

<http://doi.org/10.21555/top.v700.2663>

Rawlsianism and What Is Urgent

El rawlsianismo y lo que urge

Enrique Camacho-Beltrán

Estación Noroeste de Investigación y Docencia, Universidad
Nacional Autónoma de México
Mexico

enrique.camacho@comunidad.unam.mx

<https://orcid.org/0000-0001-6876-8460>

Recibido: 19 - 10 - 2022.

Aceptado: 28 - 01 - 2023.

Publicado en línea: 14 - 08 - 2024.

Cómo citar este artículo: Camacho-Beltrán, E. (2024). Rawlsianism and What Is Urgent. *Tópicos (México)*, 70, 39-70. DOI: <http://doi.org/10.21555/top.v700.2663>.



This work is licensed under a Creative Commons Attribution
-NonCommercial-ShareAlike 4.0 International License.

Abstract

In *Justice as Fairness: A Restatement*, Rawls asserted that his conception of justice should provide guidance for a non-ideal theory of real-world injustices. Nonetheless, he established at the same time an enigmatic limit to his framing in which neither political philosophy nor justice as fairness should be understood as applied approaches. It is unclear how guidance against real-world injustices is possible, since it would presumably require insights into applied ethics. This paper focuses on drafting a possible alternative by offering a partially comprehensive reconstruction of the reasonable as a virtue. The hope is to use the virtue of reasonable citizens to transfer Rawlsian ponderations into a special kind of applied ethics controversies concerned with urgent public debates and resistance politics. The idea of this construction is to allow the virtue of the reasonable to establish a standard for citizen conduct.

Keywords: Rawls; reasonability; political conception of justice; approaches in applied ethics.

Resumen

En *Justice as Fairness: A Restatement*, Rawls afirmó que su concepción de la justicia debería servir de guía para una teoría no ideal de las injusticias del mundo real. No obstante, estableció al mismo tiempo un enigmático límite a su planteamiento, según el cual ni la filosofía política ni la justicia como equidad deben entenderse como enfoques aplicados. Como resultado, no queda claro cómo podría fungir como orientación contra las injusticias del mundo real, ya que presumiblemente se requeriría de conocimientos de ética aplicada. Este artículo se centra en esbozar una posible alternativa ofreciendo una reconstrucción parcialmente exhaustiva de lo razonable como virtud. La esperanza es utilizar la virtud de los ciudadanos razonables para transferir las reflexiones rawlsianas a un tipo especial de controversias de ética aplicada relacionadas con debates públicos urgentes y políticas de resistencia. La idea de esta construcción es permitir que la virtud de lo razonable establezca una norma para la conducta ciudadana.

Palabras clave: Rawls; razonabilidad; concepción política de la justicia; aproximaciones en la ética aplicada.

1. Introduction¹

In the first part of *Justice as Fairness: A Restatement*, Rawls shocks with an interesting warning: according to him, neither political philosophy in general nor justice as fairness in particular should be used in any way to tackle problems of applied ethics (2001, p. 14). In this vein, justice as fairness seems to be understood as an effort to simplify and reduce the structure of political reality in an abstract-theoretical model we call *ideal theory* (the realistic utopia of distributive justice for a well-ordered society composed by fully compliant citizens).²

Interestingly, most problems of applied ethics (such as migration, discrimination, corruption, etc.) cannot be properly raised within this understanding of ideal theory, since they simply disappear when applying the restrictions introduced by the model: citizens are not corrupt since they have a sense of justice, they do not discriminate as society is made of equals, etc. In contrast, many problems of applied ethics belong to what Rawls calls *non-ideal theory*, which assess what justice requires when the background institutions of a society are unjust and citizens do not fully comply with the principles of justice (1993, pp. 5 & 90).

However, according to Rawls, non-ideal theory “is worked out after an ideal conception of justice has been chosen” (1971, p. 216), because without a solid ideal outlook, non-ideal theory wouldn’t have a clear objective to pursue (Rawls, 1993, p. 90). In this way, “the idea of a well-ordered society should also provide *some guidance* in thinking about

¹ This study is indebted to the POLETH seminar series at Mexico City. I am also grateful to Itzel Mayans, Jesús Rodríguez Zepeda, Imer Flores, Andrés Moles, Luis Muñoz Oliveira, and Moisés Vaca Paniagua for their insightful comments, objections and suggestions to previous versions of this paper. I benefitted from audiences of undergraduate courses at the UNAM Faculty of Philosophy and Universidad Panamericana. I am grateful to Gabriela Sofía Llanos and Binnui Navarro Romo for their perceptive comments. Finally, many thanks to Jesús Alvarado for his editorial and research support, as well as to Francisco García González for making this paper readable in English.

² See Rawls (2002, pp. 4-10). For a persuasive account of the limits of ideal theory regarding applied problems such as discrimination, see Vaca Paniagua (2022) and Valentini (2012).

non-ideal theory [...]” (Rawls, 2001, p. 13). An extensive literature tackles the suitability of this distinction and how useful or illuminating it can be to the overall field (Stemplowska, 2016, Cohen, 2008; Miller, 2011 Valentini, 2012); but I will not engage in these debates. Instead, the problem I will pick up here is that it is not at all clear in which way and to what extension or capacity ideal theory could perform this *guiding role* for non-ideal theory.³

Ideal theory should help evaluate different kinds of policies and laws to address injustices (Fuller, 2012, p. 312). But Rawls does not provide any guidance on how justice as fairness could cross the threshold of non-ideal theory. As Rawls puts it, “the measure of departures from the ideal is left importantly to intuition” (1971, p. 216). With this in mind, in this paper I will sketch a strategy for constructing an approach that is able to transfer knowledge from ideal theory into non-ideal theory, thus guiding intuitions in at least a specific area of urgent public debates on applied ethics. It is in these public debates that I believe the Rawlsian virtue of the reasonable can be introduced and be particularly explanatory.⁴

In this paper I am going to argue we can use the virtue of reasonableness as a standard of citizen epistemic conduct, which establishes the requirements of care towards the rest of the citizenry. Failing to fulfil that standard of reasonable citizen conduct may allow us to establish a case of culpable ignorance, when such ignorance harms specific groups of vulnerable people. In turn, this may allow us to establish the plausibility of assigning responsibilities to people who violate the standard of epistemic conduct that constitutes the virtue of what is reasonable.

This paper has the following sections. In the second section, I will invoke the core of the Rawlsian account of reasonableness. In order to do that, I will recall Lief Wenar’s distinction between two stages of analysis within Rawls’ account. In the third section, I will highlight the general difficulties of establishing a standard of citizen conduct to be used in the

³ I’m grateful to an anonymous reviewer for pressing me on this matter so I could clarify the problem.

⁴ I do not use the adjective “urgent” in a deep or technical way. The term is used as a general reference to public debates with the kind of scope that is significant and broad enough in allowing groups traditionally subject to oppressive relationships to be treated more and more as free and equal citizens.

reality of urgent public controversies. At the same time, I will suggest invoking the controversial traits of reasonableness in Rawls' account. The fourth section will allow me to argue how the Rawlsian virtue of the reasonable can be used to determine the plausibility of moral change over a public controversy in a given context. Finally, I address two objections to my proposal: first, I question whether my account is in any way explanatory or useful, or rather if it is already analytically contained in Rawls own clues about how ideal theory should guide non-ideal theorizing. Also, the opposite can be challenged: how do you know if *R* in this context is still Rawlsian in any meaningful sense? I believe we may use the core of Rawls' application of the reasonable in the case for pluralism in liberal societies as a guide to ensure *R*'s pedigree in the urgent cases I have mentioned.

