a middle ground leads Rawls to endorse an ambiguous position with respect to the status of the general facts and of the shared ideas he refers to. As for ideas, it has already been argued that their character is uncertain and that they are hardly uncontroversial. Rawls's position is equally ambiguous with respect to the facts. While introducing the five general facts, Rawls asserts that «any political conception of justice presupposes a view of the political and social world, and recognizes certain general facts of political sociology and human psychology» (1989, 474). Yet, Rawls's general facts are not *empirical* facts. Considering, as an instance, the fact of pluralism, it is clear that it cannot be intended as such. Rawls does not refer to the mere fact of pluralism, since he adds a clearly evaluative qualification: it is *reasonable* pluralism he is concerned with. Therefore, Rawls is not advancing a simple factual claim. The presence of reasonable pluralism «is not simply a historical or sociological claim»: as Krasnoff correctly notices, «we must understand the diversity of moral and religious doctrines not as a failure of rationality, but as a consequence of rationality in its diverse exercises» and «this is a very strong claim» (Krasnoff 1998, 277). Indeed, Rawls present reasonable pluralism as itself valuable. Thus, the fact of reasonable pluralism, which is the most relevant among the general facts Rawls accounts for, should not be considered as a bare fact, but, let us say, as a *qualified* fact.

In Rawls's political approach, feasibility frames desirability. Feasibility constraints are derived on the basis of considerations that are presented as factual. Yet, on a more careful consideration, these same feasibility constraints show a non-properly factual character, or at least, a twofold character: they are intended, at once, as feasibility constraints and as desirability requirements. Therefore, on the one hand, Rawls's later methodology implies the impossibility of asserting genuine claims of desirability, as shown with respect to the case of liberalism. On the other hand, Rawls's attempt at finding a middle-ground position leads to ambiguities that are difficult to get on with. This suggests that the very attempt to find a middle ground between realism and utopianism, which entails the refusal to rank desirability and feasibility, is unreliable. That is why it is necessary to address more directly Rawls's idea of a realistic utopia, in which similar ambiguities are possibly more apparent.

4. THE SHORTFALLS OF RAWLS'S REALISTIC UTOPIANISM

Rawls's search for a middle-ground conception of justice comes from the aspiration to grant political philosophy a major practical relevance – a major feasibility – without excessively downgrading the normative appeal – the desirability – of its claims. The idea of a *realistic utopia*, is the outcome of such an attempt at reconciling desirability and feasibility and at finding a synthesis between

them. The critical literature concerning *The Law of Peoples*, where the idea of a realistic utopia is most fully elaborated, qualifies the principles Rawls proposes for the international dimension as neither realistic nor utopian. As Andrew Kuper puts it, «Rawls is neither sufficiently utopian nor sufficiently realistic» (Kuper 2000, 659). From the perspective of realism, Rawls's understanding of the international dimension may be criticised for its reliance on the Westphalian model of international relations, which is regarded as no longer suitable for the description of an increasingly interdependent scenario²⁸. From the point of view of utopianism, instead, it may be claimed that the content of Rawls's proposed law of peoples is hardly different from that of the conventional precepts that have been regulating interactions among sovereign states over the last centuries and that, therefore, there is nothing utopian in it. Thus, Rawls seems unable to meet his own standards for a satisfactory normative model, for a realistically utopian model.

While dissatisfaction with the content of Rawls's law of peoples is plausible, the analysis here proposed will not follow this line of criticism. For evaluating the potentiality and tenability of Rawls's approach, it is more interesting to investigate upon the very concept of a realistic utopia by singling out both the presuppositions that render it meaningful and its implications. To these ends, Rawls's failure is scarcely illuminating: it might simply be a *contingent* failure due, maybe, to Rawls's incapacity of identifying realistic conditions or of elaborating utopian solutions, or both. Nonetheless, a more in-depth examination shows that, if considered from Rawls's own perspective, *The Law of Peoples* is not a failure: it perfectly fits Rawls's overall project for political philosophy. Yet, it is precisely this very project that appears unsatisfying since it does not allow political philosophy to meet its normative commitments. As this section argues, a realistically utopian political philosophy, as envisaged by Rawls, is up to vindicate neither the desirability nor the feasibility of its principles and models.

The enlightenment of realistic utopias is one of the tasks, the fourth task, Rawls envisages for political philosophy in the opening section of *Justice as Fairness. A Restatement*. He asserts that «we see political philosophy as realistically utopian» when it probes «the limits of practicable political possibility» (2001, 4). Political philosophy is *realistic* in that it is committed to remain within the boundaries of the feasible set and, at the same time, it is *utopian* in that it is able to recommend options that are not immediately realizable. As Rawls correctly points out, «there is a question about how the limits of the practicable are discerned» and he states that «the limits of the possible are not given by the actual, for we can to a greater or lesser extent change political and social institutions, and much else» (2001, 5). Accordingly, Rawls seems to consider the boundaries of the sphere of possibility as not determined by the actual, by actual arrangements and institutions. Yet, this impression is disconfirmed by further investigation on Rawls's position.

