

# PUBLIC REASON

Journal of Political and Moral Philosophy

Volume 10, Number 1, 2018



# PUBLIC REASON

## Journal of Political and Moral Philosophy

*Public Reason* is a peer-reviewed journal of political and moral philosophy. *Public Reason* publishes articles, book reviews, as well as discussion notes from all the fields of political philosophy and ethics, including political theory, applied ethics, and legal philosophy. The Journal encourages the debate around rationality in politics and ethics in the larger context of the discussion concerning rationality as a philosophical problem. *Public Reason* is committed to a pluralistic approach, promoting interdisciplinary and original perspectives as long as the ideal of critical arguing and clarity is respected. The journal is intended for the international philosophical community, as well as for a broader public interested in political and moral philosophy. It aims to promote philosophical exchanges with a special emphasis on issues in, and discussions on the Eastern European space. *Public Reason* publishes two issues per year, in June and December. *Public Reason* is an open access e-journal, but it is also available in print.

### Editors

Editor in Chief

Romulus Brancoveanu, *University of Bucharest*

Sorin Baiasu, *Keele University*

### Associate Editor

Thomas Pogge, *Yale University*

### Editorial Team

Assistant Editors

Mircea Tobosaru, *Politehnica University of Bucharest*

Cristina Voinea, *University of Bucharest*

Carmen-Viviana Ciachir, *University of Bucharest*

### Editorial Board

Ovidiu Caraiani, *University Politehnica of Bucharest*

Luigi Caranti, *University of Catania*

Radu Dudau, *University of Bucharest*

Mircea Dumitru, *University of Bucharest*

Adrian - Paul Iliescu, *University of Bucharest*

Ferda Keskin, *Istanbul Bilgi University*

Valentin Muresan, *University of Bucharest*

Alina Romascu, *Université de Corse Pasquale Paoli*

Constantin Stoenescu, *University of Bucharest*

Ion Vezeanu, *University of Grenoble*

### Advisory Board

Radu J. Bogdan, *Tulane University*

Paula Casal, *University of Reading*

Fred D'Agostino, *University of Queensland*

Rainer Forst, *Goethe University, Frankfurt am Main*

Gerald Gaus, *University of Arizona*

Axel Gosseries Ramalho, *Catholic University of Louvain*

Alan Hamlin, *University of Manchester*

John Horton, *Keele University*

Janos Kis, *Central European University, Budapest*

Jean-Christophe Merle, *University of Tübingen*

Adrian Miroiu, *SNSPA Bucharest*

Adrian W. Moore, *University of Oxford*

Philippe Van Parijs, *Catholic University of Louvain*

Mark Timmons, *University of Arizona*

*Public Reason* is available online at <http://publicreason.ro>

ISSN 2065-7285

EISSN 2065-8958

© 2018 by *Public Reason*

# PUBLIC REASON

Journal of Political and Moral Philosophy

Vol. 10, No. 1, 2018

## ARTICLES

*Julian Katz*

Kant's 'Bund': A Voluntary Reading ..... 3

*Leonidas Makris*

Mill's Political Perception of Liberty: Idiosyncratic, Perfectionist but essentially Liberal ..... 15

*Alexandros Manolatos*

Rawlsian Stability and the Hazards of Envy. .... 33

*Jens Peter Brune*

Capabilitarianism without Paternalism? ..... 49

*Hossein Dabbagh*

The Problem of Explanation and Reason-Giving Account of *pro tanto* Duties in the Rossian Ethical Framework ..... 69

*Adelin Costin Dumitru*

Feminism, Agency and Objectivity. .... 81

## Kant's 'Bund': A Voluntary Reading

Julian Katz  
Tulane University

**Abstract:** In 'Kant's Changing Cosmopolitanism' and *Kant and Cosmopolitanism: The Philosophical Ideal of World Citizenship*, Pauline Kleingeld argues that, in 'Idea for a Universal History with a Cosmopolitan Intent,' Kant meant for the *Bund* (usually translated as 'federation') of states to be a coercive federation. Kleingeld admits that there is a disparity between this earlier coercive idea of the *Bund* and Kant's talk of a voluntary (i.e., non-coercive) congress in *Toward Perpetual Peace* and *The Metaphysics of Morals*. She explains this disparity by: (1) appealing to a semantic ambiguity in the term '*Bund*'; (2) making claims about with which contemporary 18<sup>th</sup> century works Kant was acquainted; and ultimately (3) attempting to draw a parallel between the unsocial sociability of individual people within a state and the unsocial sociability of individual states in a larger community of states. In this paper, I argue that while Kleingeld's claims are superficially supported by the text, her claims depend on her apparent conflating of teleology and morality.