Before moving on to the analysis, it is important to introduce some preliminary clarifications. This is not a work on distributive justice and therefore does not have the restrictions of a political conception of justice. Rather, I extract Rawls' conception of the reasonable from its proper ideal distributive context and discuss its application in the domains of applied ethics and epistemic injustices. This means I begin with the idea that, like any moral right, our right to believe things without proper evidence has limits, particularly when we have a tendency to encourage other people to share our beliefs or when our beliefs may harm other people, or even when our beliefs could be instrumental in harming other people. For example, I am allowed to believe that onions are as sweet as ripe bananas. Although many people disagree with that belief, and I have no evidence to support it, my belief does not harm anyone but my culinary reputation and myself. But suppose instead that I believe it is safe to decorate Christmas trees with burning candles, and I convince my family and friends to do so. If someone has an accident involving the violent combustion of their artificial Christmas tree, it may have been my responsibility to ensure the truth of my belief before spreading it.

Second, controversies in applied ethics differ from more abstract philosophical controversies in that they are not just about deciding between abstract principles or values. Controversies in applied ethics also involve controversies about different principles of action. In other words, I assume the discussion at hand is not only about what we should believe, but also about the things we should believe in order to act in accordance with those beliefs. For example, knowing what a reasonable person would believe about candles on Christmas trees

includes knowing what a reasonable person would do about them. So, before recommending decorations to our friends, it seems convenient that we know what a reasonable person would believe about candles on Christmas trees.

Finally, it is also important to note that, to some extent, this is a paper of philosophical methodology. The used examples offer only a frame of reference but are not discussed at length. In addition, it is important to insist that this work is not intended to be a scholastic exercise or scholarly interpretation of Rawls' work. Rather, it is a first sketch of a heuristic or interpretative tool to discuss a possible methodology for the reconstruction of Rawlsian arguments, which apply to issues of applied ethics that are ordinarily considered to be outside the restrictions of the Rawlsian model. The hope then is to draw attention to one part of the approach, so it continues to inspire us.

2. The Rawlsian role of the reasonable

In *Political Liberalism*, Rawls claims reasonable citizens are those who accept two basic normative ideas: the idea that one belongs to society as a cooperative member with free and equal standing, and the idea that everyone's judgment about moral and political judgments is seriously limited (1993, pp. 81-82 & 86).⁵ Only citizens who meet that standard

⁵ Recall that, in *Political Liberalism*, Rawls describes the political conception of justice as a group of principles and ideas of political philosophy designed to regulate — through many generations — the main political and social institutions. It is a case of pure procedural justice that consists of two principles about permissible liberties and inequalities, which are the result of an ideal decision model or procedure articulated through different abstract elements, such as the original position and reflective equilibrium. The decision model envisions free and equal citizens who are cooperative members of a society that in turn is conceived as a liberal and democratic system of social cooperation. This trans-generational stability would be the product of an overlapping consensus of many visions, doctrines and comprehensive theories (1993, pp. 12 & 144-145). Justice as fairness can be the object of this type of consensus only because it is itself a political conception of justice grounded and developed independently of the variety of doctrines or comprehensive theories that citizens may hold (1993, pp. 10-13, 140-141 & 144). It is precisely this type of ideas of political philosophy which Rawls claims for no reason should be used in problems of applied ethics (2001, p. 14). The reason is that justice as fairness has a single abstract or ideal purpose, which is to establish the content of the political conception

of conduct can effectively interact politically under pluralistic social conditions in the manner appropriate to make political consensus stable and lasting (1993, p. 41). Conversely, my claim is reasonableness is not just a corollary of the liaison between individual moral standing and the moral character of society: it may also be constructed as a virtue for guiding epistemic conduct in certain cases of applied ethics.

Here I am thinking about cases of decisive public debates with the potential to change the landscape of the structure of a political regime, because they represent substantive changes in the moral conscience of citizens. Take, for instance, historical cases: at what point was it plausible to attribute epistemic responsibilities to English citizens who did not change their minds in the face of evidence presented by the suffragettes and by John Stuart Mill, explaining the moral requirement that women be given the right to vote? Did the English of the late 19th century have a responsibility to be reasonable and consider the evidence of the suffragettes?⁶ Can we identify unreasonable citizens and plausibly hold them accountable for being oblivious and ignorant of the kinds of things they should have known about people's rights to fulfil their duties of care to other citizens? I will argue we can, at least on the level of epistemic conduct, but first, in this section, it is worth recalling Rawls' own characterization of reasonableness.

According to Lief Wenar and others, Rawls employs the virtue of reasonableness at least in two distinct segments of analysis (Caney, 2012, pp. 273-283, n. 44 & p. 294; Wall, 2013, p. 429; Doorn, 2009, p. 131; Wenar, 1995).⁷ He of course famously underscores *the content* of justice as fairness as a conception of justice comprised only by political values in a way that it gathers its justification independently of any comprehensive or metaphysical doctrine (1993, pp. 10-13 & 154-157). At that slice of analysis, reasonableness mainly explains the connection between the conception of persons as free and equal and the conception of society as a well-ordered one. But, importantly for my purposes here, at a different

of justice: going from values, ideas, and conceptions, to principles, all through the restrictions of method (I have mentioned this in the introduction) in order to ground principles of justice of a pure procedural case for basic institutions.

⁶ See Mill (2018).

⁷ Note that the distinction between levels of analysis and the concepts used to distinguish them varies with each author. In this paper, for simplicity, I prefer Wenar's twofold distinction.

stage of the argument he engages in *the presentation* of certain traits of reasonableness in order to show justice as fairness as a stable conception of justice across generations.

At that stage, reasonableness serves the purpose of supporting the plausibility of the conception of justice in more practical terms by explaining why overlapping consensus is necessary, how it may be achieved, and how it may persist (Wenar, 1995, p. 33). Importantly, under this interpretation, the political conception could be used in an applied manner, but exclusively in a single case, which is to serve as “[...] at best but a guiding framework of deliberation and reflection which helps us reach political agreements on at least the constitutional essentials and the basic questions of justice.” According to this, parallel to the theoretical objective of *the content*, Rawls could be pursuing the practical purpose of articulating *the presentation* of the standards that make possible the consensus on the basic structure; in other words, to show how *the content* can truly be a stable political conception (1971, p. 75; Wenar, 1995, pp. 32-33).⁸

Recall that a reasonable comprehensive theory or doctrine is one that is the product of the exercise of theoretical and practical reason within a tradition of thought; its primary task is to organize and characterize outstanding values and principles of society (Rawls, 1993, p. 59). This means that a comprehensive doctrine conveys an intelligible view of the world that claims to be, all things considered, the correct outlook, and it is thus clearly distinguished from other rival theories or doctrines which are wrong. As we will discuss later, one single comprehensive doctrine cannot be the focus of an overlapping consensus around a political conception of justice, as citizens defend various, mutually exclusive, such doctrines. So, saying that Rawls introduces partially comprehensive elements means he employs fragments of comprehensive doctrines in

⁸ From now on, I am going to refer to these different levels of Rawlsian argumentation as *the content* and *the presentation*. It must be remembered that this distinction is an interpretation Wenar loosely bases on the difference between procedural justice and pure procedural justice, but it is not a distinction Rawls explicitly makes. It is certainly a controversial interpretation and it may be rejectable. But here it is attractive because it seeks to shield the political conception of justice from certain attacks, it being the most important normative core for Rawls.

order to explain the plausibility of overlapping consensus, even if that defence itself cannot be the object of such consensus.

If this is true, maybe we can theoretically separate what it means to be reasonable from the point of view of a conception of justice, from the more broadly questions of what, in reality, plausible morality could require from a reasonable person and how is it possible to be in fact reasonable. The presence of elements of a comprehensive theory, even if integrated in a partial form, will be helpful for us in order to extract a conception of reasonableness for discussions of applied ethics. I will say more about this in the next section. But first, let's see whether the distinction is convincing and we can separate the merely political elements from the partially comprehensive particles of reasonableness. The following traits (i)-(iv) are part of *the content*—I will invoke traits (v)-(vii) in the next section.⁹

Rawls describes the epistemic standard of conduct of reasonable citizens through the following features (1993, pp. 15-35 & 81). First, the central feature of reasonable citizens is that:

- (i) they understand the paramount importance of *social cooperation* for social life, so that they are willing to be, or at least to become, as fully as possible cooperative members of the society (see Rawls, 1993, p. 49).

