On the Rawlsian Anthropology and the "Autonomous" Account

Jared Mayer
Johns Hopkins University
On Autonomous Accounts and Rawlsian Anthropologies

Jared I. Mayer – Johns Hopkins University

In his book *Political Liberalism*, John Rawls attempts to provide an account for how to implement principles of justice in a diverse and ideologically pluralistic society. Critical to achieving this goal is the implementation of a political conception of justice that is “presented as a freestanding view.”¹ In other words, “it is neither presented as, nor as derived from, such a [comprehensive] doctrine applied to the basic structure of society.”² Joseph Raz, however, believes that Rawls’ account leaves him choosing either the kind of autonomy he wishes to have, or the hope of having any practical value when the autonomous account conflicts with other political and moral values. In response to Raz, Rawls can say that Raz is making a category mistake by assigning the political conception of justice the properties of other moral and political values, something which Rawls explicitly denies with his conception of a freestanding political conception of justice. To further insulate his freestanding account from ordinary moral and political values, Rawls creates a “political conception of the person” to match his political conception of justice.

It is not clear, however, that Rawls can evade Raz’s critique so quickly. In this paper, I argue that Rawls creates a “dual aspect” philosophical anthropology. Such an account presupposes not a metaphysical claim, but a “metaphysical enough” claim about the nature of persons and their moral capacities. This is precisely the kind of conjecturing Rawls needs to avoid if he wishes for his theory to remain – indeed, be in the first instance – autonomous. And if my understanding of Rawls’ dual aspect philosophical

¹ In this essay, I use “freestanding” and “autonomous” interchangeably.
anthropology is correct, then, in fact, Raz’s supposed “category mistake” is no such thing at all. Since Rawls must provide some account of philosophical anthropology, Rawls finds himself confronted with either keeping the autonomy or the practicality. Something’s got to give.

We begin with a brief sketch of Rawls’ notion of a political conception of justice. Early on in the First Lecture, Rawls provides a crisp program for his conception of justice:

[The political conception of justice] presents itself as a conception of justice that can be shared by citizens as a basis of a reasoned, informed, and willing political agreement. It expresses their shared and public political reason. But to attain such a shared reason, the conception of justice should be, as far as possible, independent of the opposing and conflicting philosophical and religious doctrines that citizens affirm.  

Here, a few key elements of Rawls’ project shine through. First, citizens are bound to subscribe to mutually exclusive comprehensive doctrines. Second, in order to create a just society, one that is workable given the set of mutually exclusive comprehensive doctrines, there must be some epistemic and political mechanism by which we are able to cast aside, or perhaps transcend, our comprehensive views. Third, once those epistemic and political mechanisms are in place, “reasoned, informed, and willing political agreement” will ensue. The best way to reckon with these claims, according to Rawls, is to present his political conception of justice “as a freestanding view.” This means that while the account can and should be supported by a range of reasonable comprehensive views, “it is neither presented as, nor as derived from,” any of those views. What’s more, Rawls’ account is not intended to apply to all persons in all circumstances at all times;

---
3 Rawls, ibid. At 9.
4 Rawls, ibid. At 12.
5 Ibid.
rather, its intended use is solely for the “basic structure” of a society. Given all of this –

especially the fact that we are trying to create a just society despite pervasive
disagreement – Rawls claims that because his account is to be endorsed by all, such that

“the political values expressed by its principles and ideals normally have sufficient
weight to override all other values that may come in conflict with them.” Since all
citizens can endorse the political conception of justice, and no comprehensive account
can possibly serve in that capacity, it is imperative upon them to uphold the political
conception of justice as supreme over their own comprehensive doctrines.

In an essay responding to Rawls, Joseph Raz claims that Rawls wishes to enshrine
a “strongly autonomous” account of the political conception of justice. “Strongly
autonomous,” for Raz, denotes that the conception may only “appeal to individual well-
being in justification of political values and institutions.” Rawls’ account is autonomous
precisely because he wants to claim that it is not in the business of evaluating the truth or
goodness of any reasonable moral or political value besides individual well-being. If
this is so, then, according to Raz, two possible objections can arise. First, when one is in
the business of providing and justifying moral and political values of any kind, one is