In *The Law of Peoples*, the idea of a realistic utopia plays a crucial role: Rawls affirms «I begin and end with the idea of a realistic utopia» (*LP*, 6). Indeed, at the very

²⁸ See, for instance, Buchanan 2000; Hurrell 2001; O'Neill 2001.

beginning of his book, Rawls spells out the conditions for a normative model to constitute a realistic utopia and, in the closing section, he assesses whether his own model qualifies as a realistic utopia. Rawls lists seven pairs of conditions for a realistic utopia, each condition being specified for a just liberal democratic society, on the one side, and for a society of well-ordered peoples, on the other. Accordingly, the idea of a realistic utopia not only guides Rawls's elaboration of the law of peoples, but it also may offer a new key for reading *Political Liberalism*.

The first two conditions put forward by Rawls in *The Law of Peoples* express requirements of realism. With reference to the domestic case, Rawls writes:

There are two necessary conditions for a liberal conception of justice to be *realistic*. The first is that it must rely on actual laws of nature and achieve the kind of stability those laws allow, that is stability for the right reasons. It takes people as they are (by the laws of nature) and constitutional and civil laws as they might be, that is as they would be in a reasonably just and well-ordered democratic society. [...] The second condition for a liberal political conception of justice to be realistic is that its first principles and precepts be workable and applicable to ongoing political and social arrangements. (*LP*, 12-13)

The two parallel conditions for the international dimension are as follows:

The reasonably just Society of well-ordered Peoples is *realistic* in the same way as a liberal or decent domestic society. Here again we view peoples as they are (as organized within a reasonably just domestic society) and the Law of Peoples as it might be, that is, how it would be in a reasonably just Society of just and decent Peoples. [...] The Law of Peoples is also realistic in a second way: it is workable and may be applicable to ongoing cooperative political arrangements and relations between peoples. (*LP*, 17)

Different levels of realism are involved in these two pairs of conditions. First, a conception of justice must not imply the violation of natural laws. This is a very minimal requirement of feasibility that allows to distinguish what is simply *logically* possible from what is *practically* possible, what is possible since it does not entail the violation of logical laws and what is possible to be realized since its enactment does not imply the violation of physical constraints. Second, a conception of justice must be able to achieve *stability for the right reasons*. Let us simply note that, since he relies on a *qualified* notion of stability – stability for the right reasons – Rawls seems to introduce normative considerations in the conditions presented as genuine requirements of realism. This already qualifies Rawls's project as implying certain inconsistencies, but, in order to explain why Rawls proceeds in this manner, it is necessary to highlight some further aspects of the conditions under study.

As far as the formulation of the first condition is concerned, Rawls explicitly relies on Rousseau's statement according to which he intends to take «men as they are and laws as they can be» (Rousseau 1762, 41). Rawls specifies that he assumes Rousseau's «phrase "men as they are" refers to persons' moral and psychological

natures and how the nature works within a framework of political and social institutions; and that his phrase “laws as they might be” refers to laws as they should, or ought, to be» (*LP*, 7). Therefore, Rawls draws a distinction between the individual dimension and the institutional sphere and he seems to endorse realism with respect to the former. In order for political philosophy to be realistic, it must keep constant the individual dimension, while it is on the institutional dimension that political philosophy may normatively intervene.

Rawls is not particularly clear about the individual dimension. He seems to suggest that realism requires political philosophy to refrain from asking individuals attitudes or capacities that, given their physical and psychological endowments, they do not possess or cannot attain. Some further clues for understanding what Rawls has in mind are offered by his appeal to the *subjective circumstances of justice*²⁹ that are connected to the fact that «persons and associations have contrary conceptions of the good as well as of how to realize them, and these differences set them at odds, and lead them to make conflicting claims on their institutions» (1980, 323). As Onora O’Neill notices, from *Kantian Constructivism in Moral Theory* on, «Rawls has always seen ethical pluralism as (in part) constitutive of modern circumstances of justice» (O’Neill 1988, 717). Therefore, *individuals as they are* seems to hint not only at individuals as they *naturally* are – as characterised by their natural endowments – but also at individuals as they are under modern conditions. Thus, pluralism is, once again, the relevant constraint of feasibility. In addition, Rawls usually refrains from drawing a sharp distinction between the individual dimension and the institutional sphere. Indeed, the individual dimension is not completely detached from the institutional one since individuals’ motivations, attitudes and even plans of life are quite substantively affected by the framework of practices and institutions they inhabit, at least in the sense that individuals are required to shape their plans of life by taking into account their compatibility with principles of justice³⁰. Consequently, it seems possible to point out some confusion between descriptive and normative plans, between requirements of realism and normative claims: within the limits of natural endowments and the fact of pluralism, how individuals are depends on the practices and institutions they live in; in turn, practices and institutions are not given and they are to be modelled by normative principles.