For this, citizens must establish equal relations of dialogue and deliberation with the rest of the citizenry. Rawls apparently believed that this is only possible when citizens also display a set of psychological and intellectual virtues. Thus, citizens must also be able to:

- (ii) exercise at least the *basic competencies of judgment, thought and inference*. For this they must also consider themselves fallible (see Rawls, 1993, p. 82).

This means that it is necessary to interpret the evidence and the judgments collectively, accepting what Rawls calls the burdens of judgment, that is:

⁹ Wenar uses this distinction to show that the aim in *Political Liberalism* of providing a free-standing political conception of justice is incoherent. However, in this paper I will remain agnostic regarding that particular criticism.

- (iii) the challenges, difficulties and *burdens of the ordinary exercise of the judgment*, in matters that imply the interpretation of the relations between facts and values (1993, pp. 55-59).

Finally, also:

- (iv) despite all these burdens, each citizen must develop for herself her own *sense of what is good and fair* (see Rawls, 1993, p. 52).

Now notice that, according to the distinction I invoked above between what Rawls says about *the content* of a conception of justice and what he claims for his *presentation* as stable, requirements (i)-(iv) are sufficient to understand what is reasonable as a political standard of citizen epistemic conduct (Wenar, 1995, pp. 38-39). This is because these conditions establish what Rawls means for the reasonable as a practical and normative virtue and not only as an epistemic virtue: citizens are reasonable when they have the set of psychological and mental competencies sufficient for a cooperative disposition and behaviour with respect to other citizens—first, in *the way they form beliefs*, and then in the *principles of action* related to those beliefs.

The quality of beliefs is particularly regulated by *the burdens of judgment* that rule out the adverse effect of biases caused by positionality and upbringing. The reason is that, as I explain next, burdens of judgment account for the kind of fallibility our comprehensive moral theories face. So, it seems this restricted characterization also establishes the capacity for affirming reasonable comprehensive doctrines, since burdens of judgement restrict citizens' beliefs.

The burdens of judgment require that citizens, when forming their beliefs, consider that (a) the available evidence is often contradictory and complex, (b) it is difficult to know how to weigh different considerations, (c) our concepts are often vague and we are not always certain how to interpret them, (d) we are highly vulnerable to our biases when evaluating evidence, (e) it is difficult to make a final judgment on a matter if we have normative considerations pulling in opposite directions, and (f) each society selects the most important values among many that can be selected (Rawls, 1993, pp. 55-67).

Now, perhaps (i)-(iv) could be sufficient to characterize reasonableness; the problem is that (i)-(iv) explain *what reasonable*

citizens are like, but they do not explain *the motivation that people might have to become reasonable and continue to be so* throughout their lives (Rawls, 1993, p. 84). This condition is important if we want a conception of reasonableness that shows not merely conceptual coherence, but also practical relevance.

In order to see this, consider the case of John and how he can use (i)-(iii) to revise his judgments. John is a prosperous slave trader and, during his time, he may believe he is doing his part as a member of his society (i). Yet, during John's voyages, the conditions are such that under a third of the enslaved people die, suffer torture, rape and other aggressions with John's consent or sometimes at his own hands. Does this mean John fails in (iv) and does not have a sense of justice and morality? John is not a psychopath or an amoral person; that is, the damage John does to the enslaved people is not the product of perversity or lack of empathy. But for John and his contemporaries, what is obvious to us is not at all transparent: slavery is repulsive and morally unacceptable. That explains why, at the same time, John is an ardent man of faith racked with disproportionate guilt for being foul-mouthed. Every time John has escaped from the clutches of death, he has interpreted that fate as a sign of God's love, who saves him so he can lead a more Christian life. This is why John is constantly trying to be a better Christian; but in his attempts throughout his life, John has never felt the slightest remorse for the dreadful faith of the enslaved people under his command.¹⁰ So John has a sense of morality and justice, albeit his morality does not consider the moral standing of enslaved people.

What about (ii) and (iii)? Is John failing to access information and consider existing evidence? In John's day, people were sceptical of the radical experiment in free labour. Three hundred years ago, for most white Europeans, slavery was perfectly moral and just: part of the natural order of things (Anderson, 2014, p. 2). Furthermore, most educated people were sure regular people would not work unless forced to (Anderson, 2014, p. 15). So, it seems John cannot be accused easily of failing (ii) and (iii). According to the moral and scientific tradition of the time, without forced labour, the economy would collapse, and both enslaved people and salaried workers would fall into savagery: reduced in the best of cases to self-sustaining production and vagrancy. Taken

¹⁰ This example is based on the life of John Newton. Elizabeth Anderson (2014) takes it from Newton's own autobiographical accounts (1788 & 1830).

further, without forced labour, there would be no wealth creation and, as a result, civilization would cease to exist.

So, should we simply conclude that John is as reasonable as he could have been at the time? One alternative is to reconsider the conception of reasonableness to see if we can demand more from John. Indeed, as we saw, Wenar claims that, at least in *Liberalism*, one central role of reasonableness is to help explain how reasonable an abstract person needs to be in order to choose between conceptions of justice that are already reasonable. But perhaps a conception like this fails if what we want to identify is the journey of reasonable people who took us from John's perspective at one time in history, to our current *ethos* that clearly sees slavery as morally impermissible. Accordingly, let's consider in the following section the other traits of the reasonable that Wenar believes introduce a partially comprehensive element into Rawls' account.

3. Finding the contours of the problem

Before discussing the remaining elements of reasonableness, it is worth noticing important contextual elements of John's times. In terms of moral consciousness, John's times were very demanding because the legitimacy of forced labour had just begun to be manifestly questioned. Indeed, some 250 years ago, some enlightened groups began to question the validity of the laws that allowed slavery.¹¹ The public needed to become aware and consider the arguments above in order to question slavery. But notably, at the time the sugar producers that refrained from using forced labour had to compete with those that did not. This competition was illegitimate, since production with forced labour was obviously cheaper, at least in the short run. In the face of it, the general public needed to weigh the facts as they understood them, without the refinements of our current philosophical, social and economic theories,

¹¹ According to Anderson (2014, p. 17), two types of arguments predominated. Arguments of the first type stressed that it was morally impermissible for society to use enslaved people's bodies and their lives to serve the ends of civilization. Others, such as Adam Smith, highlighted that the premise from which the slaveholders started was incorrect: enslaved people and forced labourers (and almost anyone) would not work freely if the pay was not enough to cover their subsistence and if they lacked some free time to enjoy the pay. But if workers were to receive sufficient pay and time to significantly improve their living conditions, then most will work voluntarily.

and side with either for or against forced labour.¹² It seems to me that, in order to accept the moral weight of the reformists' speculations about free labour, John and his contemporaries needed to exhibit a special moral and epistemic virtue: they were required to be reasonable; that is, they needed to exhibit a special kind of virtue of judgment that generally enables anyone to mediate between the teachings of their moral tradition, that at the time validated slavery, and the innovations that the progressive people suggest, in this case facts and projections about free labour.

But for the majority of people to be reasonable in this way and for allowing such reasonable people to effectively undertake a sensible change of moral conscience, a great deal of time is required, allowing for trials and errors that cost lives and much suffering.¹³ So we might ask ourselves, wouldn't it be better if this process of trial and error for moral consciousness were shorter and less painful? Can we learn to be reasonable in the way required to process these debates much more clearly and sooner? Therein lies the urgency: in the hope that these kinds of changes could be perceived and enjoyed in the course of a human lifetime. In the eighteenth century, the urgency for activists, philosophers, and theorists was to make John—alongside his contemporaries—to reasonably and critically reconsider what otherwise his slave-owning tradition told him was perfectly just and morally correct.

