

# Rawlsian Public Reason and the Democratic Form of Life

Cyril Hédoin\*

*University of Reims Champagne-Ardenne*

**Abstract:** This paper outlines and critically assesses John Rawls’s account of the democratic form of life that he develops in the context of his political turn. My overall view about Rawls’s achievements is in general skeptical. On the one hand, I argue that this account constitutes a major step toward a better understanding of the political morality of liberal democratic societies. On the other hand, I claim that Rawls’s account of the democratic form of life is inappropriate, as far as it fails to acknowledge the full implications of the kind of pluralism that Rawls himself identifies. The roots of the problem are to be found in a conception of the political that puts too much emphasis on agreement and too little on the conflictual dimension and, more generally, in a problematic account of the idea of public reason. Rawls, therefore, fails to establish the possibility of just liberal democracies. The failure is however valuable because it points toward a new direction of research: the issue of the robustness of liberal societies, and the related relevance of polycentric governance.

**Keywords:** Rawls – Stability problem – Democratic form of life – Public Reason - Liberalism

## 1. Introduction

John Rawls’s so-called “political turn” (Weithman 2011), which started in the 1980s and culminated in the publication of *Political Liberalism* (Rawls 1993, henceforth *PL*), consists of a reinterpretation of his theory of justice (“justice as fairness”) in terms of a political conception not relying on any beliefs, values, and ideals constitutive of a specific comprehensive doctrine or conception of the good. As he himself indicates, this reinterpretation has been largely motivated by the problematic treatment of the stability problem in *A Theory of Justice* (Rawls 1971, henceforth *TJ*). By interpreting justice as fairness as a political conception, Rawls’s objective is to establish that a liberal conception of justice can gather a consensus within

---

\* Professor of economics, UFR de sciences économiques, sociales et de gestion, and REGARDS economics and management research center.

Contact: [cyril.hedoin@univ-reims.fr](mailto:cyril.hedoin@univ-reims.fr)

Paper prepared for the WINIR annual virtual conference on “Polycentric Governance and the Challenges of the 21<sup>st</sup> Century”.

This version: 28 June 2022.

democratic societies that are characterized by a reasonable pluralism of views about the good life.

In the pursuit of this objective, Rawls makes several key assumptions and claims about the nature of what he dubs the “democratic political culture”, or what I shall call more generally in this paper, the *democratic form of life*. Rawls’s characterization of the democratic form of life builds in particular on his account of the idea of *public reason*. The idea of public reason serves the key role of ensuring the “stability for the right reasons” of the political conception of justice encapsulated in the basic structure by fostering an overlapping consensus between diverse conceptions of the good. The resulting “well-ordered society” can then be interpreted as an ideal normative model of a just constitutional democracy. It follows that if Rawls’s account of the democratic form of life, and in particular his account of the idea of public reason, are viewed as a good depiction of the practices through which western democracies regulate conflicts and disagreements in the public sphere, then his political liberalism successfully establishes the *possibility* of liberal democracy.

My objective in this paper is to outline Rawls’s account of the democratic form of life and to provide a critical assessment of it. My overall view about Rawls’s achievements on this matter is in general skeptical. On the one hand, I argue that this account constitutes a major step toward a better understanding of the political morality of liberal democratic societies. On the other hand, I also claim that Rawls’s move remained incomplete and that his account of the democratic form of life is inadequate. Public reason may not be as determinate as Rawls thought and the distinction between the comprehensive (where diversity rules) and the political (where consensus prevails) is probably illusory. Based on this, I examine the possibility that Rawls’s political liberalism may not be political enough to grapple with the contemporary challenges that confront liberal democratic societies. If the essence of the democratic form of life is the consensus – as Rawls seemed to think, then a possibility is that liberal and open societies may not sustainably remain democratic in this sense. This invites us to consider the possibility of liberal but non-fully democratic societies, especially in the context of polycentric governance.

The rest of the paper is organized as follows. Section 2 provides more detail about the stability problem and Rawls’s political turn. Section 3 discusses Rawls’s account of the idea (and the ideal) of public reason and explains its centrality in the political reinterpretation of justice as fairness. Based on this, section 4 makes explicit Rawls’s mostly implicit characterization of the democratic form of life. I note in particular in this section an unexpected but welcome side-effect of this characterization, i.e., the fact that it acknowledges that the stability of liberal democracies partially relies on the existence of a proper ethos. Section 5 starts the critical discussion by considering the criticism that Rawls’s political liberalism is actually not political enough by underestimating the extent of public disagreement in liberal democracies. Section 6 reconsiders the idea of public reason, suggesting for it a different scope, structure, and content. Following Gaus (2019), I coin the concept of a *minimal liberal social world* to capture the form of minimal public consensus that should prevail in a liberal society. Section 7 builds on these considerations to push for disentanglement of the concepts of liberalism and democracy. Section 8 takes stock and suggests that if liberal societies are also open, robustness more than

stability is the relevant issue to be considered. Polycentric forms of governance, eventually including non-democratic components, are particularly relevant to consider in this context. Section 9 briefly concludes.

## 2. The Stability Problem and Rawls's Political Turn

Most of the discussions that have followed the publication of Rawls's *TJ*, especially from economists, have focused on the principles of justice advanced and the device through which Rawls has argued for them – the original position. A significant portion of Rawls's *magnum opus* – its whole third part – is however dedicated to the separate but still important issue of the *stability* of the scheme of cooperation that results from a conception of the justice and the basic structure in which it is encapsulated. Basically, the problem is to determine whether individuals, once the veil of ignorance is lifted, would be willing to abide by the principles of justice chosen in the original position. This is a fundamental issue, as Rawls (1971: 498) points out that considerations of stability are relevant for the choice of principles by parties in the original position.

The issue of stability can be decomposed into two separate problems that Rawls (1971: 269) identifies as the isolation and the assurance problems.<sup>1</sup> The former “arises whenever the outcome of the many individuals' decisions made in isolation is worse for everyone than some other course of action, even though, taking the conduct of the others as given, each person's decision is perfectly rational” (Rawls 1971: 269). In this prisoner's dilemma type of situation, individuals will realize that it may be to their advantage to deviate from the prescriptions following the adoption of principles of justice. The latter reflects the fact that participation in a cooperative scheme may be rational only conditional on the expectation that others will also do their part. Everyone should have the assurance that everyone else intends to behave in accordance with the principles of justice and to respect the agreement from the original position.

Rawls's initial treatment of the stability problem is in terms of congruence between the right and the good. Justice as fairness – i.e., the conception of justice corresponding to Rawls's two principles of justice – is thought to be “inherently stable” thanks to its ability to work through institutions reinforcing the sense of justice nudging individuals toward cooperation. The sense of justice corresponds to the recognition of a moral duty to do one's part in a cooperative scheme, even when defection could be to one's personal advantage. Within the appropriate institutional framework, the sense of justice becomes part of the individuals' conception of the good, in such a way that cooperation is rational. Rawls suggests in this perspective that justice as fairness is inherently stable because it is constituted by a small number of ideals (friendship, personal conduct, association) that, in the context of the moral psychology he is assuming (Rawls 1971: 490-6), trigger and reinforce the required sense of justice.

---

<sup>1</sup> See also Gaus (2011) who distinguishes between the justificatory instability problem and the assurance problem.

Starting in the 1980s and culminating with the publication of *PL* (Rawls 1993), Rawls's so-called "political turn" consists in a reinterpretation of justice as fairness as a *political* conception of justice. Following this reinterpretation, justice as fairness is no longer viewed as expressing an ethical view belonging to a particular comprehensive doctrine but rather as providing a political framework within which persons holding different views about the good life can agree on a conception of justice. Though several reasons and motivations underlying this political turn have been singled out by commentators, Rawls is actually quite explicit about the objective of the reinterpretation of his theory of justice. In the introduction of *PL*, Rawls writes

"Certainly, as I have indicated, there are important differences [between *TJ* and *PL*]. But to understand the nature and extent of these differences, one must see them as arising from *trying to resolve a serious problem internal to justice as fairness, namely from the fact that the account of stability in Part III of Theory is not consistent with the view as whole...* To explain: the serious problem I have in mind concerns the unrealistic idea of a well-ordered society as it appears in *Theory*. An essential feature of a well-ordered society associated with justice as fairness is that all its citizens endorse this conception on the basis of what I now call a comprehensive philosophical doctrine" (my emphasis).

