

Ricœur's Rawls: Constitutive Antecedence and Reflective Equilibrium

Benjamin Coy Hutchens

Rutgers University, Newark

Abstract

This article aims to stimulate dispute about the justification of Paul Ricœur's hermeneutic reading of John Rawls. Offering a close, methodically point-for-point textual engagement, I shall propose that Ricœur's misreading of certain hermeneutic circularities in Rawls is owed to some confusion about the role of the procedural nature of Rawls' theory. Generally speaking, the problems with Ricœur's interpretations center on the question of whether there is something "pre-understood" within the formal theoretical understanding of the procedural theory of justice and the substantive convictions and judgments that figure within the reflective equilibrium of deliberations about the terms of justice. Arguably, Ricœur has not made a satisfactory case that the difference and liberty principles are considered convictions that anticipate their discovery and establishment. Ultimately, Ricœur has not demonstrated that there is a single presuppositional form that renders Rawls' procedure self-defeating. Instead, he has proposed to us several potential forms of damaging presupposition, each of which is based on a questionable reading of Rawls' text.

Keywords: Hermeneutics; Justice; Reflective Equilibrium; Convictions; Maximin Rule.

Résumé

Le but de cet article est de susciter la discussion autour de la justification de la lecture herméneutique que Ricœur fait de Rawls. En proposant une approche textuelle serrée, méthodique et procédant point par point, je soutiens que la mécompréhension ricœurienne de certaines circularités herméneutiques chez Rawls est due à une confusion concernant le rôle de la nature procédurale de la théorie de Rawls. Généralement, les problèmes soulevés par les interprétations de Ricœur sont focalisés sur la question de savoir s'il y a quelque chose de "pré-compris" dans la compréhension théorique formelle de la théorie procédurale de la justice et dans les convictions bien pesées et les jugements qui interviennent dans l'équilibre réfléchi des délibérations sur les questions de justice. Sans doute Ricœur n'a-t-il pas suffisamment pris en compte le fait que le principe de différence et le principe de liberté sont considérés comme des convictions qui anticipent leur découverte et leur établissement. Au final, Ricœur n'a pas démontré qu'il existe une seule forme de présupposition qui mette en échec la procédure rawlsienne. Au lieu de cela, il nous a proposé plusieurs formes potentielles de présuppositions préjudiciables, lesquelles reposent sur une lecture discutable du texte de Rawls.

Mots-clés: Herméneutique; Justice; Équilibre Réfléchi; Conviction; Règle du Maximin.

Études Ricœuriennes / Ricœur Studies, Vol 11, No 1 (2020), pp. 130-143

ISSN 2156-7808 (online) DOI 10.5195/errs.2020.388

<http://ricoeur.pitt.edu>



This work is licensed under a Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 United States License.



This journal is published by the [University Library System](#) of the [University of Pittsburgh](#) as part of its [D-Scribe Digital Publishing Program](#), and is cosponsored by the [University of Pittsburgh Press](#).

Ricœur's Rawls: Constitutive Antecedence and Reflective Equilibrium

Benjamin Coy Hutchens

Rutgers University, Newark

Rawls' *A Theory of Justice* has always been ripe for hermeneutic picking. Developing a procedure that draws attention to its every theoretical move, its textual composition nonetheless leaves sufficient space to wonder whether *all* of its presuppositions have been laid bare, especially where the basic structure of political society is at issue. We might wonder whether the procedural theory of justice it offers actually discovers and establishes basic principles of justice without already presupposing precisely such principles in order to do so. Hermeneutics is interested not only in whether something is presupposed, but also in whether that presupposition renders the procedure self-defeating.

Such issues are nothing new for political philosophy. For example, in his immensely influential work *Being Singular Plural*, Jean-Luc Nancy avers that Rousseau exposed the "aporia of a community that would have to precede itself in order to constitute itself."¹ Insisting that Rousseau did not resolve this difficulty, Nancy asks: what of the "originary division [*déliation*]" between the individuals who enter into a social contract with one another? In other words, what reveals itself in this political context is a *hermeneutic* problem of *constitutive antecedence*: something entailed from or consequent upon a truth would also have to precede and constitute that truth in order for it to be a consequent.²

If Rousseau imagined that his macro-view of history and his micro-conception of the role of the legislator addressed this problem satisfactorily in *The Social Contract* (Book II, section 7), the social contract tradition from Kant to Rawls has not agreed. As is clear from his lectures on the history of political philosophy, Rawls regarded himself as improving upon Rousseau's solution to the problem (though of course, this was certainly not the primary intention behind his procedural theory of justice). And in his articles on Rawls' work in *The Just* and in *Oneself as Another*, Paul Ricœur was intrigued by the *textuality* of the procedural theory in *A Theory of Justice*. For Ricœur, Rawls' project is inherently self-defeating as long as it fails to overcome the problem of constitutive antecedence. And as long as material "contingencies" have not been stripped away by the formal procedure, this problem can always potentially arise: criteria, however formal, can disguise prejudices, "ideological" or otherwise.