Looking at the first condition of realism for the society of peoples, confusion seems possibly greater. Rawls affirms that, in order to be realistic, the principles of the law of peoples should take «peoples as they are». Yet, peoples’ distinguishing features, as envisaged by Rawls, are hardly realistic: they do not rely on empirical observation. Rather, they are highly idealized. In Rawls’s definition, differently from states, peoples have «a reasonably just constitutional democratic government that serves their fundamental interests» (*LP*, 23), their citizens are «united by common sympathies and a desire to be under the same democratic government»

²⁹ For Rawls’s account of the circumstances of justice and his distinction between *objective* and *subjective* circumstances of justice, see *TJ*, 126-130.

³⁰ See Rawls’s discussion on the priority of the right over the good: Rawls 1988 and *PL*, 173-211.

(LP, 24), and they «have a moral character» (LP, 25). It is apparent that peoples' profile is shaped by normative criteria. It is also evident that, with respect to states, the empirical reference for peoples is not immediately clear. In addition, Rawls seems to suggest it is possible that peoples, as he describes them, do not already exist: as he writes, «the idea of a reasonably just society of well-ordered peoples will not have an important role in a theory of international politics until such peoples exist» (LP, 19). This further illuminates the paradoxical character of Rawls's requirements of realism. In fact, it is natural to ask how can requirements of realism be in any sense plausible if they neither refer to observable states of affairs nor to empirically ascertainable entities?

The answer to this question is connected to the fact that Rawls's conditions for a realistic utopia are worked out from within ideal theory. This is confirmed by Rawls's reference to laws as «they would be in a reasonably just and well-ordered democratic society» and to the law of peoples as it «would be in a reasonably just Society of just and decent Peoples», that is, to how they would be under *favourable* circumstances. Yet, Rawls's second condition, which, as well as the first one, puts forward requirements of realism, seems to disconfirm such an understanding. Indeed, introducing a third level of realism, this condition requires normative principles to be applicable to ongoing institutional arrangements: in order to be realistic, that is in order to be feasible, principles and models must be applicable to *current* social and political arrangements. Reference to current arrangements seems to contradict both the claim that Rawls is appealing to ideal theory and the already mentioned statement according to which the sphere of the possible is not determined by the actual. In addition, it seems, it must be that either Rawls works from within ideal theory and, accordingly, he disregards actual circumstances or he does not. And it also must be that either the limits of the possible are determined by ongoing arrangements or they are not. Actually, Rawls envisages a third possibility. Rawls *is* working from within ideal theory, but he is trying to make ideal theory as realistic as possible, so that ideal principles result closer and immediately applicable to current states of affairs. And Rawls *is* fixing the limits of the possible by reference to the actual, but the actual is not to be intended as static, it does not coincide with what is simply there, and it possesses some intrinsic value. Two interlinked elements show the plausibility of this interpretation. The first has to do with the location of the boundary between ideal and non-ideal theory, the second with the facts that are to be intended as feasibility constraints.

The Law of Peoples is clearly divided into ideal and non-ideal theory. Ideal theory addresses *liberal democratic peoples* and *decent societies*. As Rawls specifies, these latter, «though they are not liberal democratic societies, have certain features making them acceptable as members in good standing in a reasonable Society of Peoples» (LP, 5). Non-ideal theory, instead, «deals with noncompliance, that is, with conditions in which certain regimes refuse to comply with a reasonable Law of Peoples» and with «unfavourable conditions of societies whose historical, social and economic circumstances make their achieving a well-ordered regime, whether liberal or decent, difficult if not impossible» (LP, 5). The border between ideal and

non-ideal theory is drawn by Rawls in a peculiar way. Non-compliance and unfavourable circumstances are undeniably part of non-ideal theory, but, at first, it is not clear why decent non-liberal societies should be included in ideal theory. Their inclusion reveals that ideal theory is shaped by Rawls with the precise intent of making it more accommodating: «whenever the scope of toleration is extended», he writes, «the criteria of reasonableness are relaxed» (1993b, 561). Nonetheless, as Rawls asserts, not all peoples are included in ideal theory:

We seek to include other than liberal societies as members in good standing of a reasonable society of peoples. Hence, when we move to these societies, their domestic regimes are less, often much less, congenial to us. This poses the problem of the limits of toleration. Where are these limits to be drawn? Clearly, tyrannical and dictatorial regimes must be outlawed, and also, for basic liberal reasons, expansionist states like those of the Wars of Religion. The three necessary conditions for a well-ordered regime – that it respect[s] the principles of peace and not be expansionist, that its system of law meet the essentials of legitimacy in the eyes of its own people, and that it honor basic human rights – are proposed as an answer as to where those limits lie. These conditions indicate the bedrock beyond which we cannot go. (1993b, 561-562)