So, it is his dissatisfaction with the treatment of the stability problem that has pushed Rawls to reinterpret his account of justice.<sup>2</sup> What is the nature of the problem exactly? In a nutshell, it finds its roots in the tension between two contradictory effects that the institutions of a well-ordered society ruled by the principles of justice as fairness would have on its members. On the one hand, as stated by the congruence solution to the stability problem, these institutions are assumed to foster a sense of justice relying on the endorsement of a small set of ideals. The endorsement of these ideals has the effect of making members of the well-ordered morally autonomous in the Kantian sense. The well-ordered society would be inherently stable because morally autonomous agents would desire to live by ideals promoted by justice as fairness. On the other hand, however, Rawls also acknowledges that the institutions of the well-ordered society would encourage its members to think of themselves as free in their exercise of practical reason. Because of that, a well-ordered society is inevitably a society where a *reasonable pluralism* prevails. Free and autonomous persons will develop different and non-compatible views about the good life, grounded on non-compatible metaphysical beliefs and ethical convictions. As far as constitutional democracies approximate the ideal of the well-ordered society, reasonable disagreement due to the "burdens of judgment" is the normal outcome under the institutions of the Western world.

The implication is that it is incoherent to insist that inherent stability should result from the endorsement of justice as fairness. As far as it is constituted by a desire to live by some moral ideals, justice as fairness is just one among the many comprehensive doctrines that rational and reasonable persons can freely endorse in a constitutional democratic regime. Rawls's

---

<sup>2</sup> Given Rawls's explicit own account of his political turn, it is surprising that so many commentators have looked for a different rationale. Weithman (2011) provides what I regard as the definitive treatment of Rawls's political turn, based on a careful study of Rawls's texts. There is no longer any ambiguity, if there was before, about Rawls's motivation for it.

solution consists in reinterpreting justice as fairness as a “free-standing” political conception that no longer goes along with the endorsement of such ideals. Justice as fairness just consists of the two (political) principles of justice. Obviously, this does not solve the stability problem *per se*. The core of Rawls’s political liberalism is to revise the model of the well-ordered society to provide an answer to the following question: “how is it possible for there to exist over time a just and stable society of free and equal citizens, who remain profoundly divided by reasonable religious, philosophical, and moral doctrines?” (Rawls 1993: 3). Rawls’s use here and elsewhere of the term “possibility” is not anecdotal. It denotes a concern with establishing that our (at least, Westerners’) belief that the natural and social world is friendly to just liberal democracies is well-founded (Weithman 2009). The revised model of the well-ordered society can be seen in this perspective as an “existence proof” of just liberal democracies.<sup>3</sup> This is in this context that the idea (and the ideal) of public reason makes its appearance in Rawls’s theoretical apparatus.

### 3. The Idea and Ideal of Public Reason

It would take too much time and space to provide the detail of Rawls’s revised solution to the stability problem following his political turn. I shall only offer a sketch of it, focusing instead on the role played by the concept of public reason.<sup>4</sup> This is indeed the latter that says the most about Rawls’s essentially implicit conception of the democratic form of life.

Rawls’s revised account of the stability of just constitutional democracy relies on three separate ideas, each of which is an innovation or at least an alteration of justice as fairness as stated in *TJ*: (i) the idea of justice as fairness as a “freestanding” political conception; (ii) the idea of an overlapping consensus; and (iii) the idea of public reason. I have already quickly mentioned the first above. By freestanding, Rawls means that the two principles of justice constitutive of justice as fairness do not depend for their justification on any metaphysical or epistemological doctrine or belief. A political conception of justice is separate from comprehensive doctrines not in the sense that it denies their existence or asserts a discontinuity between political and other values. It rather means that the same political conception of justice can presumably be endorsed and supported *from within* a great diversity of comprehensive doctrines but does not depend on any specific one (e.g., Rawls 1993: 10). Obviously, no political conception of justice can be completely “freestanding”. In the case of justice as fairness, it starts “from the fundamental ideas of a democratic society” (Rawls 1993: 40), ideas that I shall make more explicit in the next section. These ideas provide the justificatory background for the original position and the related notion of political constructivism. As I shall discuss below in more detail, Rawls nonetheless leaves open the possibility that other freestanding political

---

<sup>3</sup> As Weithman (2009) rightly notes, Rawls’s political liberalism is not concerned with the issue of the *transition* from unjust to just societies. Filling the mathematical analogy further, we can therefore view the well-order society as a “fixed point” model of just liberal democracies.

<sup>4</sup> An authoritative source is again Weithman (2011). Samuel Freeman (2006: chapters 5 and 6) provides a detailed account of Rawls’s treatment of the stability problem in *PL*, also emphasizing the role played by the concept of public reason.

conceptions could be justified from a democratic standpoint. Even though the range of such justifiable conceptions stays limited to “liberal” ones,<sup>5</sup> it is a significant concession made by Rawls that will turn out to have important implications for the plausibility of his account of the democratic form of life.

The identification of a freestanding political conception of justice corresponds to the first stage of Rawls’s political liberalism. This stage specifies “the fair terms of cooperation among citizens and specifies when a society’s basic institutions are just” (Rawls 1993: 133). The idea of an overlapping consensus corresponds to the second stage and is more directly concerned with the stability problem. The overlapping consensus in question is between “reasonable” comprehensive doctrines, i.e., doctrines from which a liberal political conception of justice can be endorsed. Following Rawls, it can be contrasted with the notion of *modus vivendi*. In the latter, stability is the result of a compromise between conflicting views about the good. While this compromise can be stable in some sense, it will generally not be inherently stable.<sup>6</sup> From the perspective of political liberalism, the problem is that the stability fostered by a *modus vivendi* is *not for the right reasons*. While Rawls (1993: 158-68) does conjecture that an overlapping consensus might emerge from a *modus vivendi*, passing through a stage of “constitutional consensus”, its nature is fundamentally different at least in two respects (Rawls 1993: 147). First, the object of the consensus, namely the political conception of justice, is itself a moral conception. The agreement that prevails in a *modus vivendi* does not have this status, except under a deflationary account of morality. Second, the consensus is affirmed on moral grounds, as it depends on the range of values, principles, and ideas that underlie the political conception of justice. As a result, the main contrast between a *modus vivendi* and an overlapping consensus is that the latter but not the former is robust with respect to a change in the balance of relative power between the conflicting views.

As noted by Freeman (2006: 190-2), the idea of the overlapping consensus somehow plays the role of a substitute for the congruence argument that appears in *A Theory of Justice*. A major difference is of course that the congruence does not occur among the same elements: between the right and the good in the latter, between comprehensive doctrines and a political conception of justice in the former. The overlapping consensus provides however only an incomplete account of the stability of the well-ordered society and more generally of the justificatory endeavor of a political conception of justice. Indeed, after the first stage, the freestanding political conception is only *pro tanto* justified, based only on the political values that one may find in the democratic political culture. Once achieved, the overlapping consensus realizes the *full justification* of the political conception: “the citizen accepts a political conception and fills out its justification by embedding it in some way in the citizen’s

---

<sup>5</sup> According to Rawls, any plausible political conception that can be adopted from the original position or through another theoretical device derived from democratic ideas and values would satisfy three conditions. First, it would specify a set of rights, liberties, and opportunities similar to those we usually find in constitutional democratic regimes. Second, an assignment of a priority to this set, especially with respect to perfectionist values and claims about the good. Third, the definition of measures guaranteeing to all citizens the all-purpose means to make effective use of their liberties and opportunities (Rawls 1993: 6).

<sup>6</sup> At least within a Hobbesian outlook, stability will rather tend to be imposed.

comprehensive doctrine as either true or reasonable” (Rawls 1993: 386). A last step is however the *public justification* of the political conception; it “happens when all the reasonable members of political society carry out a justification of the shared political conception” and “take one another into account as having reasonable comprehensive doctrines that endorse that political conception, and this mutual accounting shapes the moral quality of the public culture of political society” (Rawls 1993: 386). Public justification is intimately related to the idea of public reason and to the requirement of stability for the right reasons of the well-ordered society, the latter replacing the notion of inherent stability of *TJ*.

The concept of public reason is central in several contemporary social contract theories and has been articulated through many specific conceptions.<sup>7</sup> It appears lately in Rawls’s writings, especially in the sixth lecture of *PL* and in a subsequent article refining Rawls’s account of it (Rawls 1997). The Rawlsian idea of public reason refers to two sets of political values.<sup>8</sup> On the one hand, the political values that underlie a liberal political conception of justice. On the other hand, the values that serve as guidelines for public inquiry and guide the application of principles of justice in the well-ordered society. Hence, public reason “specifies at the deepest level the basic moral and political values that are to determine a constitutional democratic government’s relation to its citizens and their relation to one another” (Rawls 1997: 574). In relation to its content which basically consists of the principles of justice and principles guiding their applications, Rawls indicates that the scope of the idea of public reason is restricted to “constitutional essentials and matter of basic justice”. Correspondingly, it only applies to discussions within the “public political forum”, which is divided into three parts: the discourse of judges in their decisions, the discourse of government officials, and the discourse of candidates for public office (Rawls 1997: 575). Discussions among citizens rather belong to the “background culture” and are dispensed from the requirement of public reason, though Rawls notes that citizens voting on constitutional essentials and matters of basic justice “are to think of themselves *as if* they were legislators” and use the requirements of public reason to assess government officials and candidates for public office (Rawls 1997: 577). I put aside for the time being these restrictions about the scope and the content of the Rawlsian idea of public reason, as I wish to focus on two key routes through which it is realized.