This kind of reasonable reconsideration of tradition is no easy task. Being reasonable usually means holding oneself to a certain standard of conduct. In this case, it is very important to note that this presumably would require the use of a standard of epistemic conduct that would go against the established standard of practical conduct: the duty to believe things for which there is more or less adequate evidence. The quality of

¹² Perhaps at the time production of sugar using an enslaved workforce was cheaper than free-labour sugar. However, in contemporary capitalism, under certain circumstances things may be different given considerations of efficiency and compliance. I'm grateful to an anonymous reviewer for this remark.

¹³ Indeed, since the independence of Haiti, which gave strength to emancipatory movements, it took another ninety-three years for the entire continent to be freed from legal slavery. Nevertheless, Mauritania didn't abolish slavery until 1981 and even today there are forms of forced labour involved in the production of many items we consider essential, such as clothing, cell phones and computers.

our beliefs is based on such a standard of conduct: reviewing our beliefs to confirm that they are based on appropriate evidence.

But then a series of interesting questions arise: how can John be reasonable like that? How is it that John can take the necessary precautions to not uncritically adopt his slave-owning tradition, but at the same time to not recklessly adopt any changes proposed to him (this means prudently considering the danger of plunging his society as he knew it into chaos and famine, as we saw the defenders of slavery claimed)? What is needed in order to properly weigh the evidence presented by the defenders of slavery and the reformists? How is it possible to introduce some order into the debate and arrive at a judgment? What kind of reasons should reformers present to John to help him evolve his moral conscience? If John refuses to weigh the evidence, is he responsible for the evil suffered by enslaved people? What standard of conduct could be appropriate for John and his contemporaries?

As we can see, it is not at all clear how we can answer these questions by means of a conception of the virtue of the reasonable (i)-(iv) appropriate only for purely conceptual political controversies about modelled free and equal citizens in ideally ordered societies. In contrast, in a real-life practical case, we know how to identify the appropriate standard of conduct for a sailor or a merchant, because the standard of conduct is determined by the practices related to what a good sailor or a good merchant must do, in a given context, according to the skills and responsibilities of sailors or merchants (Mosqueda Esparza, 2019, p. 59). The ability to learn about the humanity of enslaved people and the normative inconsistencies of slave societies is not specifically present in the role of the good sailor or the good salesman; but perhaps it is part of the role of virtuous citizens.

Here is where Wenar's distinction becomes interesting. If it is true that certain traits of reasonableness introduce a partially comprehensive element to the conception, perhaps we can use them to ask John to be reasonable in a more demanding way. Wenar stresses that Rawls added three other requirements, which, as we will see next, are central to using the virtue of the reasonable as a standard of conduct regarding beliefs. Citizens must also:

- (v) have a notion about what is correct or, in other words, about the *essential aspects of objectivity*.

Also, in order for them to adjust their behaviour to what they have identified as objectively correct, citizens must:

(vi) have an appropriate *moral psychology*.

Finally,

(vii) in accordance with their standards of objectivity and their ability to obey principles, citizens interpret their conception of what is good in life based on a *comprehensive (reasonable) theory*.

With (v) to (vii), Rawls introduces two controversial elements: Objectivity and Moral Psychology. The essential aspects of Objectivity (O) must specify the arrangement of reasons according to relevant principles and criteria in order to assign those reasons to (individual or otherwise corporate) agents as reasons that the agents are required to consider and follow in certain specified circumstances. They are compelled to act on these reasons, whether or not they are persuaded of them; in this way, these reasons thusly assigned will displace the very reasons that the agents have or those that, from their point of view, they believe they have. By introducing a standard of objectivity, it is shown that basic conceptions of justice as fairness, such as the conception of the person, can be articulated through the appropriate use of practical reason (Rawls, 1993, pp. 110-112).

To understand the requirements of moral psychology, one must distinguish between object-dependent desires, such as those we have for food or other people's bodies, and principle-dependent desires, such as those we have for following rules or maxims, such as treating others as we want to be treated. The reasonable principles that regulate how a plurality of agents establish relationships among themselves are those that require certain normative psychological dispositions, such as the inclination to follow rules of morality. According to this, the most important thing is to note that there may be desires dependent on reasonable conceptions. The principles that we wish were the reasons for our actions in this case are those that serve to articulate a political conception of justice (Rawls, 1993, pp. 83-84). The central conception that serves this role is that of a free and equal cooperative citizen. From these elements we obtain a conception of Moral Psychology (MP). Having a reasonable moral psychology is then being the possessor of the desires dependent on the conception of citizen, which implies that "they want

to realize in their person, and have it recognize that they realize, that ideal of citizen" (Rawls, 1993, p. 84).

Now, comparing (i)-(iv) against (v)-(vii), it is relatively easy to see how the first group approaches what it means to be reasonable or what is the nature of a reasonable citizen, whereas the second group tells us more about what a reasonable person does or what is expected from a reasonable person. This, in turn, clarifies why anyone would be inclined to be reasonable, provided it can be a realistic option. If Wenar's interpretation is correct, it would be a flagrant mistake to take both analytic levels as the same. The reason is that, as we saw, in the content, Rawls deals with features derived from the concept of pure procedural justice that have been manipulated by Rawls himself to deliver fair results (the political conception of justice) in the modelled conditions of a liberal democracy (1993, pp. 156-157). In the presentation, Rawls deals with features derived from the principles of justice, which give substantive content to justice as fairness, and which defend it to demonstrate that it can be stable (1971, pp. xv-xxi).¹⁴ But, crucially, to defend stability from sceptics, Rawls seems to believe that he also has to argue in a partially comprehensive way, as when we defend order, cooperation, and democracy from bigots.

With all these conditions, by pairing the nature of reasonableness with the concern about its practical plausibility, the conception of the virtue of the reasonable citizen can be stated as follows.

R: The reasonable is the political and moral virtue that citizens of liberal democracies must display so that justice as fairness can be a stable conception of justice. The reasonable, then, is a standard of conduct that people must meet when they exhibit traits *i-iv*. It is a virtue that people and institutions have the duty to develop because their defence complies with *O* and people regularly exhibit *MP*.¹⁵

¹⁴ I am going to assume this distinction has sufficient interpretive anchorage and it is reasonable to understand Rawls in this way. The full discussion can be found at Wenar (1995). Specialized discussions on specific partially comprehensive issues in Rawls can be found at Mandle & Reidy (2013), particularly in the chapters by Krasnoff, Macleod, Dagger, and Wall.

¹⁵ See Rawls (1971, pp. xv-xxi).

The central idea of introducing *R* to controversies in applied ethics is that identifying what is plausible to expect from citizens makes it possible to identify what we can transform or reform here and now, something that matters to us in many cases of applied ethics. Therefore, in what follows, I will try to explore applications of *R* to a specific case. To do this, in the next section let's consider John's example again and put it in the context of the time in history in which we could appreciate among a significant part of the population the internal conflict between being on one side or the other of the controversy.

4. Employing *R* for problems of applied ethics

I begin by noticing that, just over three hundred years ago, almost no one conceived a social world other than the one that considers the institution of slavery as a fundamental piece of its economic framework (Anderson, 2014, p. 2). Perhaps, all things considered, people at that time can be partially excused from being unaware that slavery is an evil. To understand this, one must consider the distinction between the way in which moral responsibility is attributed in ordinary cases of negligence and non-culpable inadvertence (Mosqueda Esparza, 2019, pp. 58-63). Imagine that you are at the beginning of the 19th century living next to Fulgencio, who suffers a form of pericarditis and is a few days away from having a heart attack. In the garden there is a leafy white willow tree. Finally, Fulgencio suffers a heart attack and dies without you knowing that you could have ground the willow bark and prepared an infusion to save Fulgencio's life. In this case, it is very difficult to attribute responsibility to you for the death of Fulgencio, since for you the symptoms of pericarditis and the curative properties of the white willow go unnoticed and you are not responsible for knowing those properties. In contrast, imagine also that today you are Fulgencio's family doctor. In that case you have responsibilities to know that Fulgencio's symptoms are typical of pericardial disease and that they can be controlled with daily doses of aspirin, which happens to be the active ingredient found in white willow.