---

6 Ibid. At 11. The basic structure, for Rawls, is “a society’s main political, social, and economic
institutions, and how they fit together into one unified system of social cooperation from one generation to
the next.” Ibid.
7 Ibid. At 138.
8 I use the term “supreme” not so much to connote a sense of the all-encompassing nature of the political
conception of justice – that would just be incorrect – but rather to connote a sense that the political
conception of justice must be seen as overriding comprehensive doctrines with respect to fundamental
questions that affect the basic structure. In this sense, supremacy is taken the same way that it is taken in
the U.S. Constitution: “This Constitution, and the Laws of the United States which shall be made in
Pursuance thereof; and all Treaties made, or which shall be made, under the Authority of the United States,
shall be the supreme Law of the Land; and the Judges in every State shall be bound thereby, anything in the
Constitution or Laws of any State to the Contrary notwithstanding.” See U.S. Constitution, Article VI, Sec.
2 (emphasis added).
10 Ibid. At 153nn.
11 Rawls, ibid. At 94; Raz, ibid. At 152-153.
required to fit those values into the web of other beliefs that make up a larger account of 
human well-being. Second, since the first point holds, the practical implication of any 
given value is determined by its interaction with conflicting values and the method by 
which we decide which value trumps the other. Yet because the practical implication of a 
value is itself dependent on that value’s place in the greater constellation of values in a 
single conception of human wellbeing, there is no practical implication to any value 
without a conception of human wellbeing. Rawls‘ strongly autonomous account, then, is 
left disarmed.

But Rawls need not surrender so easily. He might respond to Raz by saying that 
Raz has made a category mistake. In making his two objections to Rawls’ “strongly 
autonomous” account, Raz compares the account with “moral and political values” writ 
large. But who’s to say that Rawls would accede to that comparison? Indeed, Rawls’ 
project is premised on a distinction between, broadly speaking, the political and the 
personal. The political realm is the basic structure, and the basic structure is governed by 
the political conception of justice. Comprehensive moral doctrines, by contrast, may not 
serve as the guiding spirits for the state and its coercive powers. This is precisely because, 
as we have mentioned earlier, they are hotly contested, and indeed may be strongly 
opposed by, fellow citizens. And while the political conception of justice ought to be 
supported by various reasonable comprehensive doctrines, it does not derive its 
normative force from that support.\textsuperscript{12} So in comparing the political conception of justice 
with other kinds of moral and political values, Raz is incorrectly assuming that the 
properties of the other kinds of moral and political values can apply to the political 
conception of justice.

\textsuperscript{12} Rawls, ibid. At 12.
There is, however, one missing link from Rawls’ account as provided above. As we just noted, Rawls bases his project in _Political Liberalism_ on his distinction between the political conception of justice and comprehensive moral views. Yet in order to fully cement this distinction, Rawls needs to provide an account of how citizens, given their comprehensive doctrines, can set these doctrines aside when deliberating fundamental questions. In other words, he needs to provide an explanation for how citizens ought to see _themselves_ with respect to the basic structure and the political conception of justice.

Rawls’ philosophical anthropology is an attempt to do just that. He begins by noting that this philosophical anthropology is created in order to provide some understanding of what persons in the original position are like.\(^\text{13}\) In light of this, Rawls importantly notes that “in virtue of their two moral powers (a capacity for a sense of justice and for a conception of the good) and the powers of reason… persons are free.”\(^\text{14}\) Fair enough, I suppose. But persons have another aspect to them, a political aspect.

In this vein, Rawls proposes three senses in which citizens – in short, persons in their political aspects – see themselves as free. The first is that they are free “in that they conceive of themselves and one another as having the moral power to have a conception of the good.”\(^\text{15}\) In this respect, persons have two aspects to their identity – broadly speaking, the personal and the political.\(^\text{16}\) In the personal realm, citizens can have profound changes of heart, ones that press them to radically alter their comprehensive

---

\(^\text{13}\) Ibid. At 29. The original position is the thought experiment wherein persons therein are behind a veil of ignorance, where they are blind to their own future life conditions, such as their race, ethnicity, class, sex, and so forth. The purpose is to serve as the proper environment in which the political conception of justice can be born. Since each person is unaware of the conditions in which he or she may be born into, there can be no claim that the political conception of justice is biased towards someone’s own aims. By the same token, however, it is unclear _just what a person in the original position would be like and how their epistemic processes would operate_. Hence, a philosophical anthropology is needed.