First, Rawls makes an explicit connection between the idea of public reason and what he calls the “liberal principle of legitimacy”,

“our exercise of political power is proper and hence justifiable only when it is exercised in accordance with a constitution the essentials of which all citizens may reasonably be expected to endorse in the light of principles and ideals acceptable to them as reasonable and rational” (Rawls 1993: 216).

---

<sup>7</sup> See Quong (2018) for a recent and useful overview. The concept is also central in Amartya Sen’s (2009) theory of justice, despite its explicit rejection of social contract accounts.

<sup>8</sup> In *Political Liberalism*, Rawls distinguishes between the political values underlying the political conception of justice and the “values of public reason” (Rawls 1993: 224). In Rawls (1997), both sets of values are now encompassed under the same heading of public reason.

According to the liberal principle of legitimacy, the exercise of political power is legitimate only if it is grounded on principles that every reasonable person is expected to rationally endorse. In other words, only *publicly justified political principles* can license the legitimate use of political power, especially coercion, in a well-ordered society. This entails that the use of nonpublic reasons, issued from particular comprehensive doctrines, cannot provide the required form of justification. The idea of public reason follows from the liberal principle of legitimacy, which Rawls locates in the very nature of constitutional democratic regimes.<sup>9</sup>

In turn, featuring the idea of public reason as being part of the overlapping consensus over a political conception of justice imposes the recognition of a moral duty, the “duty of civility” ruling interactions within the public forum. This duty requires judges, government officials, candidates for public offices, and citizens voting on constitutional essentials and matters of basic justice to only use public reasons. It excludes, for instance, the use of religious or other metaphysical convictions to justify the use of political power over others.

The full realization of the duty of civility in the well-ordered society corresponds to what Rawls more specifically characterizes as the *ideal* of public reason. As Rawls (1993: 219) acknowledges, in a society ruled by this ideal, citizens voting on constitutional essentials and matters of basic justice would be close to expressing Rousseau’s general will. This is not the case in contemporary liberal democracies, and in this sense, the latter are not well-ordered societies. This does not invalidate the regulative function of the idea of public reason in liberal democracies, however. Provided that the political conception of justice and the idea of public reason are mutually supportive within an overlapping consensus,<sup>10</sup> the use of public reason provides a solution to the stability problem. The existence of the overlapping consensus guarantees that no one has any decisive nonpublic reason to rationally reject the prevailing political conception of justice. Moreover, the fact that this consensus and more generally the legitimacy of the use of political power are public provides everyone the assurance that everyone else recognizes her duty of civility and therefore abides by the publicly endorsed political conception of justice.<sup>11</sup>

---

<sup>9</sup> In his 1997 article, Rawls endorses what he calls the wide view of public reason. It licenses the use of nonpublic reasons in the public political forum as long as proper political (public) reasons are introduced in due course.

<sup>10</sup> Rawls (1993: 252) acknowledges the speculative character of this assumption. In essence, its status depends on the validity of the moral psychology put forward by Rawls in the eighth chapter of *TJ* (and briefly reminded in *PJ*, pages 84-5) and thus on the propensity of members of a well-ordered society to acquire a sense of justice inclining them to meet their duty of civility. This is one of the two components (the other being the thin theory of the good and the related conception of rational plans) of Rawls’s early treatment of the stability problem that remains in his political liberalism.

<sup>11</sup> A fairly recent literature has cast doubt over the effectiveness of the revised Rawlsian treatment of the stability problem. Most of the criticisms emphasize the inadequacy of the idea of public reason (at least in its Rawlsian version) in this perspective. See in particular Gaus (2011), Thrasher and Vallier (2015), Kogelmann and Stich (2016), Kogelmann (2019), Chung (2019), Chung (2020). See however Weithman (2011), Hadfield and Macedo (2012) and Quong (2010) for a defense of Rawls’s account. I have myself proposed a revision of Rawls’s model of public reason along Wittgensteinian lines to meet these critiques (Hédoin 2022). I shall ignore for the rest of the paper this issue and assume that Rawls’s solution is at least partially successful.

#### 4. The Rawlsian Account of the Democratic Form of Life, or Rawls's "Wittgensteinian Step"

It already transpired in the preceding section that Rawls's revised solution to the stability problem and more generally his political liberalism are grounded in some view about the values and political culture that we find – or should expect to find – in contemporary constitutional democratic regimes. While this point has already been noticed by commentators, its implication has not necessarily been properly worked out. The goal of this section is to make more explicit how Rawls characterizes the democratic form of life that serves as a template for his political liberalism. The following sections will critically assess the implications for the normative project (which is not only Rawls's) of establishing the possibility of a just liberal society.

Looking at the indexes of *TJ* and *PL*, it is interesting that the word "democracy" is not given a proper entry in either of them.<sup>12</sup> Nonetheless, the use of the words "democracy" and "democratic" is pervasive in Rawls's later book, especially in conjunction with the newly introduced concepts of a political conception of justice, overlapping consensus, and public reason discussed above. This marks a major shift in the focus of Rawls's political philosophy as a whole. *TJ* is mostly concerned with constructing a theory of justice figuring as a superior alternative to utilitarianism, intuitionism, and perfectionism. There is no doubt that in this theoretical endeavor, which is essentially pursued in the first two parts, Rawls is assuming a lot about the nature of the society in which justice as fairness is thought to apply and the corresponding considered judgments about the good and the right that serve as a starting point in the search of reflective equilibrium. But nowhere in the 1971 book are these assumptions explicitly, or even unambiguously even if implicitly attached to a definite characterization of a democratic regime. Even in section 13 titled "Democratic equality" which essentially consists of a derivation of the difference principle from a state of initial equality, it is unclear what the label "democratic" stands for. "Democratic equality" just figures as a name to designate a particular interpretation of the locutions "equally open" and "to everyone's advantage" in the second principle. This is not to say that Rawls thought that justice as fairness can or should prevail in any kind of society – the stability problem evidently shows that this is not the case – but the defense of this conception of justice is not grounded in any precise account of what a constitutional democracy is.<sup>13</sup>

The pattern is quite different in *PL*. Here, Rawls repeatedly tells his readers that the major concepts and ideas constituting his theoretical framework find their roots in the "democratic political culture" of constitutional regimes. This is the case for instance for the political/comprehensive distinction, as stated at the very outset of *PL*: "The dualism in political

---

<sup>12</sup> The closest in *TJ* is the entry "democratic equality". In *PL*, there is an entry "deliberative democracy, its idea of public reason".

<sup>13</sup> Rawls's subsequent reinterpretation of justice as fairness presented in *TJ* as a comprehensive doctrine confirms this. As a comprehensive doctrine, justice as fairness relies on metaphysical, ethical, and epistemological beliefs and principles that are not constitutive of the democratic form of life as Rawls will later characterized it.

liberalism between the point of view of the political conception and the many points of view of comprehensive doctrines doctrine is not a dualism originating in philosophy. Rather, it originates in the special nature of democratic political culture as marked by reasonable pluralism" (Rawls 1993). As reasonable pluralism is itself a feature of liberal democratic societies as Rawls conceived them, this suggests that the political/comprehensive distinction is not relevant as part of the justification of a conception of justice in other kinds of human societies.

More surprisingly, Rawls (1993: 8) also seems to suggest that the object of the overlapping consensus is itself to be found not through philosophical reflections, but rather "by looking to the public culture itself as the shared fund of implicitly recognized basic ideas and principles". A few pages later, pursuing the same idea, Rawls (1993: 13) indicates that one of the constitutive features of a political conception of justice is

"that its content is expressed in terms of certain fundamental ideas seen as implicit in the public political culture of a democratic society. This public culture comprises the political institutions of a constitutional regime and the public traditions of their interpretation (including those of the judiciary), as well as historic texts and documents that are common knowledge."

Rawls (1993: 35-6) gives some detail about the content of the political culture of a democratic society. In addition to the fact of reasonable pluralism, two other facts are publicly recognized: the "fact of oppression" and the "fact of majority". The former states that nothing but state power can maintain a shared understanding on a comprehensive doctrine; in other words, stability can only be imposed if its object is a comprehensive doctrine rather than a political conception. The latter consists of the observation that an enduring and secured democratic regime must rely on the free and willed support of its politically active citizens. The conjunction of these three facts of reasonable pluralism, oppression, and majority directly implies that a political conception of justice cannot serve as the basis for public justification in a constitutional democratic regime if it is not widely endorsed by citizens with diverse and conflicting comprehensive doctrines.