So, the evil of slavery may go unnoticed by almost everyone before John's time. This is not to say, of course, that slavery was morally permissible before John's time (in the same way that acetylsalicylic acid was still curative for pericarditis in Fulgencio's time even if no one knew it). That is why you and I may, instead, judge the experience of John and his predecessors from ideal judgments and evaluations—such as

those offered by fundamental human rights—that make slavery morally unacceptable, all things considered. In the same way, today’s doctors can lament the tragic deaths that occurred when the healing properties of white willow went unnoticed.¹⁶

This explains why John is not a sociopath without empathy. John’s reasonableness is fuelled, in principle, by a moral tradition that preceded him and that just then began to be questioned. But the important thing to note is that the excuse of inadvertence is no longer available to John. This is why we place John a moment later, precisely when the radical moral consciousness shift may be possible: that is, between being sure of the legitimacy of slavery at one point in John’s recent past, and some years later, when he can be convinced of the opposite. This is the moment that makes the applied ethics controversy possible and, at the same time, what makes it urgent: in contrast to the years just before John’s life, some 250 years ago, some people and groups began to question the legitimacy of slavery as an institution. The urgency lies precisely in the fact that all those people who previously had no hope, now have a small chance to radically improve their lives.

This process of questioning, refutation, and deliberation led to the beginning of the process of the abolition of slavery in different areas of the Americas around 1888 or even earlier, until it became generalized in almost the entire world by 1981. Thus, in this second stage of John’s life, plausibility assessments changed, precisely because they allow us to assess the life of John and many of his contemporaries as carrying out the demanding task of correcting one’s own moral judgments or transforming one’s moral conscience in the light of new information and in ways more coherent and cohesive of explaining their own values and principles.

What these plausibility assessments allow us is precisely to avoid applying our ideal assessments directly to John’s life, reducing thus his process of moral deliberation and that of his contemporaries to madness, complacency, or ignorance. If we instead make plausibility assessments with the *R* principle, this allows us to introduce a standard of epistemic

¹⁶ We also judge of course from our social context, in which slavery is not an openly recognized or accepted social practice. However, if we investigate well, some of the goods that we enjoy are produced in conditions to a certain extent similar to those of slavery in Newton’s time. For a discrete list of these goods, see Bureau of International Labor Affairs (2022).

conduct to them, which allows us to formulate a normative claim of what is reasonable to expect from them, regardless of our own views (Dewey & Tufts, 1981, p. 216).

In order to see this, let's come back to *R*. As highlighted in the previous section, Wenar (1995, pp. 33-34) proposes that in the presentation Rawls transgresses the limits of a political conception to clarify why overlapping-consensus is necessary, how it can be obtained, and how it can be stable. In Wenar's interpretation, *O* is used by Rawls to show the process of birth and evolution of liberal institutions as a historical necessity, which concludes in the interlocking consensus itself. Wenar (1995, p. 55) suggests Rawls thought that, without an independent theory of objectivity (or without a historical and cultural root as in *Restatement*), the overlapping consensus could not last indefinitely. According to this, for Rawls, both the historical need for pluralism and the need for a modern *ethos* constitute part of a process of historical and moral evolution going from absolutism and reform to contemporary pluralism. Crucially, that process is shown by Rawls as a necessary product of bearing the burdens of judgment, that is to say, an almost inevitable result of the history of the towns of reasonable people (Wenar, 1995, p. 47). In other words, Rawls might be suggesting that stable reasonable pluralism is itself also a necessary result of the history of peoples inhabited by reasonable people over time, or at least that it is an inevitable result of human reason developing into increasingly free institutions (1993, pp. xxiv, 37 & 144). To that end, Wenar suggests Rawls intended to use *MP* to stipulate the material conditions of the human mind, under which overlapping consensus is possible and not utopian (1993, pp. 168-169). In this sense, *MP* fulfils the function of explaining human motivation, which Rawls needs to suggest that the overlapping consensus can be stable. Reasonable people seek to belong to social arrangements that allow them to actualize, in their persons, in increasingly better and fuller ways, the ideals of citizen freedom and equality (1993, p. 84). In this way, as a necessary corollary of *R*, stability will be secured.

In my proposal, on the other hand, the aspiration is more modest, but it is of the same kind. Overlapping consensus is a decision model. As we have seen, through the presentation Rawls hopes to demonstrate that justice as fairness is endorsed by consensus in a stable manner by idealized citizens. This entails the application of a political conception to one single real-life case: the constitutional essentials. My interest focuses

on another decision model suitable for real controversies: studying the change of moral consciousness in urgent public controversies requires identifying the plausibility of establishing epistemic responsibilities that help to form agreements about controversies in applied ethics. Establishing epistemic responsibilities in some groups in society obviously does not by itself solve the controversy, but it helps to identify the tipping point on which the pressure of the various resources of the politics of resistance can be exerted. But let's go step by step.¹⁷ Let's first follow Rawls's footsteps. Extrapolations are dangerous, and we must proceed with care.

Now, perhaps Wenar goes too far in attributing a historical determinism of sorts to Rawls' deployment of the presentation.¹⁸ For my purposes here, it is enough to consider the specific role that the burdens of judgment performed in Western history as a kind of *moral fallibilism* of modern institutions (delivered by the Reformation, but which continued to fragment philosophical and religious systems throughout history). In particular, we must consider how this role connects with

¹⁷ The politics of resistance consist of collectively coordinated containment efforts sharing a common agenda vis-à-vis governments that play the role of targeting that agenda, initiators of claims, or implicated third parties. See Tarrow (2008).

¹⁸ I am grateful to a blind reviewer for raising this point to me. But perhaps we can find support for this kind of purported notion of progress in the roles that Rawls assigns to political philosophy at the beginning of *Reformulation*. There, Rawls recovers the Hegelian ideal of moral change through reconciliation. In this sense, political philosophy is in charge of processing ideological and moral conflicts to develop a more enlightened conception of ourselves at each step (Rawls, 2001, § 1). Perhaps Rawls himself was applying this ideal of change to modern political history when he suggested the authoritarianism of the Catholic Church was tempered by the Reformation in the sixteenth century and by religious conflict (Rawls, 1993, pp. xxii-xxviii). Although religious freedom was initially adopted as a *modus vivendi* due to the exhaustion and debacle of religious wars, Rawls seems to suggest European culture eventually embraced pluralism as a kind of *ethos* or at least a broader public moral fallibilist political culture, which is embraced not for convenience but for its own sake (1993, p. 213). Thanks to this, the stability of justice as fairness and its overlapping consensus can be reliably conditional on the fact that citizens share, from their first socialization, the main ideas of justice as fairness. This is because these ideas are extracted from a certain idiosyncrasy typical of some peoples (reasonable peoples). But then you still have to explain that cultural and social mechanism.

real-world moral deliberation by means of *O* and *MP*. Under this softer interpretation, even if the process is contingent, people like John are now bound to sustain the burdens of judgment in a way that becomes bounded by a certain kind of reasons (*O*) that are accessible to him because he should have a normative psychological disposition (*MP*).

Now, there are at least two major problems with this kind of application of *R*. The first is knowing whether *R* in this context remains Rawlsian in a significant sense beyond its pedigree or its origin in the political conception of justice.¹⁹ Particularly, this use of the Rawlsian conception seems pragmatist. It could be questioned if the Rawlsian pedigree is lost when applied to cases of social cooperation, where it seems that the correct reasons disappear, and moral principles pursue the goal of solving recurring and structural problems in our cooperative relationships.²⁰ Secondly, one must consider whether what I have suggested in this paper is, instead, already present in Rawls' original account. Rawls definitely establishes something like *R* as a paramount requirement, so I am not really providing any insight at all. I will examine these objections in the next and last sections.