From the outset, a certain comment about Ricœur's work on Rawls is due. The two papers on Rawls in *The Just* are similar in most important particulars and cover material also repeated nearly word-for-word in the eighth study of *Oneself as Another*. This article refers to the first of these papers, entitled "Is a Procedural Theory of Justice Possible? John Rawls' Theory of Justice,"³ though reference to other sources will be made when helpful.

Affirming that the bond between a deontological ethics and a social contract politics is not a “contingent” one, Ricœur nonetheless queries whether a contract theory such as Rawls’ can “substitute a procedural approach for every attempt to ground justice on some prior convictions concerning the good of all, the common good of the *politeia*, the good of the republic or the *Commonwealth*?”⁴ As he sees it, there is an unfortunate circularity in the way that Rawls handles this question, especially inasmuch as he aims to surmount the difficulty of the passage from the Kantian principle of autonomy in the *Groundwork* to the idea of the social contract in *The Metaphysics of Morals* (§§46-47).⁵ The problem arises for Ricœur when one discovers that a certain “moral sense of justice founded on the Golden Rule” is “always presupposed by a purely procedural justification of the principle of justice” itself. If so, the underlying dynamic of Rawls’ constructive approach to justice is borrowed from “the very principle it claims to engender by its purely contractual procedure.”⁶ It is precisely this circularity that is at stake in Ricœur’s approach to Rawls’ procedure, and my task here is to query whether there is indeed such circularity and whether anything like this notion of the Golden Rule is “presupposed” by Rawls’ procedure.

Ricœur covers similar terrain in *Oneself as Another*. The question of whether a non-historical contract can be binding on a historical society confirms in his mind that there is a significant difference between the social contract and the principle of autonomy. The problem is that the social contract involves a group of people imposing laws upon themselves, whereas autonomy involves a self-legislating personal freedom. A question of *principle* arises here, according to Ricœur, as to whether even a deontological theory of justice “calls upon” a certain “ethical sense of justice” especially when presented procedurally. He asks: “does a purely procedural conception of justice succeed in breaking all ties to a sense of justice that precedes it and accompanies it all along?”⁷ In other words, this problem of circularity presupposes two theoretical axes: on one axis there is the alleged problem of the relation between the social contract and autonomy; on the second axis, the alleged problem of the presupposition of an ethical foundation (or principle) by any procedural theory intending to arrive at such a foundation (or principle).

To address this, Ricœur takes note of a basic distinction between fairness and justice: the former “characterizes the procedure of deliberation that should lead to the choice of those principles of justice recommended by Rawls, whereas justice designates the content of the chosen principles.”⁸ If the situation in which rational individuals deliberate is fair, then the institutions they choose ought to be just. If Rawls’ work is deontological in nature, it is nonetheless without transcendental foundation. And being without such a foundation entails that there is no commitment to any objective criteria of justice. This in turn opens the matter of circularity concerning presuppositions: something must be presupposed about the conditions of fairness in order for those presuppositions, now in the form of agreed terms, to reemerge. And so, although Ricœur does not mention it here, justice comes in the form of the *material content* of the principles agreed upon, content which is stripped away by the veil of ignorance so that the *form* of the fair condition of deliberation can be secured, ultimately for the purpose of reintroducing the material content presupposed later. And with this procedure of formalizing the basic situation, we find ourselves at the center of the constructivist project of all contractualism. When “the just” is constructed by means of a procedure, it could not have been known in advance of the procedure. Yet, something must be presupposed even in the absence of this knowledge, and precisely what that presupposition is poses a difficulty for any deontology without transcendental basis – or more specifically, any procedural conception of justice that is substituted for an ethical foundation.

Ricœur is interested in the further question of the nature of the veil of ignorance. Does it accomplish what Rawls thinks it does? For Rawls, the veil strips individuals of the knowledge of their own singularity as well as the knowledge of the singularity of their relations with others: they are “artificial persons” inhabiting the space of a thought experiment, not metaphysical persons resulting from a reduction to some essence.⁹ One thing is certain for Ricœur: the wills exercised in deliberations under the veil are not Kantian “transcendental wills,” or as he understands them, wills stripped of any empirical ground and consequently lacking any reference to ends or values. Here I think Ricœur is leveling at Rawls the same criticism that Hegel raised against the Kantian critical philosophy: it is merely “formal,” though in this case, its formalism is imperfect. Kant’s distinction in the *Groundwork* between formal practical principles abstracted from subjective ends and the material practical principles which have such ends as their basis enables him to bracket the material aspect of wills in *The Metaphysics of Morals*, where he declares that the concept of will has to do only with the “external” and “practical” relations between willful persons insofar as they can influence each other. The material content of their willful choices is husked away, leaving only the “form” of choices, that is, their freedom and their conformity with the equally “formal” universal principle of right.¹⁰ Since the Rawlsian subject does have some “earthly interests” (though of course it will not (re)-discover what they are until the veil is lifted), Ricœur asserts that this subject is situated between transcendentalism and empiricism: an empirico-transcendental subject, one might say, one divested of some kinds of material content but not all. Rawls has made this point, of course, because there are some things that such subjects must know if there is to be any agreement at all about post-veil “real world” matters. In particular, rational subjects who enter into the original position may not know what their own “empirical” interests are (the material contents of formal wills), but at least they know what it means to “have an interest.” For Ricœur this means that a problem of justice pertaining to social advantages already arises. As is well-known, four forms of knowledge must be possessed by each rational subject if there is to be agreement at all, namely, knowledge of (a) the general psychology of human passions and emotions, (b) the primary social goods, including self-respect, without which freedom would be empty, (c) the competing principles of justice, and (d) certain constraints that bear on all ethical decisions, including that of publicness and stability.¹¹ For Rawls, the veil of ignorance is thus optimally “thick”: it is thin enough for the rationally deliberative agents to be recognizable as such, but thick enough for them not to have the kinds of self-knowledge that can admit contingencies into the decision making process.¹² Since none of this knowledge presupposes or establishes an “ethical foundation” for justice, Rawls is satisfied that any circularity that might arise is merely a procedural matter that does not in any way compromise the unprejudiced nature of the endeavor. It does not follow from the unwarranted presence of material contingencies in the procedure. But of course, this is not to say that Rawls admits that any circularity does arise.