From this characterization of the democratic political culture, only a tiny step is required to fall back on the liberal principle of legitimacy and the duty of civility that are at the core of the idea (and ideal) of public reason. The implicit suggestion is that the democratic form of life as Rawls characterized it is *de facto* implicitly constituted by the recognition of values and principles that support both a liberal political conception of justice and the use of public reason in the discussion within the political forum. As Rawls (1993: 213) puts it, "[p]ublic reason is characteristic of democratic people: it is the reason of its citizens, of those sharing the status of equal citizenship." Beyond the ideas of public reason and the political conception of justice that are the content of the overlapping consensus, the democratic form of life is constituted by the two fundamental ideas that are at the bottom of the whole argumentative structure of political liberalism, i.e., the conception of society as a fair system of social cooperation (and its related idea of reciprocity) and the political conception of citizens as free and equal (Rawls 1993: 15-21).

Before starting the critical assessment of this conception of the democratic form of life, let me briefly contemplate what it permits Rawls to achieve in the context of the stability problem. In *TJ*, justice as fairness is defended as a philosophical doctrine against other philosophical doctrines based on metaphysical, epistemological, and ethical arguments. The issue of stability arises because it has to be shown that this doctrine is compatible with plausible assumptions about individuals' moral psychology and agency capacities. With his political turn, Rawls adopts a radically different strategy. Now, it is contended that justice as fairness can be publicly justified because it depends on ideas and principles which are *actually*, even if only implicitly, realized in the political practices of reasonable citizens of democratic societies. Put differently, Rawls is not contemplating from a philosophical viewpoint whether justice principles can be justified; he is asserting that all the ingredients for the justification and thus the stability of these principles are already present in the prevailing practices of constitutional democracies. Correspondingly, the role of the political philosopher has changed. It is no longer to create a new justice account and to defend it against its rivals, but rather to help us see that this account is already contained in our democratic form of life.

In this perspective, it is no exaggeration I think to present Rawls's political turn as a *Wittgensteinian step* that achieves partial naturalization of justice as fairness by uncovering the "language games" we are participating in our democratic political practices. This Wittgensteinian step has several intriguing implications. For instance, it trades on a subtle understanding of the relationship between the normative dimension of political philosophy and the role played by descriptive statements in this normative endeavor – being understood that it would be implausible to accuse Rawls of committing any kind of naturalistic fallacy. More to the point here, the Wittgensteinian step reveals that the justification of justice as fairness partially depends on the existence of a "democratic ethos" encapsulated in the practical use of public reason and the related acknowledgment of a duty of civility. This is intriguing because Rawls's initial statement of justice as fairness has been attacked for its failure to engage with the role of such an ethos at the level of individual conducts.<sup>14</sup> Whether it is a progress depends on the perspective taken about the proper justification of a theory of justice. I would tend to view it as a positive evolution as far as it reflects the need to better account for the social roots of liberal political morality. But this evolution also exposes Rawls's account to a new set of criticisms targeting the relevance of its characterization of the democratic form of life.

---

<sup>14</sup> See in particular Cohen (2008). Melenovsky (2016) convincingly rebuts the criticism as far as it is intended to imply that the *same* normative principles should apply to the basic structure and to individual conducts. It is actually apparent however that the idea of public reason does indeed apply only to individual conducts and that the duty of civility is very similar to Cohen's ethos. Quong (2010: 43, fn. 70 and 257, fn. 5) points out that there is a tension here between the rejection of Cohen's egalitarian ethos and the endorsement of the duty of civility that Rawlsians have failed to properly engage with.

## 5. Is Political Liberalism Political Enough?

Rawls's political liberalism builds on a sharp distinction – not a dichotomy as Rawls (1993: 10) does not assume a discontinuity between political and other values – between the domain of the political and what belongs to the “comprehensive”. This distinction, which it is argued is part of the democratic form of life, is essential in the perspective of solving the stability problem. It is however not easy to precisely characterize the political as this term is used in conjunction with key concepts of Rawls's theoretical apparatus (e.g., political conception of justice, political conception of persons, political constructivism). Several possible interpretations are available (see for instance Gaus 2003: 180-97). A first interpretation relies on the three features of political conceptions of justice that Rawls (1997: 584) identifies: (i) they apply to the basic structure; (ii) they can be presented independently from comprehensive doctrines; and (iii) they can be worked out from ideas that are implicit in the democratic form of life, especially the conception of free and equal persons and the idea of society as a fair system of cooperation. This is not very helpful, however. Indeed, these features are self-referential as they contain in their very statement either the notion of the political or of its complement (the comprehensive).<sup>15</sup>

A second interpretation relies on the way the basic structure can be characterized, especially through examples of political values that Rawls gives: justice, domestic tranquility, general welfare, equal liberties, ... (Rawls 1997: 584). It is not satisfactory, however. There is no a priori reason to assume that the basic structure of society (its “political and social institutions”) must be restricted to the regulation of behavior in the context of the values that Rawls identifies. Further, if it is of course perfectly permissible to analytically identify the domain of the political with the basic structure as Rawls conceptualizes it, this may not reflect the view of every citizen. Consider the case for instance of the moral and legal status of non-human animals or the issue of the social acceptability (and not only legality) of homosexuality or other nonmainstream sexual orientations. These issues do not refer to political values as Rawls seems to circumscribe them (i.e., those mentioned in the preamble of the U.S. Constitution). But there is no doubt that they nowadays are perceived and lived as political issues by a significant fraction of the population in constitutional democracies. The point is that the domain of the political may be evolving and differ across societies, even liberal-democratic ones. No a priori conception of the political seems to make sense.<sup>16</sup>

We are led to a third interpretation that is the one most in line with the “Wittgensteinian step” I have identified above. The characterization then takes place in two steps: (a) the political just is the sphere of values and principles that reasonable persons agree on, (b) as a matter of fact, in constitutional democratic regimes, given their comprehensive views, reasonable persons agree exactly on those values and principles referring to liberal political conceptions of justice and the idea of public reason as Rawls characterizes them. This interpretation has at least two

---

<sup>15</sup> The basic structure is indeed made of “basic political and social institutions” (Rawls 1997: 584). As indicated above, what I call the democratic form of life directly refers to the democratic political culture.

<sup>16</sup> This is Habermas's (1995) main point in his critical discussion of Rawls's political liberalism. Rawls's answer to it (published in the second edition of *PL*) suggests unsurprisingly that he does not endorse this interpretation.

advantages. First, it eschews the arbitrariness of identifying a priori the domain of the political and avoids any kind of self-reference. Second, it becomes an analytical truth that the liberal principle of legitimacy is satisfied in constitutional democratic regimes. By definition, the exercise of *political* power is backed up by an agreement over values and principles that justify it.

But this “Wittgensteinian” interpretation of the political in terms of the democratic form of life is fragile at two levels. It first introduces a circularity in the definition of another key concept in the Rawlsian apparatus, the concept of the “reasonable”. Rawls defines reasonable persons as persons who have developed a sense of justice that makes them willing to honor the reciprocity requirements following their participation in the cooperative venture of society. In turn, that implies that reasonable persons endorse reasonable comprehensive doctrines, i.e., comprehensive doctrines that can be part of the overlapping consensus over a political conception of justice. In other words, to be reasonable is to have views that support the political conception of justice in the context of an overlapping consensus and to be disposed to use public reason in the public political forum. But that seems to imply that assessing the reasonableness of a person and her comprehensive views depends on having at hand the political conception that gives its content to public reason.

This is not a decisive objection because it can be answered that the identification of the domain of the political and of reasonable persons proceed at the same time in a fixed-point relationship. To be reasonable is just to endorse the values and principles that are part of the overlapping consensus and the latter is identified by looking at what agree reasonable persons agree on.<sup>17</sup> But now, it can be argued that this fixed-point in actual liberal democratic societies is not located at the coordinates of the Rawlsian well-ordered society in the space of possible (just) societies, i.e., citizens do not agree on the same set of values and principles as claimed in Rawls’s political liberalism. This criticism can be developed in several directions. In a relatively benign form, it can be argued that not justice as fairness but other liberal conceptions of justice tend to be objects of the overlapping consensus. As I have highlighted above, this is a concession that Rawls is ready to make. Different liberal democratic societies would agree over different conceptions of justice that nonetheless satisfy the same set of conditions.<sup>18</sup> The trouble is that once this concession is made, it seems that others are difficult to avoid. For instance, discussions in the public political forum of actual liberal democracies suggest that several political conceptions are competing *within* the same society. Considerations of social justice are among the topics that are hotly debated on the occasion of elections or between political representatives in legislative and executive bodies. Even if the disagreement stays within the range of liberal conceptions, that implies that the content of public reason is partially indeterminate. And while we may be tempted to stigmatize officials and citizens holding *nonliberal* political conceptions of justice as unreasonable, the rise of populist

---

<sup>17</sup> Such a move is for instance advanced by Quong (2010).

<sup>18</sup> See footnote 5 above. See Gaus and Schoelandt (2017) for a discussion of Rawls’s various “models” of political liberalism, some of them allowing for a range of political conceptions of justice.

movements is nonetheless an enduring feature of liberal democracies that keep on strengthening (Mounk 2018).