Thus, there is a limit beyond which it is not possible to relax criteria of reasonableness. Yet, Rawls assumes this limit quite arbitrarily: he does not offer any justification for it. Neither does he offer justification for including decent non-liberal peoples in ideal theory: he simply asserts that criteria of reasonableness should be relaxed. Actually, the only plausible justification available to Rawls is that, since the international dimension is characterised by a plurality of peoples with different histories, traditions and conceptions of justice, taking into account only liberal peoples would amount to nothing more than useless idealization, which would lead to an unrealistic and infeasible law of peoples, one lacking any practical relevance. Therefore, it is plausible to maintain, it is Rawls's concern with feasibility that forces him to extend ideal theory's boundaries: ideal theory is made more realistic in order to include circumstances of justice that are as similar as possible to actual circumstances and in order to make the transition from the status quo to a society of well-ordered peoples, as described within ideal theory, workable³¹.

In order to be meaningful and practically useful, ideal theory is required to give a reliable account of the circumstances of justice and, as a consequence, it must incorporate the fact of pluralism. Yet, Rawls specifies that reasonable pluralism is not «a historical fate we should lament» (2001, 5): reasonable pluralism is intrinsically valuable. In the context of liberal democratic societies, a similar claim might be partially warranted: reasonable pluralism is the outcome of the *free* exercise of reason within the framework of *free* institutions and, consequently, its inclusion in ideal theory can be justified. On the contrary, once Rawls's focus shifts to the international dimension, a similar claim about the worth of pluralism

³¹ See *LP*, 89-90.

looses most of its persuasiveness: in *The Law of Peoples*, as Thomas McCarthy asserts, reasonable pluralism is «tacitly replaced by de facto pluralism» (McCarthy 1997, 211), by mere *diversity*. If ideal theory strives for accommodating diversity, it is not *ideal* theory, it is just a depiction of the merely actual. Indeed, in *The Law of Peoples*, Rawls engages in ideal theory not for the purpose of depicting how a *perfectly* just society *ought* to be – as he asserts in *A Theory of Justice* – but in order to illustrate how a *minimally* just society *can* be. Rawls aims at showing that «the social world allows at least a decent political order, so that a reasonably just, though not perfect, democratic regime is possible» (2001, 4). Rawls's approach turns out to be, therefore, unbalanced towards realism, and feasibility considerations take priority over desirability: normative criteria are downgraded in order to meet feasibility constraints and ideal theory is re-shaped in order to lead to feasible principles.

Rawls's realistic utopianism seems unbalanced towards realism also because his utopian claims are not very demanding. According to Rawls, «a necessary condition for a political conception of justice to be *utopian* is that it use[s] political (moral) ideals, principles and concepts to specify a reasonable and just society» (*LP*, 14) and, parallelly, «a reasonably just Law of Peoples is *utopian* in that it uses political (moral) ideals, principles, and concepts to specify the reasonably right and just arrangements for a Society of Peoples» (*LP*, 17-18). It is not necessary to add much in order to show that these conditions are quite scarcely demanding. Utopian models can be made to coincide with current states of affairs if the latter can be re-described in normative terms: the re-articulation of the actual is not only necessary, but sufficient condition for a conception of justice to be utopian. According to Rawls, utopianism does not imply a *difference* between the actual and the ideal. Rather, Rawls puts forward a much less exigent condition for utopianism: he only requires utopian models to be described in the language of ideals and principles. This explains why readers who criticize *The Law of Peoples* for its lack of utopianism are wrong: from Rawls's perspective, utopianism is something different from what the expression commonly suggests. With this particular understanding in mind, it is easier to see why the role political philosophy covers in its realistically utopian profile is nothing more than a «variation» (2001, 4) of its reconciliatory function: by showing that the current social world – its practices and arrangements – can be understood and expressed in normative terms, political philosophy opens up the possibility for reconciliation.

It is fair to recognise that Rawls's utopianism is not exhausted by similar requirements. At the beginning of *The Law of Peoples*, he writes:

Political philosophy is realistically utopian when it *extends* what are ordinarily thought to be the limits of practical political possibilities and, in so doing, reconciles us to our political and social condition. Our hope for the future of our society rests on the belief that the social world allows a reasonably just constitutional democracy existing as a member of a reasonably just Society of Peoples. (*LP*, 11, emphasis added)

A realistic utopia, then, is not only intended to rearticulate actuality, but is also meant to enlighten *latent possibilities*, that is, possibilities that are not immediately on hand, but that are just beyond the actual and may be brought about by the development of the relevant tendencies on hand. In turn, the acknowledgement of *positive* possibilities latent in the current states of affairs allows for reconciliation, allows to «affirm our social world positively, not merely to be resigned to it» (2001, 3).