Ricœur is not convinced that this is entirely correct. He argues that such a procedural conception of justice “at best” provides rationalizations of some specific sense of justice that is already presupposed by them. In defense of this claim he notes that, although Rawls presents his principles of justice – the liberty and difference principles – in a “serial order,” there is in fact a “circular order” that is “characteristic of all ethical reflection” (though this is not explained here). He observes:

The reader may be surprised by the fact that the principles of justice are defined and even developed (§§11-12) before the examination of the circumstances in which the choice is made (§§20-25), consequently before the thematic treatment of the veil of ignorance (§24) and, in an even more significant way, before the demonstration that these principles are the only rational ones (§§26-30). This does not prevent Rawls from characterizing the two principles of justice in advance as the ones that would be chosen in the original situation.¹³

Ricœur goes on to quote Rawls to the effect that principles of justice are in fact the principles that rational, self-interested subjects in the original position would accept as “defining the fundamental terms of their association.”¹⁴

However, closer examination of the matter shows that Rawls is speaking here rather offhandedly. He is looking ahead to what the original position will ultimately enable rational agents to understand about their association. In fact, there is no reason for the reader to be surprised by this early presentation of principles of justice. The development of concepts and theories in *A Theory of Justice* is in fact “serial,” and Rawls simply provides a sneak preview of the principles before they are operative. Comments of this kind are not to be found deeper into the exposition of the procedure. But what really counts here is that his casual observation that they “would” accept such principles to define their association does not mean that they have already done so procedurally.¹⁵ And although the reader may “anticipate” the presentation of such principles, it does not even follow from the fact that we can do so that the rational subjects under the veil would do so. Anticipating such principles is quite different from preconceiving various basic characteristics of the original position, especially those pertaining to self-interests and other-interestedness.

Concerning the question of circularity, Ricœur arguably misses the target in §11 of *A Theory of Justice*, where Rawls simply says that he will present the liberty and difference principles in a “tentative” way so that the procedure can flow naturally. However, what is most important is that Rawls concluded §10 with the observation that formal and substantive forms of justice have too often been conflated, and that his own task is to determine “what are the most reasonable principles of substantive justice and under what condition men come to affirm and to live by them,” and only once that is done can the manner in which formal and substantive forms of justice are tied together be discovered.¹⁶ Missing the point that the preliminary introduction of the liberty and difference principles is intended to lay out the conditions for doing so once the fair conditions are first presented and tested reflectively, Ricœur simply concludes from §11 that Rawls’ procedure is a “linear argument, but one that provides a progressive clarification of the preunderstanding of what justice means.”¹⁷ One might notice here the introduction of the hermeneutic notion of “pre-understanding,” the understanding that must obtain if further understanding is to take place. This is important for his principal argument against the linearity of Rawls’ position, and of course for his view that Rawls is answering Rousseau’s problem of constitutive antecedence. This principal argument, which Ricœur believes challenges “the whole contractualist school,” is that

[...] the procedural definition of justice does not constitute an independent theory, but rests on the preunderstanding that allows us to define and interpret the two principles of justice that we ought to be able to prove – if we ever got that far – would be chosen in the original situation, that is, behind the veil of ignorance.¹⁸

In other words, in Ricœur's Rawls as in the entire contractualist school, we must "pre-understand" the two principles in order to set the conditions wherein rational subjects can then later discover and establish them. Ricœur continues by noting that circularity, not linearity, favors the self-sufficiency of the procedural definition of justice that relates cyclically with the notion of the original position. He remarks that, although the original position is famously non-historical,

the constraints that define it are, to be sure, constructed as a thought experiment and create a wholly hypothetical situation with no roots in history and experience. But they are imagined in such a way that they satisfy the idea of fairness that works like the transcendental condition for all of the procedural development. Now what is fairness, if not the equality of partners confronted with the requirements of a rational choice? Do we not have here the sense of an *isotes* according to Isocrates and Aristotle, which in turn implies respect for the other as an equal partner in the procedural process?¹⁹