\(^\text{14}\) Ibid. At 19.

\(^\text{15}\) Ibid. At 30.

\(^\text{16}\) Ibid.
moral outlook. In the political realm, however, the citizen, regardless of whatever changes he or she made to his or her comprehensive moral view, maintains the same moral identity. This is because the citizen remains committed to “the values of political justice and want[s] to see them embodied in political institutions and social policies.”

The second sense in which citizens see themselves as free is that “they regard themselves as self-authenticating sources of valid claims.” This means that citizens see themselves as the root of normative claims that they make, especially (and most relevant here) those normative claims they make towards and about their governing institutions. The third and last sense in which citizens view themselves as free is that “they are viewed as capable of taking responsibility for their ends and this affects how their various claims are assessed.” In other words, with just institutions in play, citizens feel capable of altering their expectations and desires with respect to how much they are able to give. The result is practical and cohesive: citizens are capable of recognizing that the demands of justice are not responsive to the degree of one’s desire for a particular end. In these three respects, then, citizens see themselves as free.

But it is precisely in these respects that Rawls’ account proposes something of a metaphysical account of the person, one that heavily draws upon the Kantian spirit. Rawls’ notion that citizens can “conceive of themselves and one another as having the moral power to have a conception of the good,” plays upon Kant’s notion that rational

---

17 Ibid. At 31.
18 Ibid. At 30-31.
19 Ibid. At 32.
20 Ibid. This is so, so long as such claims fall within the confines of the political conception of justice. Ibid.
21 Ibid. At 33.
22 Ibid. At 34.
beings act in accordance with their conceptions of the law in mind.\textsuperscript{23} His second way of how citizens see themselves as free – that “they regard themselves as self-authenticating sources of valid claims” – smacks of Kant’s ideal of a rational being in the kingdom of ends who “is universally legislating in it, but also itself subject to [universal] laws.”\textsuperscript{24} And lastly, citizens “viewed as capable of taking responsibility for their ends [which] affects how their various claims are assessed” hints at the disposition Reason needs to “view herself as the authoress of her principles, independently of alien influences, and must consequently… by herself be viewed as free.”\textsuperscript{25} These parallels, then, suggest that Rawls’ conception of how the citizen sees herself as free is fraught with a somewhat metaphysical understanding of the citizen and her relationship to her fellow citizens.

An example might be instructive in drawing my claim out a bit further. Suppose a citizen converts from one faith to another, or in fact moves from one faith to no faith at all. Her driver’s license will in no way be affected – as many would think it should not be. Her personal status is, in this respect, immaterial with respect to her status \textit{qua} citizen. But notice how Rawls plays upon this intuitive appeal by applying it to a distinctly moral framework. Rawls’ point here is to say that, \textit{qua} citizen, everyone remains committed to the project of furthering the political conception of justice. Yet equally intuitive is the sense that this kind of identity framework – where one’s political moral identity is static, yet personal moral identity is open to (hypothetically) endless revision – is a far cry from that of the driver’s license, which posits a miniscule claim

\textsuperscript{24} C.f. \textit{Groundwork}, 4:433.
\textsuperscript{25} C.f. \textit{Groundwork}, 4:448.
about the moral identity of the driver, if any at all. And this is the upshot of it all: Rawls’ notion of how the citizen sees him or herself as free is that, at one and the same time, citizens must be committed to seeing themselves as both morally static and morally changing, not operating in a unified, coherentist process, but rather with two separate aspects.

Now, to be sure, it would be incorrect of me to claim that Rawls’ philosophical anthropology is wholly or fully metaphysical. Even if Rawls’ account is Kantian in its formulation and its spirit, the evidence for this claim is insufficient to demonstrate its overarching metaphysical character. But this is not the necessary task at hand; all one needs to do is demonstrate that Rawls’ philosophical anthropology is metaphysical enough to cause fractures between his philosophical anthropologies and competing anthropologies stemming from reasonable comprehensive doctrines. By claiming that citizens can divide themselves between their personal and political aspects, and indeed by claiming that the political aspect governs the boundaries of the personal aspect, Rawls’ philosophical anthropology is already committed to something of a metaphysical division between personal and political, one that can be certainly contested by some comprehensive moral views while supported by others. Since this is the case, Rawls’