5. Preserving the Rawlsian pedigree of reasonableness in its application

Now I will tackle the following objection: the idea that *R* is not Rawlsian enough. Perhaps the application of *R* suggested above is a rather pragmatic arrangement, as in a *modus vivendi*: people reconsidered their intuitions on forced labour when doing so became attractive or they found incentives to do so.

I believe this objection is mistaken. The reason is that the pragmatism implied in connecting principles like *R* with their function, in this case, is not taken in an essentialist way. It is not assumed that this application of *R* is the only method to approach problems of applied ethics, nor that pragmatism is the only way to understand morality. It is rather a layer of analysis that we associate with other layers, such as ideal Rawlsianism above and social sciences below, where the stability of a conception of justice must occur for the right reasons and not pragmatically as a

¹⁹ I thank Jesús Rodríguez Zepeda, Moisés Vaca Paniagua, and a blind reviewer for pressing me on these last points.

²⁰ See Dewey & Tufts (1981).

modus vivendi. Furthermore, plausibility assessments using *R* are not intended to be one-correct-answer methods. Rather, they aspire to be one method—among others—that allows us to compare different interpretations of practical action or political and moral agency.

One way to show that my account is still, to some extent, Rawlsian is to compare it with Rawls' own attempt at presenting the plausibility of overlapping consensus in a partially comprehensive way. It appears that the Rawlsian model of the decision for moral change (in order to demonstrate the stability of overlapping consensus) begins with (A) a normatively oriented change-trigger, such as *MP*. Thanks to that motivation to seek to better embody the ideals of citizenship, it is possible to articulate (B) a philosophical interpretation of social reality explained through the history of necessary ideas, in this case, the normative vision that one wishes to defend (in the case of Rawls, the stability of the overlapping consensus). This intellectual reconstruction includes a reflection on this interpretation of reality and its history, which explains how the social order that is morally satisfactory at one moment (religious absolutism) ceases to be so at another (reform). This is possible thanks to the fact that in (A) the trigger of that change is established. Finally, it establishes (C) the final objective of reconciliation between different ideologies, cultural visions, and comprehensive doctrines.

Now, just as Rawls perhaps found it necessary to show by means of (A)-(C) some vague conception of the meaning of the concrete practice of *R*-based overlapping consensus, we can show through (A)-(C) what is required in the concrete practice of other principles using *R* as a foundation. The idea is to introduce these kinds of arguments into plausibility assessments of what we should reasonably expect from citizens in urgent public controversies of applied ethics. In order to see how, let's recall John's example is based upon the case of John Newton. Perhaps in this way we can see whether *R* is heuristic or explanatory with respect to the actual history of the person the example is based upon. That might also show how we could insert *R* into our current controversies by using the Rawlsian path contained in *R* between (A)-(C), which is the generalization of the Rawlsian application I presented above.²¹

²¹ It is obvious that people in John's day did not deliberate about what Rawls considered reasonable in democratic and liberal social settings. But thinking about the abolitionist debate in terms of *R* can help us assess the

Captain John Newton was interested in the slave trade as a way of subsistence and also as a life experience. But Captain Newton also had a strong interest in improving his moral conscience. Through (A) we can see that this interest was not satisfied with the adventure but included developing his intellectual and moral capacities to be able to live according to his moral religious principles. This interest ultimately led Newton to abandon the life of the sea for a life of religious contemplation. For us here and now, from our ideal evaluations, wealth-accumulation and adventure at the cost of slavery is simply unacceptable. But the important thing is that the plausibility assessments allow us to use the Rawlsian arguments (B) to recover the Newtonian context. Through the interpretations (B) we elaborate an intellectual reconstruction of the social practice of slavery in Newton's times. Of course, morality is also a social practice (not just ideal speculation) and it is also included in the reflection we create when we mentally reconstruct (from subjective narratives and history) the way in which the practice of slavery was being reflected upon in Newton's time. And for this reason, it becomes questionable for him and his contemporaries.

Anderson (2014, pp. 4-5) explains this can happen in the following way. First, interpersonal conflicts give rise to uncertainty in the application of principles. Second, conditions may have changed to such an extent that the typical application of the principle gives rise to unexpected and unsatisfactory consequences, or people become aware of adverse consequences that were not apparent. Finally, people question the legitimacy of a principle or a social norm, highlighting objectionable features in its application or flaws in its justification. The abolitionist movement arises from and simultaneously provokes reflection on slavery in all these ways (Anderson, 2014, pp. 11-13), but for us the important thing is to note how *O* could participate in these weightings.

It is perfectly possible that people act with respect to the norms of the practice of slavery (rules of property, transfer, and treatment of enslaved people) out of habit or custom without having the faintest idea that these practices are, all things considered, morally impermissible. But once the revolts among enslaved people took place, the protests of the abolitionist movement, the political and academic exhortations, etc., then the lack of certainty and the disagreement detonate the applied moral deliberation

plausibility of holding them accountable for weighing abolitionists' evidence and changing their beliefs.

and *O* could come into play, because it establishes the conditions of the fallibility of the moral judgments through the burdens of the judgment. People motivated by *MP* seek out more complete and full ways to realize their abilities and the ideal of citizenship. This implies that people, by questioning their compliance with the rules of slavery, take control over their future behaviour, considering different alternatives to the conflict.

In order for these different alternatives to be considered, it is necessary to preserve the pluralistic *ethos*, or at least the softer form of fallibilism that seems to be implied in the burdens of judgment. Finally, the motivation and incentives provided by *MP*, on the one hand; and the interpretations of history and the current moment disciplined by *O*, can give rise to an attempt to reconcile the different demands and claims with the real world. This would correct the biases that had dominated the public discourse in defence of slavery, and against alternative social worlds.

In that sense, Newton established a dialogue with the abolitionists more or less in the terms established by *O*, in such a way that he considers his own moral judgments fallible. As Anderson notices, in Newton's famous 1781 sermon, in which he condemned England for an enormous list of deadly sins, he did not mention slavery. But the consideration of the possibility that there may be standards of objectivity to which their moral judgments may be subject, including their religious interpretations, opened the door for abolitionists thirty years later, after he had left the slave trade, to persuade Newton to write his eyewitness account of the slave trade. Finally, Newton wrote in that account that he felt humiliated by having been involved in such cruel and terrible practices that contradict the sentiments of humanity. That is, the conditions of *O* (together with the burdens of judgment) could help establish duties that bind Newton to change his beliefs, and the pluralistic *ethos* or at least the imperative of fallibility, allows this process to expand and continue in a stable manner.

The hope is that we can insert the Rawlsian pathway contained in *R* between (A)-(C) to ponder our own intuitions in applied controversies of our time, such as the recognition of trans children, adoption by same-sex couples, the ethics of borders, etc. I refrain from exemplifying the application of the Rawlsian *R* path because my aim was not to take a position regarding these and other controversies, but to show how the partially comprehensive path established by Rawls himself can be used

at least in some of these controversies to establish a standard of citizen epistemic conduct.

6. Is *R* already implicit in Rawls' account?

Another objection arises: have I provided any insight at all regarding how to judge the plausibility among distinct options for moral change in a given context by invoking *R*? This challenge is important because I have been extracting the elements of *R* from Rawls' own account.²²

In order to rebut this objection, we must remember that, famously, the problem with using *R* as a standard of citizen conduct in controversies such as the ones I described in the introduction and the first section is what Faviola Rivera (2007) identifies as a dilemma established by two contrary demands pulling in opposite directions. The first, which she calls the *acceptability condition*, requires that a conception of justice must be acceptable from citizens' point of view whatever their own moral ideas (see Rawls, 1993, p. 137). The second she calls *moral-correctness condition*, and it requires that principles of justice be morally justifiable. These conditions are dilemmatic because, as someone strives to fulfil the acceptability condition, it seems reasonable to avoid relying in moral argument and reasoning, as many citizens will uphold their support for moral outlooks incompatible with their own. Yet, as one strives to fulfil the moral-correctness condition, moral argument and reasoning is required in order to justify a conception of justice.