This leads Ricœur to the "suspicion" that a "moral principle governs the apparently artificial construction" of the original position. This suspicion, he insists, is confirmed by the role of the maximin rule, which provides an independent foundation for the choice of the two principles: there is a problem of social justice and the two principles are the solution to which reasoning by the maximin rule will lead us under conditions of uncertainty. For Ricœur, this argument merely seems to offer an ethical conclusion to non-ethical premises; reasoning by the maximin rule yields only an ethical argument disguised as a technical procedure. Thus, Ricœur argues that the maximin rule is a "tacitly ethical argument," a notion he insists was obvious from the very first pages of *A Theory of Justice*, where Rawls concludes rousing lines about justice being the first virtue of institutions with the following words: "Each person possess an inviolability founded on justice that even the welfare of society as a whole cannot override. For this reason justice denies that the loss of freedom for some is made right by a greater good shared by others."²⁰ Ricœur inquires:

Having read these lines, one can ask how it can be possible to maintain simultaneously the recognition of an ethical presupposition and the attempt to free the procedural definition of justice from every presupposition concerning the good and even the just.²¹

However, I think that Ricœur is rather unfair to Rawls here on two counts. First, this quote from the first page of Rawls is not actually very damning when one considers that Rawls is merely stating an anti-utilitarian position that has been achieved by the technical procedural theory of justice, a theory in which the maximin rule itself is pivotal. Second, just because the maximin rule recommends a certain criterion that *resembles in form but contradicts in content* the standard version of the utilitarian principle does not mean that the maximin rule *functions* as an ethical principle. When Ricœur notices this formal resemblance and informal contradiction, he retroactively assigns to the maximin rule an ethical *significance* it does not possess without specifying precisely what its ethical *content* might be. In fact, the conclusion that not even the welfare of society can override the inviolability of the individual is merely one entailment of the *usage* of the maximin rule. It offers neither utilitarian criteria nor a utilitarian solution.

Ricœur is intrigued by a certain proposal Rawls makes in section four of *A Theory of Justice*: there is an *additional* way that the original position can be "justified" in some particular form,

namely by seeing whether some particular reflective manifestation of it matches or extends our “considered convictions.” For Ricœur, Rawls’ notion of “considered convictions” serves as the mediation between the ethical presupposition of the theory of justice and the purely technical maximin rule. He writes that “it is in situations where a certain moral consensus already reigns that are [*sic*] formed what we could call a preunderstanding of the principles of justice.”²² What Rawls actually says is that one way to justify some particular form of the *original position* is to see whether the principles that would be *chosen* in it would “match our considered convictions of justice or extend them in an acceptable way.”²³ For Ricœur, such considered convictions are nothing other than the “preunderstanding” of the moral judgments that are intuitive in Rawls’ terms. Ricœur asks: “Are not these considered convictions ultimately rooted in the sense of justice equivalent to the Golden Rule applied to institutions and no longer to individuals in a face-to-face relation, and moreover to institutions considered from the point of view of their distributive function?”²⁴ Clearly, this is a rather contingent interpretation. It is one thing for such convictions to *represent a preunderstanding of the moral judgments* about which we have the strongest convictions. It is another thing altogether for them to be *rooted in a sense of justice* “equivalent to the Golden Rule,” whether applied to institutions or otherwise. And neither of these connections made by Ricœur links up neatly with the unique “distributive function” of what is presumably the maximin rule. For purposes of argumentation, I am going to refer to what Rawls writes here as the “heterodox” interpretation of the justification of the original position, since it is not only quite different than the normal justification of the original position with which we are all familiar, but figures rarely in the book and for that reason remains considerably less developed.

I would suggest that Ricœur’s reading of this heterodox interpretation of the original position concedes that the conception of the maximin rule as having ethical content in itself is a failure. Closer reading of the passage reveals that Rawls is not proposing that rational subjects in the original position *would* weigh some particular manifestation of the original position against their “considered convictions” (though of course they *could* do so); on the contrary, it is we who, when examining any such particular manifestation, compare it with our firmest convictions about what a theory of justice should do for us (for example, exclude slavery). When Ricœur tries to defend his position with the observation that the entire lexical order of the liberty and difference principles is “preunderstood” at the level of these considered convictions, we are led to believe then that there are two ways of arriving at them: “rationally” by means of the correct usage of the maximin rule under the veil of ignorance, as in the orthodox interpretation, and “intuitively” because we already understand them at the level of our considered convictions, as in Ricœur’s reading of the heterodox interpretation! Surely that is not at all what Rawls was proposing with this heterodox justification of the original position.