26 I say miniscule because it possibly posits, for example, a sense that one has an obligation to follow traffic laws.
27 And indeed, Rawls may well claim that any such attack on his account would either a) presuppose that there can be a philosophical anthropology without metaphysics, or b) require that we argue over distant metaphysical topics, idealism, principles of identity, etc. Ibid. At 29n. Such objections, however, mischaracterize the objectors. No assumption with respect to the need of metaphysics (or lack thereof) in philosophical anthropology is required. All that is being argued is that Rawls’ particular explanation needs to be argued for. And what is more, the arguments over distant metaphysical topics are just that: distant. While they may arise in debates over philosophical anthropologies, we should not pretend as though it is these topics that would arise in a critique of Rawls’ account. After all, if, say, John Finnis were to evaluate Rawls’ political anthropology, his first target of all things would not be the identity of indiscernibles.
28 Indeed, adherents to traditional faiths, such as Roman Catholics and Orthodox Jews, would see themselves as being inseparable in this respect. While their comprehensive moral views may certainly endorse the political conception of justice, that is, for the sake of evaluating Rawls’ political anthropology, neither here nor there.
philosophical anthropology, and thus political conception of justice, cannot claim to be autonomous.

The consequence of this new understanding of Rawls’ project is that Raz’s first objection to Rawls’ “strongly autonomous” account is answered by biting the bullet. If Rawls is in the business of providing a philosophical anthropology, then surely he is in the business of providing some way in which his philosophical anthropology’s underpinning commitments relate to, reckon with, and perhaps even trump other moral and political values. This is because, again, any account of the moral nature of persons, or more precisely, how one’s moral capacities are recognized and evaluated by the agent will inevitably require the agent to fit those commitments with the agent’s other moral and political values. Rawls, then, is engaged in something of a project like the one Raz mentioned in his first objection: fitting those values into the web of other moral and political beliefs that make up a larger account of human well-being.

Once Rawls does this, then it is clear that Raz’s second objection is sufficiently defused. In other words, once Rawls admits that his project, in fact, requires the agent spin a web of moral and political commitments that makes them both coherent and workable, Rawls is then able to provide an answer for just how these political and moral values are to interact with each other in times of conflict. And recall that his answer was featured earlier on in this essay: “with respect to… fundamental questions, the political values expressed by principles and ideals [in Rawls’ account] normally have sufficient weight to override all other values that may come into conflict with them.”

29 Which isn’t to say that one has an entire system of just how these relationships would work out. Rather, if one is going to provide a philosophical anthropology, one needs to answer just how those relationships will operate.
30 Raz, ibid. At 153.
31 Rawls, ibid. At 138.
understanding of Rawls’ political conception of justice and how it interacts with, and
indeed presupposes, his philosophical anthropology allows Rawls to justify the priority of
his political conception of justice over other political and moral values.

Yet such an allowance comes at a cost. For Rawls cannot maintain his “strong
autonomy” and be able to sufficiently provide an account of why his political conception
of justice is entitled to govern over other political and moral values. Once again,
“strongly autonomous” accounts are those that solely “appeal to individual well-being in
justification of political values and institutions.”32 Yet Rawls’ philosophical anthropology,
which undergirds his political conception of justice, appeals to much more than that – in
essence, to a difference between persons qua persons and persons qua citizens. Since this
distinction is metaphysical enough, Rawls’ account cannot be autonomous in the strong
sense. And because Rawls’ account is underpinned by his philosophical anthropology, it
doesn’t seem to be an option for him to discard that philosophical anthropology and
thereby open the channel for strengthening his account’s autonomy. The only option left
for Rawls is to discard the strong autonomy.

I would like to conclude by musing as to whether biting this bullet will cause
Rawls’ entire theory to shatter. I am somewhat skeptical of that possibility, for I think one
can reconstruct Rawls’ account in such a way so as to mitigate the effects of having to
surrender the strong autonomy. Indeed, it may well be the case that Rawls will need to
retreat to a “Kantian Interpretation,” similar to the one he provided in A Theory of
Justice.33 Nevertheless, if this understanding of Rawls is correct, Rawls may first have to
rethink how his account is to be legitimized in a pluralistic polity.

32 Ibid. At 153nn.
University Press, 1999), 221-227.
Works Cited