Rawlsians may of course acknowledge these political facts and then simply argue that they indicate that contemporary Western societies are not (or no longer) well-ordered. The answer is however problematical in light of Rawls's very project of establishing the possibility of a well-ordered constitutional regime. The democratic form of life that Rawls is appealing to ground his model of the well-ordered society seems to be more than an idealization of liberal democracies, but actually a *fiction*. It might be thought that the source of the problem comes from the fact that Rawls's political liberalism misidentifies the domain of the political (the agreement prevails elsewhere than on matters of justice) or its extension (there is less agreement than political liberalism assumes). More radically, it may be due to the fact that political liberalism has its conception of the political wrong. Against the Rawlsian conception of the political where agreement prevails, we could oppose a *Weberian* conception that makes radical conflict the essence of the political.<sup>19</sup> In this view, the Rawlsian depiction of the democratic form of life is a fiction; liberal democracies are quite the contrary the theater of the permanent confrontation of incompatible forms of life that make conflicting judgments and demands over the good *but also* over the right. As one of the major contemporary proponents of this view has argued, this effectively means that Rawls's political liberalism is apolitical:

“The key move in Rawlsian political liberalism, in other words, is the removal from political life of the principles specifying the basic liberties and justice in distribution. These are not upshots of political discourse, or aspects of any real settlement or agreement achieved in actual political practice, but theorems, products of the peculiar species of theoretical reasoning that Rawls's method entails. Neither the method of reasoning nor the results of Rawlsian political liberalism possess any of the marks of indeterminacy and contingency that distinguish political discourse and practice. On the contrary, the object of Rawlsian political liberalism is the removal of these distinguishing marks of the political from both its method and its results. In consequence, political life is in Rawlsian political liberalism void of substance” (Gray 1995: 124).

Reflecting on whether there is any reason to prefer the Weberian conception of the political to Rawls's would lead us too far astray. What can be noted however is that among the burdens of judgment, Rawls (1993: 56-7) identifies “Berlin's view” about value pluralism, according to which “[n]o society can include within itself all forms of life... there is no social world without loss: that is, no social world that does not exclude some ways of life that realize in special ways certain fundamental values” (Rawls 1993: 196).<sup>20</sup> Rawls does not offer any compelling reason

---

<sup>19</sup> See in particular Max Weber's vocation lectures (Weber 2004). Of course, Weber is far from being alone in having entertained a conflict view of the political. For an insightful historical account of major liberal figures of the 20<sup>th</sup> century (Camus, Aron, Niebuhr, and Berlin) whose conception of the political is inspired by Weber's, see Cherniss (2021).

<sup>20</sup> For Berlin's account of value pluralism, see in particular his essays “Two Conceptions of Liberty” and “The Pursuit of the Ideal”, both reprinted in Berlin (2000). Gray (2020) offers a provocative and insightful interpretation of Berlin's account and its implications for liberal theory.

to think that what is true for ways of life in general should not in particular apply among rights establishing basic liberties, and between these rights and considerations relative to welfare and other values entering into conceptions of the good life. Quite the contrary, the attempt to isolate rights and liberties from the radical conflicts resulting from value pluralism faces the problem that the justification for the creation of specific rights and liberties can hardly be made without reference to welfare and more generally human flourishing, and thus to values that enter into radical conflicts.

Hence, Rawls himself provides elements that work against his political/comprehensive distinction and the related view of the political as the domain of agreement in the democratic form of life. The implications are of course significant, as we may now ask what remains of the Rawlsian conception of the democratic form of life in terms of shared public reason.

## **6. Public Reason and the Minimal Liberal Social World**

Building on Berlin's value pluralism and the related conflict view of the political, John Gray (2002) has developed an account of "agonistic liberalism" according to which the only form of stability that can be achieved between mutually exclusive forms of life is a *modus vivendi*. Gray's *modus vivendi* is an evolving state in which ways of life adjust to each other, especially through political means. Gray's point is that in any society, liberal democratic or not, there is no legal or social mechanism, nor any philosophical argument, that can automatically guarantee a social harmony between conflicting values and forms of life. Peaceful cohabitation is the uncertain product of political and social life, not an axiom from which philosophical analysis can start.

A trouble with this view is that it might be doubted that a *modus vivendi* may guarantee living conditions that are acceptable from any normative point of view. Philosophical considerations in the spirit of the Hobbesian tradition, but also theoretical models from the social sciences and casual observation of national and international (geo)politics indicate that the search for a *modus vivendi* may lead to disastrous results. Granted that there is no automatic pathway toward social harmony and stability, we may still want to know under what conditions a social and political order may emerge and establish conditions favorable for the pursuit by individuals of their interests and ends. We already know that Rawls's account of the democratic form of life is problematic in this respect because it assumes what needs to be demonstrated. It nonetheless provides elements that are useful in this endeavor. I shall in particular suggest in this section that a revised conception of the idea of public reason is still relevant, once exempted from the Rawlsian political/comprehensive distinction.

Acknowledging that neither a pure *modus vivendi* nor a stable Rawlsian well-ordered society can account for the normative foundations of liberal democracies, we have to search for an alternative framework. A good place to start is the still-growing theoretical and experimental literature in the social sciences on coordination and collective action problems. This literature

is large and not fully consistent but generates some key insights.<sup>21</sup> It shows that humans' ability to solve coordination and collective action problems is grounded in our ability to create (intentionally or not) and follow systems of rules, i.e., institutions. This ability is associated with an evolved set of cognitive capacities and cultural resources that make us able to "read the mind" of others, especially when they live in the same social environment. This generates the existence of "focal points" and salience effects that account both for the origins and the functioning of conventions and other institutions.<sup>22</sup> The functioning of institutions is strengthened by the biological and cultural evolution of prosocial dispositions that, everything else equals, inclines individuals toward following rules but also punishing non-followers, even at personal cost. This is especially the case in morally loaded interactions, where individuals tend to develop strong reactive attitudes toward those who deceive their normative expectations.

These results have led to significant advancements in our understanding of the nature of institutions. Institutions are systems of rules that correlate individuals' expectations and actions. They, so to speak, shape persons' theoretical and practical reasoning because what makes a rule *is* the fact that persons are sharing forms of reasoning (Gintis 2009). At the same time, these rules are enduring because they correspond to (stable) game-theoretic equilibria. It is therefore rational for individuals to follow the rules as they are cognitively representing and understanding them [(Aoki 2001); (Guala 2016)]. A key aspect is that the existence of institutions depends on some form of publicness. The fact that a rule  $R$  applies in circumstances  $C$  entails that for all participants in the corresponding social practice  $P(R, C)$ , it is commonly believed that each will (and ought) behave in a particular way that is constitutive of the fact of following  $R$ .<sup>23</sup> Such common beliefs emerge from the fact that individuals are *symmetric reasoners*, and this symmetric reasoning reflects at the bottom that participants in  $P$  share a common understanding of the situation in which the practice takes place (Hédoin [(2014); (2017)]).

Analytically, social life is constituted and made possible by institutions, and therefore publicly shared forms of reasoning. To form a society – or a community or any other collective entity – is to coordinate our activities by participating in social practices in which it is commonly understood how each participant should behave. That does not mean that agreement is pervasive in every aspect of the social life, and even less that everyone's interests are equally satisfied in the prevailing social practices. Quite the contrary, members of the society may disagree on many things, but this disagreement cannot have meaning outside a public domain

---

<sup>21</sup> See, for instance, Bicchieri [(2005); (2016)], Bowles (2016), Bowles and Gintis (2011), Greif (2006), Henrich (2015), Smith (2009). This is only the tip of the iceberg.

<sup>22</sup> On focal points, see Schelling (1981). The interplay between conventions and salience effects has been first put forward by David Lewis (2002). Following Basu (2018), I have argued elsewhere that the functioning of institutions is fully grounded in the existence of focal points (Hédoin 2021).

<sup>23</sup> See Chwe (2003) for a general and enlightening discussion of the role of common knowledge and common belief in social practices. It has been argued that the publicity requirement is too strong for a realistic account of institutions (e.g., Binmore 2008). This objection is mitigated once we realize that believing a proposition  $p$  does not entail that one is *aware* that  $p$ . See Sillari (2008) and Hédoin (2017) for accounts of rule-following behavior and institutions that emphasize this point.

that serves as a common template for expressing disagreements. This is the core insight that emerges from Wittgenstein's account of "language games".<sup>24</sup> Of course, we may perfectly imagine a society resting on a *modus vivendi* as long as the shared institutional resources permit us to strike political compromises. Rawls's account of the democratic form of life assumes that the public domain consists of a conception of justice based on which disagreement over the good can be dealt with. But, as I have argued, we lack a reason to think that these are appropriate accounts of *liberal* societies where disagreement, including on matters of basic justice, is pervasive.