Now, Ricœur may be compounding his mistake by leaping forward along the linear or serial ordering of Rawls procedure to insist that the notion of reflective equilibrium is in fact a mutual adjustment of these considered convictions and “theory.” What Ricœur does not notice in §4 and which is even more apparent in §9 and much later in §87 of *A Theory of Justice* is that there is a semantic shift in emphasis from “considered convictions” to “considered judgments” when reflective equilibrium is presented. When one studies these passages, it becomes somewhat clear that a “considered conviction” is something like a claim whose criteria one will plug into the process of reflective equilibrium to see whether it will match with principles and thus become a

“considered judgment” that can be taken as a fixed point within the process of working towards basic principles. Rawls writes:

By going back and forth, sometimes altering the conditions of the contractual circumstances, at others withdrawing our judgments and conforming them to principles, I assume that *eventually* we shall find a description of the initial situation that both expresses reasonable conditions and yields principles which match our considered judgments *duly pruned and adjusted*.²⁵

Notice here that reflective equilibrium involves a mutual alteration of conditions and judgments, not convictions and “theory” (or “principles”). So, when Ricœur claims that principles and judgments “end up by coinciding,” he is presenting only a simplification of the reflectivity of judgment.²⁶ As Rawls himself points out, it is a legitimate usage of reflective equilibrium to revise judgments which reveal certain irregularities or distortions so that they conform to principles even though in such cases they do not agree with his principles. And one is very likely to do so if there is an explanation for these irregularities and if one finds the resulting conception of fair conditions acceptable.²⁷ So, at this point, one already has some reason to believe that “considered convictions” are neither merely some potential prejudices slipping into this procedure, nor that they are unexamined criteria that serve to determine the domain of deliberation.

Even in the context in which Ricœur cites him, it is obvious that what is “reflective” in a Kantian sense about Rawls’ theory is that we can alter the principles of justice if they do not fit the reasonable conditions of the original position, and vice versa. But neither these principles nor these conditions can be said to be “preunderstood”: each is proposed in conjunction with the other, and each is tested correlative to the other. In particular, reasonable conditions define the form the particular manifestation of the original position takes. It may be a highly desirable outcome if the result matches or supports our considered convictions, but the strength of the entire procedure Rawls offers is predicated on the fact that, desirable as it may be, this agreement is a strictly contingent matter. There is no sense in which our considered convictions qua principles of justice will “end up by coinciding” with judgments made about reasonable conditions of deliberation. Nor is there any sense to the notion that considered convictions qua reasonable conditions will ultimately coincide with the two principles. To be clear, I think that what Rawls means by “considered convictions” in this heterodox interpretation is that certain basic presuppositions of liberalism (such as the wrongness of religious intolerance) will be contingently supported by the result of the procedure *as a whole*, and that in fact religious intolerance will be rejected under the veil of ignorance quite early in the process of working towards final reflective equilibrium (a good exposition of precisely this idea is to be found in the introduction to his lectures on the history of political philosophy).²⁸

If Ricœur is mistaken about the nature of the circularity (either it does not occur or its contents have been misidentified), then there is no need to follow him when he insists that it is threatened by the “centrifugal forces exercised by the contractualist hypothesis.” However, Ricœur’s reason for arriving at this conclusion is interesting in its own right. He asserts that the difference principle alone has been favored by a certain artificial (but not procedural) tendency exercised by both the notion of the original position and the maximin rule. In other words, the difference principle, and presumably not the liberty principle, has been adequately supported by

the special role played by the original position and the function of the maximin rule. If this is the case, then there is no true seriality: if the two principles must be covered in serial order, then one cannot proceed to the difference principle if the liberty principle has not received adequate support by the procedure. Or perhaps what Ricœur has in mind is that the liberty principle is not *artificially* (i.e. not procedurally) supported by the original position and the maximin rule in this circular fashion, though the difference principle is. Yet nowhere has Ricœur bolstered his case for non-seriality by showing that the liberty principle has only artificial support.

Ricœur sees that another problem flows from these questions: the argument supporting both the liberty principle and the difference principle is “autonomous,” that is to say, independent from the reflective procedure of reflective equilibrium and presumably also legislating itself without its support. There is a certain “ambivalence” in Rawls’ work according to Ricœur:

It wants to win on two fields at once, on the one hand by satisfying the principle of reflective equilibrium, and on the other by constructing an autonomous argument introduced by the hypothetical course of reflection. This explains the apparent discord between the declarations at the beginning, which assign a regulative role to our considered convictions, and the strong plea pronounced later in favor of an independent argument, of the type of the maximin rule. It may be the burden of every contractualist theory to derive from a procedure agreed upon by everyone the very principles of justice that, in a paradoxical fashion, already motivate the search for an independent argument.²⁹