The final section of *The Law of Peoples* is devoted to the identification of the historical tendencies which allow to hold a *reasonable hope* in the concrete possibility of a just democratic society and of a well-ordered society of peoples. As Jan-Werner Müller correctly notices, «historical argument [...] is fundamental for the construction of what Rawls calls ‘realistic utopias’» (Müller 2006, 335). Rawls states that the possibility his realistic utopia enlightens is «not a mere logical possibility, but one that connects with the deep tendencies and inclinations of the social world» (LP, 128): only if it is a «distinct historical possibility» (Müller 2006, 335), Rawls’s realistic utopia can legitimately aspire to be the source of reasonable hope, of reasonable faith. Rawls relies on four basic facts – the fact of reasonable pluralism, the fact of democratic unity in diversity, the fact of public reason, and the fact of liberal democratic peace – in order to account for the plausibility of his realistic utopia. He asserts that «these facts can be confirmed by reflecting on history and on political experience» and that «they were not discovered by social theory; nor should they be disputed, as they are virtually *truisms*» (LP, 124, emphasis added). Thus, the four facts Rawls refers to are assumed as uncontroversial facts, but it seems that they are not. I refrain from commenting on each of them, but, since the fact of liberal democratic peace is particularly relevant, some brief remarks on it may suffice for highlighting the controversial character of the facts Rawls appeals to and the excessive simplification he operates in qualifying them as truisms.

Rawls asserts that «as a realistically utopian idea, the Law of Peoples must have a parallel process that leads peoples, including both liberal and decent societies, to accept willingly and to act upon the legal norms embodied in a just Law of Peoples» (LP, 44), a process that parallels the one accounting for the emergence of the sense of justice in the domestic case. To such an end, Rawls explains:

We conjecture, first, that the Law of Peoples the parties would adopt is the law that we – you and I, here and now – would accept as fair in specifying the basic terms of cooperation among peoples. We also conjecture, second, that the just society of liberal peoples would be stable for the right reasons, meaning that its stability is not a mere *modus vivendi* but rests in part on an allegiance to the Law of Peoples itself. (LP, 45)

Rawls clarifies that «this second conjecture needs to be confirmed by what *actually* happens historically» (LP, 45, emphasis added) and the necessary confirmation comes from the fact of liberal democratic peace. Rawls’s account of liberal democratic peace is based on empirical findings signalling that, «though liberal

democratic societies have often engaged in war against non-democratic states, since 1800 firmly established liberal societies have not fought one another» (*LP*, 51)³². According to Rawls's interpretation, such historical evidence shows the possibility of a peaceful and cooperating society of well-ordered peoples. Yet, Rawls's account is partial: he does not consider other sources that, while offering different interpretations of empirical evidence, highlight the scarce significance of the democratic-peace thesis, if not its deceptiveness. Therefore, appeal to democratic peace is not uncontroversial and also the other facts Rawls refers to are not truisms: they represent, at most, tendencies characterising current states of affairs. Rawls would agree on this point: he himself speaks of «tendencies and inclinations of the social world».

Let us admit the tendencies Rawls identifies are actually the relevant tendencies at hand. Rawls believes that the presence of similar tendencies justifies us to hold a reasonable hope, a reasonable faith in the possibility of the realization of just arrangements, as he says, «both at home and abroad» (*LP*, 128). The difficulty is that, without a certain *teleological* understanding of history, nothing grounds such a hope: nothing guarantees these tendencies will develop in the supposed direction. Kant, too, to whom Rawls traces back the idea of political philosophy as the defence of reasonable faith³³, relies on a purposive understanding of history³⁴. In addition, unless a sort of Hegelian account of history is available, nothing assures that the possibilities Rawls's realistic utopia enlightens are valuable: they remain mere possibilities. The problem is that, if he intends to remain faithful to his method of avoidance, Rawls cannot appeal to any philosophy of history and, consequently, his appeal to reasonable faith is unwarranted and the desirability of his realistic utopian model cannot be properly vindicated. Thus, again, beyond the ambiguities and the confusion between normative and descriptive plans due to the attempt at developing a middle-ground position, Rawls's approach is unable to properly vindicate the desirability of the proposed principles because of his commitment to avoid controversial claims, to avoid philosophical claims. A commitment originated, it is worth reminding, by the urgency of accommodating the fact of pluralism, that is a commitment motivated by feasibility concerns.