As an alternative, I submit that the public domain that serves as normative foundations for liberal democracies is more satisfactorily captured as a *minimal liberal social world*.<sup>25</sup> A social world can be defined as the set of shared and convergent representations about the nature and the functioning of the world a group of persons is inhabiting. It includes metaphysical beliefs about the existence and the properties of natural and supernatural entities and (causal and non-causal) relationships between them, as well as rules and institutions that governed its social practices. These rules and institutions give rise to normative expectations about the behavior of the member of the social world and, in conjunction with metaphysical beliefs, ground individuals' normative judgments, notably about the good and the right. A social world  $W$  is thus a form of life  $F$ , or the intersection  $\bigcap_x F_x$  of a set of forms of life, constituted by shared social practices and underlying metaphysical and normative beliefs. If we view a society  $S$  as the *union*  $\bigcup_x F_x$  of the same set of forms of life, then  $W$  forms the *domain of public reason* of  $S$ .

My contention is thus that for any society  $S$ , there is a social world  $W$  corresponding to its domain of public reason. Intuitively, the more a society is diverse in terms of metaphysical and normative views, the more relatively narrow its social world will be. If we follow Rawls and many others in characterizing a liberal society as one whose institutions encourage individuals to use their practical reason to experiment with new practices, then we may expect  $W$  to be fairly narrow. But all liberal societies do have a domain of public reason and the minimal social worlds of all liberal societies can be characterized in terms of family resemblance. First, their public domain precisely exhibits the common understanding that social experiments and innovations are beneficial, except for particular reasons that need to be worked out. Second, they tend to rely on political and economic institutions realizing at different degrees and in different configurations the value of negative liberty.<sup>26</sup> Together, these two features indicate that the social world of liberal societies will generally figure a commonly acknowledged principle of "natural liberty" according to which "[w]hatever is not prohibited (and this includes the non-performance of specific acts) is permitted" (Gaus 2019: 187). Third, to realize this principle but also to deal with the incommensurability of goods and the

---

<sup>24</sup> See in particular Wittgenstein (2010: §§ 241-2).

<sup>25</sup> I am here largely following Gaus (2019: Chapter 4).

<sup>26</sup> This includes some of the basic liberties Rawls identifies, though that does not automatically imply the kind of lexicographic priority that justice as fairness ascribes to them. Moreover, as Gray (2002) notes, there is no reason to think that these basic liberties are necessarily compossible. That suggests that from one liberal society to another, different liberties are part of the domain of public reason.

complexity inherent to large societies, the domain of public reason of liberal societies will generally be constituted of rules defining jurisdictional rights, especially property rights, as well as economic and political institutions permitting to measure the value of options on a single scale (market prices, political elections).<sup>27</sup> Fourth, and maybe more contentiously, the social worlds of liberal societies gather around a conception of the person as an entity worth respect, which commands the establishment of fundamental rights and general importance given to human flourishing.

I have only sketched what a minimal liberal social world looks like, but it is already sufficient to contrast it with the Rawlsian well-ordered society. The most significant point is that there is no special reason to think that a conception of justice should be part of the liberal social world. Issues of social justice are to be politically disputed and settled (in the Weberian sense) *based on* some shared domain of public reason, but a conception of justice is unlikely to be part of public reason. This goes in the same direction that a revision of the Rawlsian theoretical apparatus suggested by Quong (2010: 162) regarding the status of the overlapping consensus. Quong suggests that the object of the overlapping consensus is not a political conception of justice, but rather the two fundamental ideas of the conception of society as a fair system of social cooperation and the political conception of citizens as free and equal. That would be on this basis that society should settle on a specific conception of justice. As I shall argue below, while these two ideas may indeed in principle belong to the domain of public reason of a liberal society, they reflect the Rawlsian conception of the democratic form of life and are not part of the minimal liberal social world. Nonetheless, this suggests that we are on the right track by expelling conceptions of justice from the domain of public reason.

Another important insight is that the distinction between imposed and inherent stability that Rawls makes in *TJ* is not more satisfactory than the distinction between *modus vivendi* and the overlapping consensus to understand the normative foundations of liberal societies. Even though we may suppose that liberal societies will rely less on imposed forms of stability than non-liberal ones, some kind of political imposition is unavoidable. This is due to the fact that the liberal social world and a liberal society are not coextensive. Hence, the principle of liberal legitimacy cannot be, as a matter of fact, fully satisfied.<sup>28</sup>

## 7. Liberalism Without Democracy

To recapitulate. Rawls's political liberalism has the project to establish that justice as fairness provides a stable framework to settle issues related to constitutional essentials and matters of basic justice within societies marked by a reasonable pluralism of comprehensive doctrines. The argument relies on the characterization of the democratic form of life in which people motivated by a sense of justice use public reasons to debate within the public political forum. The content of public reason is given by justice as fairness as a political conception, as well as

---

<sup>27</sup> We may also add that these institutions tend to satisfy at least up to some degree a requirement of impersonality.

<sup>28</sup> See Klosko (2015) for a similar point.

by a set of guidelines and norms related to a duty of civility and the principle of liberal legitimacy. The democratic form of life is ultimately related to the endorsement in social practices of the two fundamental ideas of citizens as free and equal and the conception of society as a fair system of cooperation. I have argued that this conception of the democratic form of life is unsatisfactory as far as it is thought to reflect – in the perspective of the Wittgensteinian step Rawls has taken – the working of actual liberal democracies. I suggest that the latter are more appropriately captured by the idea of the minimal liberal social world.

The difficulties I have identified in Rawls's account point toward a more general issue. Considered with the bigger picture of the whole history of philosophical, political, and economic theories of liberalism in mind, Rawls's political liberalism sets itself apart by arguing for the symbiotic relationship between democracy and liberalism. As pointed out by John Skorupski (2017) in an insightful article, Rawls's approach is in this respect very different from the one of Jon Stuart Mill and other "old liberals":

"In contrast to Mill (and even more so to some other old liberals), Rawls tends to see liberalism and democracy as expressions of essentially the same values, as if one could describe some single late modern progressive standpoint indifferently as 'liberal' or 'democratic'. At other times, a little differently, he presents his project as an attempt to rethink liberalism from within a democratic standpoint" (Skorupski 2017: 175).

This is an important point that justifies speaking of Rawls's conception of the "democratic" rather than "liberal" or "democratic liberal" form of life. As I have emphasized, Rawls repeatedly states that his political conception of justice (and at the bottom, the very political/comprehensive distinction) as well as his interpretation of the idea of public reason are grounded in the "democratic political culture". The *liberal* principle of legitimacy is itself tightly connected to the democratic political culture. The point is that there is no indication in Rawls's writings of how we could even think of a liberal society that is not democratic, or conversely, of a democratic regime with non-liberal values and principles.

There are two possible follow-ups from this observation. The first is to conclude that Rawls's political liberalism is at the bottom based on a misconception. As we go further in the analytical structure of its theoretical apparatus, this misconception reveals itself through the problems I have identified above. Political liberalism is not "political" and inadequately deals with the pluralism and the pervasiveness of disagreement that motivated its development from the outset. There is however a second (not necessarily exclusive with the first) reading available, also suggested by Skorupski (2017: 181): this suggests that justice as fairness is not justified as a "political" conception able to constitute the object of an overlapping consensus between conflicting reasonable comprehensive doctrines, but rather is itself a comprehensive view of the democratic form of life, a view that can be called "radical democracy". This may be particularly significant regarding the interpretation of the idea of the political conception of persons as free and equal citizens:

"The slogan that citizens are free and equal conceals significant differences. There is, in particular, a clear difference between what an old liberal means by that slogan and what a radical-democrat means. For the old liberal it expresses the fundamental ethical

principle of impartial individualism. At the political level, it means equality under the rule of law which secures equal civil rights to all. These tenets do not yield democracy; for that, further argument is required. For a radical-democrat, in contrast, the slogan refers from the start, to an equal political say, implied by the very idea of a general and sovereign will of all the people” (Skorupski 2017: 181).

The “comprehensive” dimension of Rawls’s account also transpires with respect to the second fundamental idea, that of society as a fair system of social cooperation. The choice to emphasize cooperation compared to other constitutive features of liberal and democratic such as conflict, competition, and rivalry is not benign. It suggests that fairness considerations only cover cooperative interactions; what about non-cooperative interactions, however? It can be answered plausibly that in well-functioning liberal democratic societies, cooperative considerations are always present, even when interactions are also conflictual.<sup>29</sup> Moreover, Rawls does not ignore conflict altogether, as otherwise the stability problem would not be so acute in his framework. But the tremendous importance given to the idea (or even ideal) of cooperation becomes apparent when we consider the rationale that Rawls provides for the difference principle in Section 13 of *TJ*. There, the idea of cooperation together with the initial state of equality do all the work to support the claim that inequalities are justified as long as they work to the advantage of the less well-off. Transposed into the theoretical apparatus of *PL*, what was a constitutive principle of a comprehensive doctrine becomes a foundational idea of a political conception. This is not only problematic from the perspective of Rawls’s political turn; this directly leads to the tension already emphasized between equating the domain of the political as the one where agreement prevails on matters of justice and still entertaining the possibility that several political conceptions may be in competition.