Properly speaking, reflective equilibrium is not a principle, but a model of judgment consisting of a body of criteria. And even though the two principles of justice are presented *textually* before the presentation of reflective equilibrium (in conjunction with the notion of the original position and the maximin rule), that does not mean that the two principles are *independent* of the process of seeking a reflective equilibrium between principles of justice and reasonable conditions of agreement. This misunderstanding leads Ricœur to believe that the “considered convictions” of the heterodox interpretation serve a *regulative* function throughout the procedure, when in fact they are only brought in to play to rule out the most obvious misinterpretations of justice. After all, considered convictions are intended to be cleansed of contingencies by inserting them into the process of seeking reflective equilibrium. This hardly makes them “regulative.” It merits repeating that the principles of justice are not merely considered convictions that have been dressed up in rational finery. They have been arrived at in a process in which such convictions have played a modest, though noteworthy, role. Ultimately, it is felicitous but hardly necessary that our considered convictions provide some preliminary criteria for reflection in this process but ultimately agree with the findings of the maximin rule about the fair distribution of primary social goods. But surely nobody could imagine that we already had a considered conviction – to the effect that social and economic inequalities are acceptable on two conditions, one of which is that they should be to the benefit of the least advantaged – before establishing the original position through reflection, and even before the introduction of the maximin rule (after all, as we have seen, Rawls refers to the “considered judgments” as having been “duly pruned and adjusted”). And it would be difficult to accept that the maximin rule led us to the justification of the difference principle after it had been functioning as a “regulative” considered conviction all along.

Ricœur's last statement in the paragraph above has a different status altogether. It may reveal the sort of hermeneutic circle Ricœur imposes artificially on the textual presentation of Rawls' system. It is not clear that either contractarian theories, such as those of Hobbes, Locke and Rousseau, or contractualist theories such as Kant's, suffer from a particular aporetic nature: they attempt to "derive from a procedure agreed upon by everyone the very principles of justice that, in a paradoxical fashion, already motivate the search for an independent argument." Certain principles of justice motivate the desire for an independent argument about the origin of just political organizations; yet the procedure chosen to establish this origin is one from which those self-same principles can be derived. Yet it is certainly not the case that in Hobbes, Locke and Rousseau everyone agrees on the procedure (though of course it is granted that they will agree to the "law(s)" and/or "right(s)" of nature from which the procedure unfurls), and the Kantian procedure is intelligible only within the vast critical philosophy by which the metaphysics of morals is supported by the critique of pure practical reason.³⁰ If we simply remove the notion of agreement here, what remains is that we derive from a procedure for establishing principles of justice the very same principles we had in mind when we sought such a procedure in the first place: the principles are "preunderstood" when the procedure is formulated. As I understand them, contractarian and contractualist theories have a tendency to display their principles from the outset, normally in the form of "laws" or "rules" of nature, or a "universal principle of right." And they strive to formulate a social contract that agrees with such principles. Even when presenting his principles "tentatively" early in the book, Rawls takes great pains not to make the same mistake, as is clear when one compares the procedure of *A Theory of Justice* and what he writes about the contractarians in his *Lectures on the History of Political Philosophy*. In fact, that is what makes his "procedural" definition of justice different in kind from contractarian definitions (primarily because they are not procedural) and different in formal structure from Kant's procedural definition. To insist that Rawls' achievement suffers from the same circular flaw shows a surprising lack of sensitivity to that achievement and lack of charity towards its grounding intent. In general terms, the problem originates with Ricœur's understanding of the correlation between Rousseau and Rawls' positions.

Ricœur declares that, "in order to dramatize what is at stake, I want to suggest that justice as fairness – as procedural fairness – aims at resolving Rousseau's well-known paradox of the legislator."³¹ Ricœur provides a long quote in the text and two others in a footnote to illustrate this "paradox." The most apt under the circumstances would be Rousseau's claim that human beings would already have to be before the law is instituted what they should become by means of the institutionalization of law. The very spirit of sociality that agreement is intended to establish would have to inspire and preside over that establishment. Simply put, people would already have to be *x* (sociable, etc.) in order to establish the conditions of *x*. Rousseau's solution to the problem, beyond the quick fix of the "legislator," is virtually indistinguishable from his view of history: minimal but sufficient degrees of *x* made the establishment of full or efficacious *x* possible (indeed, in a sense, necessary). Now, Rawls has been aware of this problem and its proposed solutions. He dedicates parts of two lectures to them.³² In neither case is it clear that Rawls's commitment to "justice as fairness" is intended to resolve this problem, either as Rousseau formulates it or in general. One might say that the initial, formal presentation of the original position, in being nonhistorical, is also without an original event. Rather, it is the device for representing the null-site or nexus of a number of criteria involving the expurgation of certain contingencies.

Now, for Ricœur, all of his reading of Rawls inspires (and has been led back to!) a question about the role of rational arguments in ethics. Can one conceive of a hypothetical situation in which rational arguments can stand in the place of prior convictions? Or should rational arguments serve only to clarify, however critically, the meaning of such convictions? Ricœur's Rawls, again, seeks to answer both of them in the affirmative: to offer a procedural theory of justice without sacrificing reflective equilibrium. Ricœur claims that, for his own part, "the preunderstanding of the just and the unjust assures the deontological intention of the self-proclaimed autonomous argument, including the maximin rule," a rule which, if separated from the domain prescribed by the Golden Rule, would remain a merely prudential argument, a quasi-economic argument "swerving towards being a pseudo-economic argument" if uprooted from our considered convictions.³³ It merits repeating that there is a significant difference between, on the one hand, the criteria and domain of our considered convictions when exposed to critique within reflective equilibrium and, on the other, the sorts of contingencies and prejudices of the Greco-Roman and Judeo-Christian traditions Ricœur believes have informed the deontological intention behind the maximin rule. Given his work on the Golden Rule in *Oneself as Another*, it is completely understandable that Ricœur would emphasize this, but it is unclear that it serves as an adequate framework within which to defend this role for the maximin rule. The problem here is that the inspiration for Rawls' work on judgment from a common and agreeable point of view is nothing directly Judeo-Christian as Ricœur would have it, but rather Hume's "judicious spectator" theory, which most definitely implies a certain critique of the Golden Rule.³⁴