5. REALISTIC UTOPIANISM: AN ASSESSMENT

Rawls's model for political philosophy evolves from its utopian profile in *A Theory of Justice* to realistic utopianism in his latest works. In its original form, Rawls's project assigns priority to desirability: the whole justificatory apparatus of *A Theory of Justice* is framed within ideal theory and, consequently, considerations of feasibility do not constrain normative claims. Rather, feasibility considerations are intended to illustrate justice as fairness's internal consistency and plausibility. In the development of Rawls's approach, the balance and the relationship between

³² For his account of the democratic peace, Rawls mainly relies on Russett 1993 and Doyle 1997.

³³ See Rawls 2000, especially pp. 309-325.

³⁴ See Kant 1784 and 1795.

desirability and feasibility change. As Chandran Kukathas and Philip Pettit write, it is possible to perceive certain «shifts in emphasis» that «come out most clearly in Rawls's increasing reliance on the *feasibility* arguments [...] and in the corresponding down-playing of considerations of *desirability*» (Kukathas and Pettit 1990, 142). This does not mean that Rawls completely abandons the aim of vindicating the desirability of the principles he proposes. More precisely, this implies that normative claims are limited and framed by feasibility constraints. Accordingly, Rawls's appeal to a *political* conception of justice is motivated by feasibility concerns: it is the need of accommodating the *fact* of reasonable pluralism that forces Rawls to avoid metaphysical and epistemological claims.

Nevertheless, Rawls's methodology is problematic, not because it is unbalanced towards realism and feasibility, but because it attempts to find a middle ground between realism and utopianism, between feasibility and desirability. It is the search of such a middle-ground position that leads to ambiguities and inconsistencies, which emerge more clearly in Rawls's latest commitment to realistic utopianism. Indeed, Rawls's realistic utopianism is an unsatisfactory profile for political philosophy. It forces political philosophy to look for a compromise between realism and utopianism, a compromise that, in turn, leads political philosophy to distort the character of both desirability and feasibility criteria and to elaborate principles and models that are neither properly utopian nor convincingly realistic. That is, in pursuing a middle ground, a synthesis between realism and utopianism, political philosophy is bound to downgrade the desirability of its proposals, without significantly enhancing their feasibility. On the one hand, Rawls's realism is hardly recognisable as such since its requirements are developed within ideal theory and, therefore, feasibility constraints lose their plausibility in limiting the set of options political philosophy can meaningfully recommend if it intends to undertake a properly practical function. On the other hand, Rawls's utopianism is very modest since it is aimed at proposing feasible models. As a consequence, a realistically utopian political philosophy, as envisaged by Rawls, is up to convincingly defend neither the feasibility nor the desirability of its models.

Rawls seems not troubled by such implications: he does not expect political philosophy to express the normative appeal of its models. Rawls assigns to political philosophy a more urgent practical and political task: political philosophy should support reconciliation and it should promote reasonable hope in the possibility that the social world renders it feasible to enact desirable arrangements. Indeed, according to Rawls, reconciliation and hope are crucial for the character of *actual* political life:

The answer we give to the question whether a just democratic society is possible and can be stable for the right reasons affects our background thoughts and attitudes about the world as a whole. And it affects these thoughts and attitudes before we come to actual politics, and limits or inspires how we take part in it. Debates about general philosophical questions cannot be the daily stuff of politics, but that does not make these questions without significance, since what we think their answers are will shape the underlying attitudes

of the public culture and the conduct of politics. If we take for granted as common knowledge that a just and well-ordered democratic society is impossible, then the quality and tone of those attitudes will reflect that knowledge. (*PL*, lix)

Reading this and other similar statements, it seems clear that political philosophy is no more supposed to be engaged in rational argumentation aimed at illustrating how politics *ought* to be. In fact, under this respect, political philosophy is deprived by Rawls of any authoritativeness:

A liberal political philosophy which, of course, accepts and defends the idea of constitutional democracy, is not to be seen as a *theory*, so to speak. Those who write about such a doctrine are not to be viewed as experts on a special subject, as may be the case with the science. Political philosophy has not special access to fundamental truths, or reasonable ideas, about justice and the common good, or to other basic notions. Its merits, to the extent it has any, is that by study and reflection it may elaborate deeper and more instructive conceptions of basic political ideas that help us to clarify our judgements about the institutions and policies of a democratic regime. (2007, 1, emphasis added)

As described in this passage, political philosophy is not intended to investigate on how a just society ought to be. The role of political philosophy – if it has any and Rawls seems quite hesitating on this – consists in *clarification*, consists in better articulating political ideals. Curiously, this is precisely the function the analytic tradition has usually envisaged for political philosophy before the publication of *A Theory of Justice*. Is this an adequate profile for normative political philosophy? It seems that it is not. It seems that *normative* political philosophy cannot avoid engaging into questions concerning what is desirable and how politics ought to be: normative political philosophy cannot renounce neither to address such questions nor to try and vindicate the authoritativeness of its answers.