**Key words:** Kleingeld, Kant, teleology, Just State, Bund, Federation, Congress, Inter-Governmental Organization.

Kant's contribution to the peace project of the 18<sup>th</sup> century (i.e., the Just State considered both in a local civil condition and considered globally as a union of states) is still a great source of disagreement among scholars. This disagreement mirrors contemporary political debates concerning what it is about certain Intergovernmental Organizations (IGOs) (e.g., the UN, the European Union, the G8, and the United States) that leads to, or maintains, peace.<sup>1</sup> After all, we seem to have had some peace in the latter half of the 20<sup>th</sup> century and even now; but there is still the question of how IGOs contribute to this peace. IGOs could bring about peace through coercive means *or* it could be that the particular form of democracy in many of the countries that constitute the members of those IGOs promote a voluntary peace. Regarding the latter, for example: it seems that countries that hold similar political values seem to *not* go to war as easily, especially when those countries are democratic (Doyle 1983a, 212-13).<sup>2</sup> But what exactly it means for countries to share political values is itself a point of argument. For instance, in his book, *Kant's Political Legacy: Human Rights, Peace, Progress*, Luigi Caranti discusses the disagreement as to whether *liberal democracies* and *republics* are identical (2017, 9). In *Toward Perpetual Peace* Kant, himself, seems to make a distinction between democracies and republics:

So that a republican constitution will not be confused with a democratic constitution (as usually happens), the following must be noted. [...] [*D*]emocracy in the strict sense of the word is necessarily a *despotism* because it establishes an executive power in

1] I use the abbreviation 'IGO' in line with Caranti's usage of the same (2017, 10, 164, 193, 197-98).

2] There is disagreement on this claim. For further discussion see: Caranti (2017), Doyle (1983), and Babst (1964; 1972).

which all decide for and, if need be, against one (who thus does not agree), so that all, who are nevertheless not all, decide; and this is a contradiction of the general will with itself and with freedom. (*TPP*, 8:351-52)<sup>3</sup>

I will leave aside the obvious issue of what sort of voting situation Kant expects in republics that would *not* lead to a similar situation where the majority have somehow left out the minority from particular decisions. But suffice it to say, there is much one can say on whether countries that seem to share enough of the right democratic republican values can promote a peace. For my purposes, though, I wish to discuss Kant's views on what kind of IGO contributes to a world peace; and I will be including Caranti's discussion of the historical context of Kant's view (Caranti 2017, 108-15). Kant argues that a very specific kind of intergovernmental organization is necessary for the peace between nations; however, there is disagreement, as one would expect, as to what constitutes the *just* IGO for Kant's perpetual peace.

In the *Metaphysics of Morals*, Kant tells us that, even among neighboring states, we must (that is, we have a duty to) leave the state of nature among nations "to enter a lawful condition." He describes this lawful condition (what we are calling an IGO) among states as a "voluntary coalition of different states which can be dissolved at any time, [and specifically] not a federation (like that of the American states) [...]" (*MM*, 6:350-351).<sup>4</sup> In other words, it is to be a congress of states without a coercive federal government. So, even though the Just State (JS), at the local level, has this form of government, the congress of states (the JS construed as a universal state, in other words) has no such federal government.<sup>5</sup> This is in contradistinction to Pauline Kleingeld's position that Kant really did, at one time in his intellectual development, mean for his 'congress' to ultimately be one ruled by a single federal government analogous to the government of a local state (2009, 173-74, 177-79; 2012, 45-46). She explains this development by: (1) appealing to a semantic ambiguity in the term 'Bund'; (2) making claims about with which 18<sup>th</sup> century works Kant was acquainted; and ultimately (3) attempting to draw a parallel between the unsocial sociability of individual people within a state and the unsocial sociability of individual states in a larger community of states. In this paper, I argue that while Kleingeld's claims are superficially supported by the text, her claims depend on her apparent conflating of teleology and morality.