Ricœur himself states that this "first suggestion" only makes sense in conjunction with another. Our very preunderstanding of the just and the unjust requires critical evaluation. We must strive to eradicate "ideologically biased prejudices" from our considered convictions.³⁵ And what matters here is that such prejudices must be understood to include "specifying premises," the delimiting criteria that specify the boundaries of a given judgment. It seems obvious that this is precisely the problem Rawls seeks to surmount. If his work fails to provide specifying premises that withstand the kind of "critique" to which Ricœur vaguely alludes, then it is unclear whether anyone else has ever succeeded (including the Aristotle of the arithmetic and proportional equality models of justice Ricœur compliments here). It may be that Rawls' procedural theory fails to do this, yet it seems doubtful that its failure is owed to an inability to "critique" its considered convictions in the vague way that Ricœur suggests. In fact, it is arguable that Ricœur has misrepresented the role of such convictions, in particular by implying that they serve as criteria that remain binding throughout the process of reflective equilibrium, or at least are not sufficiently criticized as to eliminate potential prejudices. Even if a considered conviction is not radically transformed or refined through reflection, that does not mean it is a prejudice that has slipped through the net of the procedure. And since it seems very unlikely that Ricœur is correct to think of the maximin rule as a discrete ethical principle that was already a considered conviction of all exchange relations,³⁶ it is unlikely that any specific "ideological" prejudice is hidden within it.

Of course, if the basic contention of Ricœur's reading of Rawls is that even sharply critical procedural theories of justice should be exposed to sharply critical procedural evaluations, then no-one is more aware of that difficulty than Rawls himself. His later reformulations of the position of *A Theory of Justice* in *Political Liberalism* and *Justice as Fairness* serve as the most obvious evidence of that. But that would require a more careful and nuanced reading of Rawls than Ricœur himself provides. And moreover, if part of the critique would be that Ricœur's own conception of ethics –

ethical presuppositions and ethical argumentation – would be taken for granted, then we have reason to believe that it is Ricœur's approach to Rawls, and not Rawls' perspective on justice, that is the most potentially biased. Most specifically, if something like the "ethics" of the "Golden Rule" Ricœur has in mind is being used to critique Rawls on the grounds that it presumes it while denying the presumption, then careful argumentation would be needed to show that. Under the circumstances, the most that can be conceded as plausible is that the approach to the Golden Rule that Rawls inherits critically from Kantian constructivism (and Hume "judicious spectator" theory) is considerably sharper and more nuanced than Ricœur's own more aretaic conception of it.

Perhaps the problem with Ricœur's hermeneutic reading of Rawls is that he is not sufficiently clear about what is being presupposed. Among the various candidates of an illegitimate or "artificial" (non-procedural) presupposition are: on the one hand, the liberty and/or difference principles, either as "considered convictions" or not, and on the other, an ethical foundation pertaining to "the good" or an ethical argument implying some truth about the Golden Rule. It seems that Ricœur has an idea of a constellation of such presuppositions being antecedent to the procedural theory of justice. Intriguing as that may be, we still need to be clear on precisely what is at stake if any such presuppositional scheme is in place. Arguably, Ricœur has not made a satisfactory case that the difference and liberty principles are (or agree with) considered convictions that anticipate their discovery and establishment. Equally, even if there is some conception of the good serving as a silent ethical foundation in the background, it does not follow that it is anything like the Golden Rule Ricœur has in mind. Ultimately, Ricœur has not demonstrated that there is a single presuppositional form that renders the procedure self-defeating. Instead, he has proposed to us several potential forms of damaging presupposition, each of which is based on a questionable reading of Rawls' text.