REFERENCES

- Ackerman, B. (1994), «Political Liberalisms», *The Journal of Philosophy*, vol. 91, n. 7, pp. 364-386.
- Baier, K. (1989), «Justice and the Aims of Political Philosophy», *Ethics*, vol. 99, n. 4, pp. 771-790.
- Barry, B. (1995), «John Rawls and the Search of Stability», *Ethics*, vol. 105, n. 4, pp. 874-915.
- Brighouse, H. (1994), «Is There Any Such Thing as Political Liberalism? », *Pacific Philosophical Quarterly*, vol. 75, n. 3-4, Special Double Issue: *John Rawls's Political Liberalism*, pp. 318-332.
- Buchanan, A. (2000), «Rawls's Law of Peoples: Rules for a Vanished Westphalian World», *Ethics*, vol. 110, n. 4, pp. 697-721.
- Cohen, G. A. (2003), «Facts and Principles», *Philosophy and Public Affairs*, vol. 31, n. 3, pp. 211-245.
- (2008), *Rescuing Justice and Equality*, Harvard University Press, Cambridge Mass.
- Cohen, J. (1993), «Moral Pluralism and Political Consensus», in D. Copp, J. Hampton and J. E. Roemer (1994) [eds.], *The Idea of Democracy*, Cambridge University Press, Cambridge, pp. 270-291.
- Copp, D. (1996), «Pluralism and Stability in Liberal Theory», *The Journal of Political Philosophy*, vol. 4, n. 3, pp. 191-206.
- Daniels, N. (1979), «Wide Reflective Equilibrium and Theory Acceptance in Ethics», *The Journal of Philosophy*, vol. 76, n. 5, pp. 256-282.
- (1996), *Justice and Justification. Reflective Equilibrium in Theory and Practice*, Cambridge University Press, Cambridge.
- Doyle, M. (1997), *Ways of War and Peace*, Norton, New York.
- Freeman, S. (2003a), «Introduction», in S. Freeman (2003) [ed.], *The Cambridge Companion to Rawls*, Cambridge University Press, Cambridge, pp. 1-61.
- (2003b) «Congruence and the Good of Justice», in S. Freeman (2003) [ed.], *The Cambridge Companion to Rawls*, Cambridge University Press, Cambridge, pp. 277-315.
- Habermas, J. (1995), «Reconciliation through the Public Use of Reason: Remarks on John Rawls's Political Liberalism», *The Journal of Philosophy*, vol. 92, n. 3, pp. 109-131.
- Hampton, J. (1989), «Should Political Philosophy Be Done without Metaphysics? », *Ethics*, vol. 99, vol. 4, pp. 791-814.
- (1994), «The Common Faith of Liberalism», *Pacific Philosophical Quarterly*, vol. 75, n. 3-4, Special Double Issue: *John Rawls's Political Liberalism*, pp. 186-216.
- Hardimon, M. O. (1994), *Hegel's Social Philosophy and The Project of Reconciliation*, Cambridge University Press, Cambridge Mass.
- Hegel, G. W. F. (1821), *Grundlinien der Philosophie des Rechts*; Eng. tran. *Elements of the Philosophy of Right*, Cambridge University Press, Cambridge [1991].
- Hill, T. E. Jr. (1994), «The Stability Problem in Political Liberalism», *Pacific Philosophical Quarterly*, vol. 75, n. 3-4, Special Double Issue: *John Rawls's Political Liberalism*, pp. 333-352.
- Hurrel, A. (2001), «Global Inequality and International Institutions», in T. Pogge (2001) [ed.], *Global Justice*, Blackwell Publishing, Malden, pp. 32-54.