---

3] *TPP* refers to *Toward Perpetual Peace*.

4] *MM* refers to the *Metaphysics of Morals*.

5] This is contextually important to a larger project which discusses the local and global conditions of the JS. In the course of my larger project, I raise the point that Kant's teleological view, portrayed in 'Idea for a Universal History with a Cosmopolitan Intent,' is meant only to be an optimistic view of how our teleology leads us to the same outcome as that which we ought to bring about. I will not discuss the local or global conditions of the JS. There will be some allusion, though, to the discussion of these conditions and some allusion to this agreement of our teleology with our moral goals.

## I. AN INTRODUCTION TO THE INTERNATIONAL CONDITIONS OF JS

Kant asks us to think of ourselves as cosmopolitan citizens of the world (and not just as citizens of our individual and respective countries) in both the *MM* and ‘Idea for a Universal History with a Cosmopolitan Intent’ (*MM*, 6:352; *UH*, 8:22).<sup>6</sup> In the latter, Kant has a couple of theses regarding our natural teleology that are relevant to this point. Note the fifth thesis: “The greatest problem for the human species, whose solution nature compels it to seek, is to achieve a universal civil society administered in accord with the right” (*UH*, 8:22). This thesis refers to a *universal* civil society, or in other words: a civil society “for the [whole] human species.” Skipping forward to the seventh thesis: “The problem of establishing a perfect civil constitution depends on the problem of law-governed external relations among nations and cannot be solved unless the latter is [solved]” (*UH*, 8:24). So clearly there are some implicit global conditions on what counts as a JS.

Similarly, to our duty to leave the state of nature and enter in the civil condition between individual humans in close proximity, we have a duty to leave the state of nature among nations (*MM*, 6:344-45). I write ‘similarly’, because the duties to leave the respective states of nature are actually very different. The conditions that bring on the duty are, however, the same (hence I use ‘similarly to’ and not ‘exactly as’). Just to expound briefly on the difference: The duty to enter in a civil condition involves the implementation of the irresistible constitution of the civil condition itself (the general united will) uniting all in that local civil condition. But the duty to enter into a civil condition with other nations specifically restricts that union from having such an irresistible authority over the nations in that global civil condition. Consider the following:

The elements of the right of nations are these: (1) states, considered in external relation to one another, are (like lawless savages) by nature in a non-rightful condition. (2) This non-rightful condition is a *condition* of war (of the right of the stronger), even if it is not a condition of actual war and actual attacks being constantly made (hostilities). Although no state is wronged by another in this condition (insofar as neither wants anything better), this condition is in itself still wrong in the highest degree, and states neighboring upon one another are under obligation to leave it. (3) A league of nations in accordance with the idea of an original social contract is necessary, not in order to meddle in one another’s internal dissensions but to protect against attacks from without. (4) This alliance must, however, involve no sovereign authority (as in a civil constitution), but only an *association* (federation); it must be an alliance that can be renounced at any time and so must be renewed from time to time. (*MM*, 6:344-45)

In other words, the countries are in a state of nature with each other until they form this league of nations. Even if the countries are not at war, the condition is considered non-rightful because the states are in close proximity to one another and might war with each other in the future. The potential for the sovereignty of any particular general united will (of even one state among all nations) to be overtaken by some other general united

---

<sup>6</sup>] *UH* refers to ‘Idea for a Universal History with a Cosmopolitan Intent.’

will is enough to obligate the nations to enter into a sort of civil condition with the other nations of the world. Furthermore, and in the same vein, the reason they cannot have a sort of federal government forcing them to stay in a union is that to create such a super-government would be to create something more sovereign than all the sovereign general united wills (all of the states) in such a federation. Thus, this league of nations must be a voluntary congress of states where each state can choose to leave that union at any time.<sup>7</sup>