This points out therefore that Rawls’s account of the democratic form of life is comprehensive in his very own sense of the term. This should not surprise us because I have suggested above that the distinction between the political and the comprehensive in Rawls’s sense is dubious. Once this is acknowledged, we can develop the same reasoning for all other key concepts of Rawls’s political liberalism.<sup>30</sup> It would then appear that Rawls’s conception of the idea of public reason, and especially his account of the duty of civility, makes sense from a radical-democratic viewpoint, but fits uneasily in a liberal outlook where restrictions on public discourse seem to go against the principle of natural liberty that – I have argued at least – belongs to the minimal liberal social world. If we push the reasoning to its end, we may be tempted to conclude that Rawls’s radical democratic understanding of the democratic form of life is not only not “political”, but also not *liberal*.<sup>31</sup> Instead of pushing for this controversial

---

<sup>29</sup> In economics as well as in political science, it is indeed fairly standard to model economic (market) and political interactions in terms of mixed-motive games, e.g., the prisoner’s dilemma or the hawk-dove game. This mix between cooperative and conflict features is actually not an artefact of liberal democratic societies, but is more generally to be found in the natural and social world (as indicated by the biologists’ use of the very same games).

<sup>30</sup> As nicely done by Skorupski (2017).

<sup>31</sup> This conclusion is made even more plausible once we realize that Rawls’s political turn and idea of public reason are significantly marked by the influence of Rousseau’s political philosophy. See Ege et al. (2022).

(and I think, not ultimately correct) conclusion, I will end this paper by briefly considering the possibility to reframe the problem of stability in terms of “robustness”, especially in the context of polycentric governance.

## 8. Polycentricity and the Robustness of Open Societies

It is time to take stock of what has been said until now. The stability problem has been the main motivation for Rawls’s political turn. The latter builds on an account of the democratic form of life that is thought to prevail in liberal democratic societies. I have argued that the latter is not satisfactory; as a result, Rawls’s political liberalism is not political in some meaningful (and normatively relevant) sense, and maybe even not liberal. Rawls’s ambitious attempt to establish the possibility of liberal democracy therefore fails. As it is however often the case in political philosophy and elsewhere, Rawls’s attempt is insightful *especially because of its failure*. More specifically, I shall suggest that beyond the suggestion I have made above to tackle the problem with the idea of the minimal liberal social world, it is the whole approach in terms of “stability” that may be inappropriate. A more fruitful perspective may be to think in terms of the *robustness* of liberal societies. This is especially the case if we consider that liberal societies will generally be open, i.e., they will tolerate or even push for the introduction of new ideas, practices, and more generally ways of life coming from “the outside”, in particular via immigration.

The suggestion that robustness is a more relevant concept than stability to reflect on the kind of issues Rawls was concerned about has been recently made by John Thrasher (2020) in the context of what he calls the “open society model of political contractualism”. As he states, a “system robust is robust when it can continue to function in a stable state in the fact of some range of perturbations” (Thrasher 2020: 1148). It is thus similar to the definition of stability in dynamic systems, where an equilibrium is said to be stable when any infinitesimal deviation is self-correcting, bringing the system back to its initial state. He goes on to argue that diversity, far from being a problem, is quite the contrary an asset for the robustness of society, provided that it is organized to ripe the benefits that diversity generates. Liberal societies are, it is argued, precisely those societies whose constitutional structure is the most likely to satisfy this requirement.

Though I broadly agree with this view, I would push for a slightly different approach, and more particularly for a stronger concept of robustness. In addition to being organized such as it settles on a dynamically stable state, we may want that society favors practices that over time strengthen, or at least do not weaken, its dynamic stability. This will in particular be the case if its institutions are not *self-undermining* in the following sense.<sup>32</sup> Consider a set of rules

---

<sup>32</sup> This analytical conception of self-undermining institutions is very close to the one developed by Avner Greif (2006) to account for the unravelling of economic and political institutions, despite the fact that they successfully solved coordination and collective action problems in the first place. In Greif’s game-theoretic framework, mechanisms making an institution self-undermining are captured by “quasi-parameters” whose values change as the interaction proceeds, without the players being aware of and forming expectations about them.

$R$  applying in a range of circumstances  $C$ . This defines a social practice  $P(R, C)$ . Following the account of institutions and rule-following behavior I have put forward in section 6 above, this practice corresponds to an equilibrium that – because we are considering a dynamic setting – is dynamically stable. That means that participants in the practice have indeed all-things-considered reasons to follow  $R$  and that the practice is robust in face of deviant behaviors, provided that they remain exceptional. Denote  $\theta = (\theta_1, \dots, \theta_n)$  the set of parameters formally describing the social and natural environment in which the practice is taking place. These parameters are independent of  $P$  and here are assumed to be exogenous. They may reflect, for instance, the available quantity of some natural resources, the political regime, the degree of protection of property rights, and so on. Next to these exogenous parameters, we also have a set of *quasi-parameters*  $\phi^P = (\phi_1^P, \dots, \phi_m^P)$  whose value is a function of the practice. In a dynamic setting, the value may change as time unfolds even though the practice remains the same. For instance, the stock of a specific natural resource may decrease through time because of  $P$ . Finally, denote  $\theta$  and  $\theta^*$  the full range of values  $\theta$  can take and the subset of values of  $\theta$  for which the practice corresponds to a stable equilibrium respectively. We say that a practice  $P$  is self-undermining exactly when it makes the values of  $\phi^P$  evolve such that  $\theta^*$  is shrinking. At the limit,  $P$  may at some point in time no longer be stable for any value of  $\theta$ .

Why is this stronger concept of robustness relevant? A liberal society, even more if it is open, will favor the development of a great diversity of ways of life and practices. By the principle of natural liberty, there is a presumption in favor of novelty in the sense that everything that is not explicitly prohibited is permitted. Obviously, what is an exogenous parameter for a practice  $P$  can be a quasi-parameter for a practice  $P'$ . That means that any institution can be indirectly affected by the development of a new way of life because the latter will change the social and natural environment in which the institution has settled. Self-undermining institutions are particularly fragile because they not only potentially weaken the branch on which they sit; they are also increasingly vulnerable to changes in the environment introduced by new institutions.<sup>33</sup> Non liberal/open societies may be less concerned by this problem because, in principle, they impose greater control on the emergence of new ways of life – at the expense of the benefits brought by social innovation and exploration. But liberal and open societies cannot ignore it. In a way, the problem of strong robustness is indeed at the root of the difficulties faced by contemporary Western liberal democracies. As Rawls, following Berlin, acknowledges, there is a limit in the range of forms of life that a society can make cohabiting.

The idea of polycentric governance acquires special relevance in the context of the strong robustness problem. Intuitively, a partial solution to it is to proceed to “compartmentalization” by restricting as much as possible the overlap between exogenous parameters (for a practice  $P$ ) and quasi-parameters (for a practice  $P'$ ). This compartmentalization cannot be obviously complete because of public goods and other externalities that cannot be fully privatized and internalized. Nonetheless, in the context of a liberal social world defined by the minimal elements I have outlined in section 6 and eventually to others specific to the liberal society under scrutiny, polycentric governance aims at robustness by favoring a diversity of political

---

<sup>33</sup> Obviously, institutions that are not self-undermining are also potentially vulnerable.

and economic social practices while making them as much as independent as possible.<sup>34</sup> Moreover, by favoring diversity, a polycentric liberal society lessens the risk of settling on a majority of self-undermining institutions, acknowledging that the self-undermining property is often hard to spot before the institution is already declining.

Note two interesting implications that follow from the endorsement of polycentric governance as a partial solution to the strong robustness problem. The first is the role that competition plays in a polycentric liberal society. I have noted in the preceding section that Rawls's account of the democratic form of life and thus of liberal societies is emphasizing the cooperative dimension. This is not to say that cooperation would not be central in a polycentric liberal society but that indicates that the stability and robustness of the social and political order are also fostered by political organization of the competition between forms of life. As a result, the latter idea would be part of the domain of public reason of this society. The second implication follows from the necessity to disentangle liberalism and democracy. *Prima facie*, there is no reason to consider that political regimes belonging to the same polycentric order should all be democratic, neither in the Rawlsian radical-democratic sense, nor in the more formal sense following from the "one person, one vote" rule. John Stuart Mill's (2015) plural voting system or recent proposals about "quadratic voting" (Posner and Weyl 2018) could for instance be experimented with, as they are not in contradiction with the minimal liberal social world. More controversially, various forms of "epistocratic" mechanisms allocating political power according to individuals' knowledge and competencies (Brennan 2016) could also be used, as long as individuals are effectively free at a sufficiently low and foreseeable cost to switch to a different political unit. At least, the liberal domain of public reason does not seem to rule out a priori this kind of proposal, even though they may afterward prove to be unacceptable or unattractive.