- Kant, I. (1784) *Idee zu einer allgemeinen Geschichte in weltbürgerlicher Absicht*; Eng. tran. *Idea for a Universal History with a Cosmopolitan Purpose*, in I. Kant, *Political Writings*, Cambridge University Press, Cambridge [1991], pp. 41-53.
- (1785), *Grundlegung zur Metaphysik der Sitten*; Eng. tran. *Groundwork of the Metaphysics of Morals*, Cambridge University Press, Cambridge [1997].
 - (1795), *Zum Ewigen Frieden*; Eng. tran. *Perpetual Peace: a Philosophical Sketch*, in I. Kant, *Political Writings*, Cambridge University Press, Cambridge [1991], pp. 93-130.
- Klosko, G. (1994), «Rawls's Argument from Political Stability», *Columbia Law Review*, vol. 94, n. 6, pp. 1882-1897.
- Krasnoff, L. (1998), «Consensus, Stability, and Normativity in Rawls's Political Liberalism», *The Journal of Philosophy*, vol. 95, n. 6, pp. 269-292.
- Kukathas, C. and Pettit, P. (1990), *Rawls. A Theory of Justice and its Critics*, Polity Press, Cambridge.
- Kuper, A. (2000), «Rawlsian Global Justice. Beyond The Law of Peoples to a Cosmopolitan Law of Persons», *Political Theory*, vol. 28, n. 5, pp. 640-674.
- Larmore, C. E. (1999), «The moral Basis of Political Liberalism», *The Journal of Philosophy*, vol. 96, n. 12, pp. 599-625.
- McCarthy, T. (1997), «On the Idea of a Reasonable Law of Peoples», in J. Bohman and M. Lutz-Bachmann (1997) [eds.], *Perpetual Peace. Essays on Kant's Cosmopolitan Ideal*, Massachusset Institute of Technology Press, Cambridge Mass.
- Müller, J. W. (2006), «Rawls Historian: Remarks on Political Liberalism's Historicism», *Revue Internationale de Philosophie*, vol. 60, n. 237, pp. 327-339.
- Neal, P. (1994), «Does He Mean what He Says? (Mis)Understanding Rawls's Practical Turn», *Polity*, vol. 27, n. 1, pp. 77-111.
- O'Neill, O. (1988), «Ethical Reasoning and Ideological Pluralism», *Ethics*, vol. 98, n. 4, pp. 705-722.
- (2001), «Agents of Justice», in T. Pogge [ed.], *Global Justice*, Blackwell Publishing, Malden, pp. 188-203.
 - (2003), «Constructivism in Rawls and Kant», in S. Freeman (2003) [ed.], *The Cambridge Companion to Rawls*, Cambridge University Press, Cambridge, pp. 347-367.
- Rawls, J. (1971), *A Theory of Justice*, Harvard University Press, Cambridge Mass.
- (1980), «Kantian Constructivism in Moral Theory», *The Journal of Philosophy*, vol. 77, n. 9, pp. 515-572; now in J. Rawls (1999), *Collected Papers*, Harvard University Press, Cambridge Mass., pp. 303-358.
 - (1985), «Justice as fairness: Political not Metaphysical», *Philosophy and Public Affairs*, vol. 14, n. 3, pp. 219-251; now in J. Rawls (1999), *Collected Papers*, Harvard University Press, Cambridge Mass., pp. 388-414.
 - (1987), «The Idea of an Overlapping Consensus», *Oxford Journal of Legal Studies*, vol. 7, n. 1, pp. 251-276; now in J. Rawls (1999), *Collected Papers*, Harvard University Press, Cambridge Mass., pp. 421-448.
 - (1988), «The Priority of Right and Ideas of the Good», *Philosophy and Public Affairs*, vol. 17, n. 4, pp. 251-276; now in J. Rawls (1999), *Collected Papers*, Harvard University Press, Cambridge Mass., pp. 449-472.

- (1989), «The Domain of the Political and Overlapping Consensus», *New York Law Review*, vol. 64, n. 2, pp. 233-255; now in J. Rawls (1999), *Collected Papers*, Harvard University Press, Cambridge Mass., pp. 473-496.
- (1993a), *Political Liberalism*, Columbia University Press, New York.
- (1993b), «The Law of Peoples», in S. Shute and S. Hurley (1993), [eds.], *On Human Rights: The Oxford Amnesty Lecture, 1993*, Basic Books, New York, pp. 41-82; now in J. Rawls (1999), *Collected Papers*, Harvard University Press, Cambridge Mass., pp. 528-564.
- (1995), «Reply to Habermas», *The Journal of Philosophy*, vol. 93, n. 3; now in J. Rawls (1996), *Political Liberalism*. Expanded Edition, Columbia University Press, New York, pp. 372-434.
- (1999), *The Law of Peoples with "The Idea of Public Reason Revisited"*, Harvard University Press, Cambridge Mass.
- (2000), *Lectures on the History of Moral Philosophy*, Harvard University Press, Cambridge Mass.
- (2001), *Justice as Fairness. A Restatement*, Harvard University Press, Cambridge Mass.
- (2007), *Lectures on the History of Political Philosophy*, Harvard University Press, Cambridge Mass.
- Raz, J. (1990), «Facing Diversity: the Case of Epistemic Abstinence», *Philosophy and Public Affairs*, vol. 19, n. 1, pp. 3-46.
- Rousseau, J. J. (1762), *Du contract social ou Principes du droit politique*; Eng. tran. *The Social Contract*, in J. J. Rousseau, *The Social Contract and Other Later Political Writings*, Cambridge University Press, Cambridge [1997].
- Russett, B. (1993), *Grasping the Democratic Peace*, Princeton University Press, Princeton.
- Weinstock, D. M. (1994), «The Justification of Political Liberalism», *Pacific Philosophical Quarterly*, vol. 75, n. 3-4, Special Double Issue: *John Rawls's Political Liberalism*, pp. 165-185.
- Wingenbach, E. (1999), «Unjust Context: the Priority of Stability in Rawls' Contextualized Theory of Justice», *American Journal of Political Science*, vol. 43, n. 1, pp. 213-232.