Conversely, our explanation of *R* must be compatible with comprehensive moral doctrines and, at the same time, it must have normative moral force provided by moral argument and reasoning. For instance, to use my opening example: it is not clear how John, who is a fervent Christian, could accept *R* if *R* conflicts with his own faith (provided that his Christian ardour does not allow him to be reasonable as characterized above, because that would entail considering that his beliefs may be fallible).²³ But if John found himself contrived by the controversy between slaveholders and reformers and turned to *R* as a

²² I am in debt to a blind reviewer for raising this objection.

²³ Of course, Rawls' account is completely out of phase with John's historic setting, but for the sake of the argument, we could notice that John's Catholic faith may prevent him for accepting the burdens of judgment (iii), as his doctrine declares pontifical infallibility. I am very thankful to a blind reviewer for suggesting this remark.

guide to determine which side to lean on, his own faith as a Christian would cause him to reject *R* whenever it conflicts with his faith in this way.²⁴ For even if we display moral arguments against slavery, when the doctrine sanctions slavery, then John has no critical point of view that allows him to examine his beliefs. To see this, imagine John comes into contact with the traditional Christian moral discourse that morally justified slavery, arguing that slaveholders freed Africans from savagery by bringing them closer to Christianity (Anderson, 2014, p. 20). If John has to reject *R* because of his faith, then the controversy for him will merely be whether or not the Bible considered slavery permissible in the end. But then we're back in the same place we started.²⁵

I believe this objection misrepresents the different levels of discussion contained above. Indeed, my claim does not add anything worth noticing in *the presentation* of a conception of political justice. Inside the model, as Rawls noticed himself with his warning, the discussions of applied ethics do not even arise, as they are cancelled-out by the restrictions of the model. But my claim becomes interesting within the presentation, where Rawls himself allows his arguments to engage with the partially comprehensive.

²⁴ The Pope can establish the Catholic Church's official stand, and as a rule it is considered that the Pope cannot be mistaken. Consider art. 24 of *The Dogmatic Constitution on the Church* (Lumen Gentium) (cited in Wenar, 1995, p. 44). Of course, John does not live in the kind of democratic setting where the Rawlsian account of justice as fairness is meant to fit. However, my concern is with the virtue of reasonableness and how it can be used to assess our judgments about justice. So even if John's conception of justice differs greatly from a conception of justice for a democratic society, the point is to reflect about how John is able to confront whatever conception of justice he has with the fact of slavery.

²⁵ It is possible to imagine reasonable versions of Catholicism and many other religious faiths. But in order to highlight the requirements of reasonable pluralism, in this paper I consider only the kind of religious practices that presumably John had, which include unreasonable central ideas, such as the dogma of eternal condemnation of infidels that assumes one's faith as the only one correct. But certainly, outside dogmatic versions of familiar faiths, there are many individuals who consider themselves Catholics, Jews, Muslims, or members of some other centrally organized faith but who also share the Rawlsian ideal of social cooperation and collective respect. See Quong (2011, p. 141).

To be sure, consider the controversy around tolerating the unreasonable (Friedman, 2000; Kelly & McPherson, 2001; Quong, 2011). Unreasonable persons reject one or more fundamental elements of a political conception of justice. This poses a problem of public justification, provided that we need to know how a conception of justice may legitimately extract obedience from citizens. But note that this problem lies outside *the content*, as people inside it are simply assumed to be reasonable. So, it seems that at least from the presentation, this problem needs to be tackled, provided that tolerance is taken as a core liberal principle. At least two reasons can be invoked. First and very straightforwardly, Rawls himself confessed that the possibility that justice as fairness could run with stability in real societies made him engage in extensive revisions of *A Theory of Justice* (1971, pp. xv-xxi). Admittedly, this also played a role in the subsequent restatements of the conception (Wenar, 1995, pp. 32-33).

Second, it is unclear how unreasonable people are included in society at all. McPherson (2001, p. 23), for instance, thinks that an implication of *the presentation* is that unreasonable people are to a point excluded from citizenship. Quong (2011, pp. 145 & 292-293), however, believes that a conception of justice that disenfranchises unreasonable individuals cannot count as liberal. For him, we can separate the “constituency for justification” from the bulk of society where everyone is subject to toleration within certain parameters. The justification of a political conception of justice is only modestly addressed to model reasonable persons whereas, outside ideal theory, stability must be worked out in itself (Quong, 2011, pp. 139-140). Using our terms, the content would address the first ideal constituency of reasonable people according to the model, while the presentation may be addressed to real people from whom we wish to demand to act reasonably in real-life pressing matters of public turmoil.

I cannot take into account the intricacies of this debate, but the point is clear: while the requirements posed over ideal citizens by the standard of reasonableness may be fully deployed within Rawls’ machinery, what the standard means in the kind of debates I depicted above is highly confusing and not at all clear, hence the need people like Friedman, Kelly, McPherson, Quong and others felt to clarify what it means to be or not to be reasonable in the real world if we take stability seriously as a constraining value.

7. Conclusion

I started off —following Wenar (1995)—with the assumption that the virtue of reasonableness unfolds in two dimensions. First, within the content of a political conception of justice, it allows Rawls to explain how a particular conception may be acceptable for citizens in modelled conditions of the pluralism of comprehensive theories. Second, the presentation provides a partially comprehensive support for the plausibility and stability of justice as fairness. The objective within the presentation is not only the application of what is reasonable in constitutional principles; rather, Rawls goes further in trying to show that this application of constitutional principles is stable over many generations.

Of course, Rawls' attempt to defend the overlapping consensus in an applied way was, for many critics, unpalatable from the point of view of other comprehensive doctrines or theories. But since my purpose here was not within the limits of ideal theory, the partially comprehensive nature of *R* was an advantage. As a result, I have shown how to use *R* as a process to discipline our plausibility judgments. Said process means determining what we can reasonably expect from a citizen in a given context. In order to see how *R* may be used, I used John Newton's example.

We moved forward and backwards through time and in each case invoked *R* to compare our judgments about the requirements of reasonableness. In this way I have shown how to apply the second role of *R* (the presentation) to determine the conditions under which a citizen has a reasonable epistemic conduct in a real public controversy. This allowed me to explain, in turn, the conditions under which we may have reasons to assign responsible negligence.

After this more or less Rawlsian path, it could be objected that the conditions under which we can attribute responsible negligence are already contained within the original stipulations made by Rawls himself for *R*. However, this objection was dismissed when the need to supplement the presentation became evident when we tried to apply *R* to the case of John.

To be sure, consider the specific requirements of *R* regarding the burdens of judgment. Believers interpret their conception of the goodness of life from their faith as required by (vii). But at the same

time, believers must remain objective with respect to conflicts that may arise between their faith and empirical evidence (v), considering that their religious beliefs may be particularly fallible (iii). Furthermore, if at any point the empirical and social evidence outweighs their religious beliefs, they must be ready to bracket their faith and act accordingly with what the objective point of view dictates (vi).

It seems like this critical path is not available for a believer such as John, since Catholics in general find difficulty accepting *R*, and being Catholic itself would mean sometimes violating *R* as a standard of conduct for citizens. At the same time, this sits awkwardly with Rawls' insistence that the main historical religions constitute reasonable comprehensive moral doctrines (1993, p. 170). This seems to highlight that *R*, or a part of *R*, is itself a theory or doctrine and at least partially comprehensive, since it claims to be the correct answer to the way in which we should form political beliefs.

If the objection above is relevant to cases like John's, then my attempt to use *R* for applied ethical controversies (where adjudication of epistemic responsibility is needed) may be disappointing: *R* does not seem like a plausible choice as a standard of citizen conduct in real debates. Which is to say that we cannot in effect successfully violate Rawls's warning not to use his approach for applied cases. But I think ending on that disappointing note may still be premature. First, this shows that my account is not really implicit in Rawls' account—that's why it is problematic to apply *R* to real-life scenarios to guide our intuitions.