## 9. Conclusion

In this paper, I have outlined and critically assessed Rawls's account of the democratic form of life, in light of his treatment of the stability problem within his political liberalism. I have argued that this account reveals what I have called a "Wittgensteinian step" in Rawls's political philosophy. This is a step in the right direction, as far as we want to achieve a better understanding of the social roots of liberal political morality.

Though it is in the right direction, the step is also tiny. Rawls's account of the democratic form of life is inappropriate, as far as it fails to acknowledge the full implications of the kind of pluralism that it favors and that Rawls himself identifies. The roots of the problem are to be found in a conception of the political that puts too much emphasis on agreement and too little on the conflictual dimension and, more generally, in the suspect political/comprehensive distinction. Rawls's account of the democratic form of life reveals an authentic comprehensive,

---

<sup>34</sup> See the insightful discussion of polycentric democracy proposed by Julian Müller (2019). This analysis starts from similar considerations as mine but is keener on remaining committed to democratic political institutions.

radical democratic, view. Ultimately, not only Rawls's political liberalism is not political in a full sense, but some may also argue that it is not appropriately liberal.

Rawls's attempt to establish the possibility of just liberal democracies is thus a failure, but a valuable one. It fails because it does not solve the stability problem with respect to the properties of actual liberal democracies. His model of the well-ordered society has therefore few things to say on contemporary Western liberal democracies and their troubles. The failure is however valuable because it points toward a new direction of research: the issue of the robustness of liberal societies, and the related relevance of polycentric governance. This issue I have barely touched. It calls for more work, as part of a post-Rawlsian endeavor to account for the normative foundations of contemporary liberal societies.

## References

- Aoki, Masahiko. 2001. *Toward a Comparative Institutional Analysis*. MIT Press.
- Basu, Kaushik. 2018. *The Republic of Beliefs*. Princeton University Press.
- Berlin, Isaiah. 2000. *The Proper Study of Mankind: An Anthology of Essays*. Farrar, Straus and Giroux.
- Bicchieri, Cristina. 2005. *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press.
- — —. 2016. *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. New York, NY: OUP USA.
- Binmore, Ken. 2008. "Do Conventions Need to Be Common Knowledge?" *Topoi* 27 (1): 17–27.
- Bowles, Samuel. 2016. *The Moral Economy: Why Good Incentives Are No Substitute for Good Citizens*. New Haven ; London: Yale University Press.
- Bowles, Samuel, and Herbert Gintis. 2011. *A Cooperative Species: Human Reciprocity and Its Evolution*. Princeton University Press.
- Brennan, Jason. 2016. *Against Democracy*. Princeton: Princeton University Press.
- Cherniss, Joshua L. 2021. *Liberalism in Dark Times: The Liberal Ethos in the Twentieth Century*. Princeton University Press.
- Chung, Hun. 2019. "The Instability of John Rawls's 'Stability for the Right Reasons.'" *Episteme* 16 (1): 1–17.
- — —. 2020. "The Well-Ordered Society under Crisis: A Formal Analysis of Public Reason vs. Convergence Discourse." *American Journal of Political Science* 64 (1): 82–101.
- Chwe, Michael Suk-Young. 2003. *Rational Ritual: Culture, Coordination, and Common Knowledge*. Princeton University Press.
- Cohen, G. A. 2008. *Rescuing Justice and Equality*. Harvard University Press.
- Ege, Ragip, Igersheim, Herrade, and Benoît Walraevens. 2022. "From the 'Sense of Justice' to 'Public Reason'. A reading of Rawls through Smith, Rousseau and Kant", *working paper*.
- Freeman, Samuel. 2006. *Justice and the Social Contract: Essays on Rawlsian Political Philosophy*. Oxford; New York: OUP USA.
- Gaus, Gerald. 2003. *Contemporary Theories of Liberalism*. London ; Thousand Oaks, Calif: SAGE Publications Ltd.
- — —. 2011. "A Tale of Two Sets: Public Reason in Equilibrium." *Public Affairs Quarterly* 25 (4): 305–25.
- — —. 2019. *The Tyranny of the Ideal: Justice in a Diverse Society*. Princeton University Press.

- Gaus, Gerald, and Chad Van Schoelandt. 2017. "Consensus on What? Convergence for What? Four Models of Political Liberalism." *Ethics* 128 (1): 145–72.
- Gintis, Herbert. 2009. *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*. Princeton University Press.
- Gray, John. 1995. "Agonistic Liberalism." *Social Philosophy and Policy* 12 (1): 111–35.
- . 2002. *Two Faces of Liberalism*. Oxford: The New Press.
- . 2020. *Isaiah Berlin: An Interpretation of His Thought*. *Isaiah Berlin*. Princeton University Press.
- Greif, Avner. 2006. *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge University Press.
- Guala, Francesco. 2016. *Understanding Institutions: The Science and Philosophy of Living Together*. Princeton University Press.
- Habermas, Jürgen. 1995. "Reconciliation Through the Public Use of Reason: Remarks on John Rawls's Political Liberalism." *The Journal of Philosophy* 92 (3): 109–31.
- Hadfield, Gillian K., and Stephen Macedo. 2012. "Rational Reasonableness: Toward a Positive Theory of Public Reason." *The Law & Ethics of Human Rights* 6 (1): 7–46.
- Hédoin, Cyril. 2014. "A Framework for Community-Based Salience: Common Knowledge, Common Understanding and Community-Membership." *Economics and Philosophy* 30 (03): 365–95.
- . 2017. "Institutions, Rule-Following and Game Theory." *Economics and Philosophy* 33 (1): 43–72.
- . 2021. "The Beliefs-Rules-Equilibrium Account of Institutions: A Contribution to a Naturalistic Social Ontology." *Journal of Social Ontology* 7 (1): 73–96.
- . 2022. "In Search of a Stable Consensus. Rawls's Model of Public Reason and its Critics." *Working paper*.
- Henrich, Joseph. 2015. *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton University Press.
- Klosko, George. 2015. "Rawls, Weithman, and the Stability of Liberal Democracy." *Res Publica* 21 (3): 235–49.
- Kogelmann, Brian. 2019. "Public Reason's Chaos Theorem." *Episteme* 16 (2): 200–219.
- Kogelmann, Brian, and Stephen G. W. Stich. 2016. "When Public Reason Fails Us: Convergence Discourse as Blood Oath." *American Political Science Review* 110 (4): 717–30.
- Lewis, David K. 2002. *Convention: A Philosophical Study*. John Wiley and Sons.
- Melenovsky, C. M. 2016. "Incentives, Conventionalism, and Constructivism." *Ethics* 126 (3): 549–74.
- Mill, John Stuart. 2015. *On Liberty, Utilitarianism, and Other Essays*. Oxford University Press.
- Mounk, Yascha. 2018. *The People vs. Democracy: Why Our Freedom Is in Danger and How to Save It*. Cambridge, Massachusetts ; London, England: Harvard University Press.
- Müller, Julian F. 2019. *Political Pluralism, Disagreement and Justice: The Case for Polycentric Democracy*. Routledge.
- Posner, Eric A., E. Glen Weyl, Vitalik Buterin, and Jaron Lanier. 2018. *Radical Markets: Uprooting Capitalism and Democracy for a Just Society*. Princeton Oxford: Princeton University Press.
- Quong, Jonathan. 2010. *Liberalism without Perfection*. Oxford ; New York: OUP Oxford.
- . 2018. "Public Reason." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2018.
- Rawls, John. 1971. *A Theory of Justice*. Oxford University Press.

- — —. 1993. *Political Liberalism*. New York: Columbia University Press.
- — —. 1997. "The Idea of Public Reason Revisited." *The University of Chicago Law Review* 64 (3): 765–807.
- Schelling, Thomas C. 1981. *The Strategy of Conflict*. Harvard University Press.
- Sen, Amartya Kumar. 2009. *The Idea of Justice*. Harvard University Press.
- Sillari, Giacomo. 2008. "Common Knowledge and Convention." *Topoi* 27 (1): 29–39.
- Skorupski, John. 2017. "Rawls, Liberalism, and Democracy." *Ethics* 128 (1): 173–98.
- Smith, Professor Vernon L. 2009. *Rationality in Economics: Constructivist and Ecological Forms*. 1st ed. Cambridge University Press.
- Thrasher, John. 2020. "Agreeing to Disagree: Diversity, Political Contractualism, and the Open Society." *The Journal of Politics* 82 (3): 1142–55.
- Thrasher, John, and Kevin Vallier. 2015. "The Fragility of Consensus: Public Reason, Diversity and Stability." *European Journal of Philosophy* 23 (4): 933–54.
- Weber, Max. 2004. *The Vocation Lectures*. Indianapolis: Hackett Publishing Co, Inc.
- Weithman, Paul. 2009. "John Rawls and the Task of Political Philosophy." *The Review of Politics* 71 (1): 113–25.
- — —. 2011. *Why Political Liberalism? On John Rawls's Political Turn*. Oxford; New York: OUP USA.
- Wittgenstein, Ludwig. 2010. *Philosophical Investigations*. John Wiley & Sons.