- ¹ Jean-Luc Nancy, *Being Singular Plural*, trans. R. D. Richardson & A. O. Byrne (Stanford: Stanford University Press, 2000), 24.
- ² The notion of “constitutive antecedence” is commonly present in hermeneutic theory since at least Heidegger and is present in Ricœur’s work by other names. I am influenced here by Jean-Luc Nancy’s “Sharing Voices,” which is to be found in *Transforming the Hermeneutic Context: From Nietzsche to Nancy*, eds. Gayle L. Ormiston & Alan D. Schrift (Albany: State University of New York Press, 1990), 224ff.
- ³ This is not to say that there are not important differences between these articles, but these differences have more to do with political liberalism than with procedure. Although it is not mentioned on the “Sources of Original Publication” page in *The Just*, this article is a translation of “Cercle de la démonstration dans *Theory of Justice* (John Rawls),” *Esprit*, no.2 (1988).
- ⁴ Paul Ricœur, *The Just*, trans. David Pellauer (Chicago: The University of Chicago Press, 2000), 37. I have found two commentaries helpful, though their goals are largely different than my own here. The first section of Molly Harkirat Mann’s Ricœur, *Rawls and Capability Justice: Civic Phronesis and Equality* (London: Bloomsbury/Continuum, 2014) is very intriguing. I found section 1.1.2.3 in Maureen Junker-Kenny’s *Religion and Public Reason: A Comparison of the Positions of John Rawls, Jurgen Habermas and Paul Ricœur* (Berlin: De Gruyter, 2014) to be especially relevant to my concerns here. I also admire David Pellauer’s “Looking for the Just,” in *Paul Ricœur and the Task of Political Philosophy* (Lanham, Maryland: Lexington, 2014), 51-64. For a fine survey, see also Bernard P. Dauenhauer, *Ricœur: The Promise and Risk of Politics* (Lanham, Md: Rowman and Littlefield, 1998), 177-82.
- ⁵ Actually, no such difficulty is revealed in these passages of *The Metaphysics of Morals*. The kind of difficulty to which Ricœur alludes is better discovered in “Theory and Practice.” See “On the common saying: ‘this may be true in theory, but it does not apply in practice,’” in Immanuel Kant, *Political Writings*, trans. Hans Reiss, ed. H.B. Nisbet (Cambridge: Cambridge University Press), 79.
- ⁶ Ricœur, *The Just*, 37-8.
- ⁷ Paul Ricœur, *Oneself as Another*, trans. Kathleen Blamey (Chicago: The University of Chicago Press), 236.
- ⁸ Ricœur, *The Just*, 39; Ricœur, *Oneself as Another*, 228.
- ⁹ John Rawls, *Justice as Fairness: A Restatement* (Cambridge, MA: Harvard University Press), pages 14-16 & 83. And John Rawls, *Political Liberalism* (New York: Columbia University Press, 1993), 27.
- ¹⁰ For the former, see Immanuel Kant, *Groundwork of the Metaphysics of Morals*, trans and ed. Mary Gregor (Cambridge: Cambridge University Press, 1997), 36 and 43-4. For the latter, see Immanuel

Kant, *The Metaphysics of Morals*, trans. and ed. Mary Gregor (Cambridge: Cambridge University Press, 1996), 23-24.

¹¹ See Ricœur, *Oneself as Another*, 232.

¹² John Rawls, *Lectures on the History of Political Philosophy*, ed. Samuel Freeman (Cambridge, MA: Belknap Press, 2007), 19.

¹³ Ricœur, *The Just*, 50; see also Ricœur, *Oneself as Another*, 236.

¹⁴ John Rawls, *A Theory of Justice: Revised Edition* (Cambridge, MA: Harvard University Press, 1999), 11.

¹⁵ Rawls, *Lectures on the History of Political Philosophy*, 15.

¹⁶ Rawls, *A Theory of Justice: Revised Edition*, 52.

¹⁷ Ricœur, *The Just*, 51.

¹⁸ Ricœur, *The Just*, 51.

¹⁹ Ricœur, *The Just*, 52.

²⁰ Rawls, *A Theory of Justice: Revised Edition*, 3.

²¹ Ricœur, *The Just*, 53.

²² Ricœur, *The Just*, 66.

²³ Rawls, *A Theory of Justice: Revised Edition*, 17.

²⁴ Ricœur, *The Just*, 54; see also Ricœur, *Oneself as Another*, 237.

²⁵ Rawls, *A Theory of Justice: Revised Edition*, 18. (Italics mine.)

²⁶ Ricœur, *The Just*, 55.

²⁷ Rawls, *A Theory of Justice: Revised Edition*, 42-3.

²⁸ Elsewhere Rawls shows he grasps the significance of this point (*A Theory of Justice*, 70). See *Lectures on the History of Political Philosophy*, §§3-4 (pp. 11-6) in particular.

²⁹ Ricœur, *The Just*, 56.

³⁰ In this context, Ricœur does not reveal that he is familiar with the difference between contractarian and contractualist theories, only some general tendencies of classical social contract theory (*The Just*, 62).

³¹ Ricœur, *The Just*, 40; see also Ricœur, *Oneself as Another*, 229.

³² Rawls, *Lectures on the History of Political Philosophy*, 197-204 and 237-41.

³³ Ricœur, *The Just*, 56.

³⁴ See Rawls, *Lectures on the History of Political Philosophy*, 184-7, where it is clear that many of Rawls' ruminations about objectivity in judgments are informed by Hume's work. This is confirmed by a reading of Rawls' *Lectures on the History of Moral Philosophy*, ed. Barbara Herman (Cambridge: Harvard University Press, 2000), 84-104.

³⁵ Ricœur, *The Just*, 56.

³⁶ Ricœur, *The Just*, 56.