Second, while Rawls thought political philosophy should not be used in any way in problems of applied ethics (2001, p. 14); he also claimed—in his celebrated discussion of the position of Dr. Martin Luther King, Jr.—that, in the non-ideal conditions of unordered and unjust societies, comprehensive theories are required to improve social and political conditions (1993, pp. xlix-l). It is precisely through comprehensive or partially comprehensive theories or doctrines that we can defend the main values of a liberal and democratic culture, such as social cooperation, legitimate order, egalitarian institutions and democratic participation—among others—with the objective to order society or have a legitimate society.²⁶

²⁶ On this role of partially comprehensive theories or doctrines, see Macedo (1995) and Clayton & Stevens (2014).

Again, this shows the account is not implicit, but it also highlights its Rawlsian pedigree. To be sure, consider the distinction between overlapping consensus and urgent controversies in disordered institutional conditions. Both are cases of decision-making; but the first requires an ideal consensus in which what is reasonable has to be acceptable to citizens despite its comprehensive doctrines, and the second requires precisely what is reasonable as a partially comprehensive theory capable of ordering society. Thus, my purpose was to transfer these arguments about the conception of the reasonable from Rawls's original approach to a conception of epistemic responsibility that helps us understand the plausibility of moral change over a public controversy in a given context through *R*.

So, to repeat the point again, if comprehensive (or at least partially comprehensive) doctrines or theories are what we need in cases of disordered societies, then perhaps the reason why *R* seems unacceptable from the point of view of comprehensive doctrines is that *R* is itself partially comprehensive. This may pose a problem of coherence within the machinery of Rawls' political conception of justice. But if we extract *R* from that context and seek to apply it to public controversies in contexts of disordered societies, then perhaps it may be an advantage that *R* is partially comprehensive. This shows both Rawlsian pedigree and that my account is not implicit, since it needs to be extracted from Rawls' account and reformulated in a way that could guide our intuitions.

This is why my aim was to study ways of judging the plausibility of reasonable moral change in a given context characterized by urgent controversies, by using the virtue of reason as a standard of epistemic conduct. I think, at least in those particular cases, invoking *R* potentially allows us to improve the *status quo* by pushing it towards some plausible version of ideal theory. For example, determining what we should reasonably expect in the next elections from the US voters of Latino origin who overwhelmingly supported Donald Trump, or what we should reasonably expect from privileged white women in matters of affirmative action; etc.

It is true that many Latino immigrant voters will vote for anti-immigrant policies and many privileged white women will discriminate against less empowered groups; but what matters here is to see the plausibility of identifying the responsibilities that these groups have to modify their beliefs and behaviour with the aim of later looking for ways to demand that responsibility from them. I believe I have shown that

this may be possible by using *R* as a standard of epistemic behaviour for citizens.

Bibliography

- Anderson, E. (2014, February 11). *Social Movements, Experiments in Living, and Moral Progress: Case Studies from Britain's Abolition of Slavery. The Lindley Lecture at the University of Kansas*. URL: <http://hdl.handle.net/1808/14787>.
- Bureau of International Labor Affairs. (2022, September 28). *List of Goods Produced by Child Labor or Forced Labor*. URL: <https://www.dol.gov/agencies/ilab/reports/child-labor/list-of-goods-print>.
- Caney, S. (2012). Just Emissions. *Philosophy & Public Affairs*, 40(4), 255-300. DOI: <https://doi.org/10.1111/papa.12005>.
- Clayton, M. & Stevens, D. (2014). When God Commands Disobedience: Political Liberalism and Unreasonable Religions. *Res Publica*, 20(1), 65-84. DOI: <https://doi.org/10.1007/s11158-013-9221-y>.
- Cohen, G. A. (2008). *Rescuing Justice and Equality*. Harvard University Press.
- Dewey, J. & Tufts, J. H. (1981). *Ethics*. Southern Illinois University Press.
- Friedman, M. (2000). John Rawls and the Political Coercion of Unreasonable People. In V. Davion & C. Wolf (eds.), *The Idea of a Political Liberalism: Essays on Rawls*. (pp. 16-33). Rowman & Littlefield Publishers.
- Fueller, L. L. (2012). Burdened Societies and Transitional Justice. *Ethical Theory and Moral Practice*, 15, 369-386. DOI: <https://doi.org/10.1007/s10677-011-9300-0>.
- Kelly, E. & McPherson, L. (2001). On Tolerating the Unreasonable. *Journal of Political Philosophy*, 9(1), 38-55. DOI: <https://doi.org/10.1111/1467-9760.00117>.
- Macedo, S. (1995). Liberal Civic Education and Religious Fundamentalism: The Case of God v. John Rawls? *Ethics*, 105(3), 468-496. DOI: <https://doi.org/10.1086/293723>.
- Mandle, J. & Reidy, A. D. (2003). *A Companion to Rawls*. Wiley-Blackwell.
- Mill, J. S. (2018). *The Subjection of Women*. Routledge.
- Miller, D. (2011). Taking up the Slack? Responsibility and Justice in Situations of Partial Compliance. In Z. Stemplowska & C. Knight (eds.), *Responsibility and Distributive Justice*. (pp. 230-245). Oxford University Press.

- Newton, J. (1788). *Thoughts upon the African Slave Trade*. J. Buckland and J. Johnson.
- (1830). *Narrative of Some Remarkable and Interesting Particulars in the Life of the Rev. John Newton*. Glasgow.
- Puddifoot, K. (2017). Epistemic Discrimination. In K. Lippert-Rasmussen (ed.), *The Routledge Handbook of the Ethics of Discrimination*. (pp. 54-67). Routledge.
- Quong, J. (2011). *Liberalism without Perfection*. Oxford University Press.
- Rawls, J. (1971). *A Theory of Justice*. Harvard University Press.
- (1993). *Political Liberalism*. Columbia University Press.
- (2001). *Justice as Fairness: A Restatement*. Harvard University Press.
- (2002). *Law of Peoples*. Harvard University Press.
- Rodríguez Zepeda, J. (2006). *Un marco Teórico para la Discriminación*. CONAPRED.
- Smith, A. (1904). *An Inquiry into the Nature and Causes of the Wealth of Nations*. Library of Economics and Liberty. URL: <http://www.econlib.org/LIBRARY/Smith/smWN.html>.
- Stemplowska, Z. (2016). Non-ideal Theory. In K. Lippert-Rasmussen, K. Brownlee & D. Coady (eds.), *A Companion to Applied Philosophy*. (pp. 284-296). Wiley.
- Tarrow, S. (2008). Charles Tilly and the Practice of Contentious Politics. *Social Movement Studies*, 7(3), 225-246. DOI: <https://doi.org/10.1080/14742830802485601>.
- Vaca, M. (2022). ¿Qué nos dice la teoría ideal sobre la discriminación? In E. Camacho-Beltrán and Luis Muñoz Oliveira (eds.), *Trato de sombras: estudios sobre discriminación incorrecta*. (pp. 123-154). UNAM-Centro de Investigaciones sobre América Latina y el Caribe.
- Valentini, L. (2012). Ideal vs. Non-ideal Theory: A Conceptual Map. *Philosophy Compass*, 7(9), 654-664. DOI: <https://doi.org/10.1111/j.1747-9991.2012.00500.x>.
- Wall, S. (2013). Perfectionist Justice and Rawlsian Legitimacy. In J. Mandle & A. D. Reidy, *A Companion to Rawls*. (pp. 413-429). Wiley-Blackwell.
- Walzer, M. (1993) *Spheres of Justice: A Defense of Pluralism and Equality*. Basic Books.
- Wenar, L. (1995). Political Liberalism: An Internal Critique. *Ethics*, 1(106), 32-62. DOI: <https://doi.org/10.1086/293777>.
- Ypi, L. (2012). *Global Justice and Avant-garde Political Agency*. Oxford University Press.