## II. KLEINGELD'S DEVELOPMENTAL VIEW

However, there is a contention in the contemporary literature that this congress, at one time in Kant's intellectual development (specifically in UH), involved a federal government. Most notably Kleingeld argues this in her book *Kant and Cosmopolitanism: The Philosophical Ideal of World Citizenship* and in her essay 'Kant's Changing Cosmopolitanism' (2012; 2009).<sup>8</sup> In the former, for example, she writes:

In [*Idea for a Universal History from a Cosmopolitan Perspective*], [Kant] advocates the establishment of a strong federation of states with coercive authority at the federal level, and [...] he appeals to the enlightened self-interest of rulers to defend the feasibility of this ideal. He does not propose a merely voluntary association but argues that a strong, state-like federation of states is required to guarantee the external security of states. (Kleingeld 2012, 45)

In the latter, she writes:

Kant uses [the term 'cosmopolitan condition'] to refer to a situation in which the external relations among states are governed by enforced laws. This of course requires a higher-level political institution, which he refers to as a 'federation of peoples' or a 'federation of states' [...] (8:24) This is to have legislative, executive, and juridical powers at the federal level, including a 'united power giving emphasis to that law' (8:26).<sup>9</sup> (Kleingeld 2009, 174)

She essentially argues in both works that in line with what Kant argues about unsocial sociability concerning the JS at the local level, unsocial sociability can only really help us

---

7] Kant says as much: "By a *congress* is here understood only a voluntary coalition of different states which can be *dissolved* at any time, not a federation (like that of the American states) which is based on a constitution and can therefore not be dissolved. – Only by such a congress can the idea of a public right of nations be realized, one to be established for deciding their disputes in a civil way, as if by a lawsuit, rather than in a barbaric way [...], namely by war." (MM, 6:351).

8] To be fair to Kleingeld, she ultimately allows that Kant rejects this federal government over states. She reads him as having changed his mind between having written UH and *TPP*. While I admit that Kant's thought had to have evolved as he wrote, I do not think his fundamental system changed as much as Kleingeld is arguing. That said, I am only arguing against her interpretation of UH, and not also against her developmental claims about Kant's thought, in general. Such an argument would require an altogether different paper.

9] All citations within this quote are meant to refer to AK pagination in UH. They are not my citations, but they are accurate to the text.

to acquire that moral whole if the unsocial aspect of individual *nations* in a *community of nations* is tempered by the irresistible authority of a constitution over and above the states in that community.<sup>10</sup>

Part of Kleingeld's argument is purely historical. She points out that Kant is presenting his cosmopolitan ideal in a long tradition of proposals, and not in a vacuum. "[Kant] himself acknowledges the Abbé de Saint-Pierre and Jean-Jacques Rousseau as his predecessors (IaG 8:24, GTP 8:313)."<sup>11</sup> Saint-Pierre, Kleingeld points out, had argued for something like this over and above the Christian rulers of Europe "with a permanent senate and an international court of arbitration, backed up by an international military force, to settle disputes between the member states."<sup>12</sup> Rousseau had apparently summarized Saint-Pierre's views and presented them, in 1761, under the title: 'Extrait du Projet de Paix Perpétuelle de Monsieur l'Abbé de Saint Pierre' (1761). However, in 'Judgement sur la Paix Perpétuelle,' written around the same time as the former, Rousseau distances himself from Saint-Pierre's views (1782).<sup>13</sup> Kleingeld argues that it follows from this information that Kant could not have read this latter text of Rousseau's since Kant clearly wants to espouse some of Saint-Pierre's views concerning international peace.<sup>14</sup> In this historical vein, she points out that Kant "does not propose a merely voluntary association but argues that a strong, state-like federation of states is required to guarantee the external security of states" (Kleingeld 2012, 44-45).<sup>15</sup>

Another line of argument (and perhaps the main line of argument) she makes is to simply interpret the creation of a league of nations as exactly like the creation of a local JS, in that they both teleologically explain how the unsocial sociability of humanity leads to the creation of a moral whole (in the sense of the JS, anyway).<sup>16</sup> Here is a notable quote to this effect from both *Kant on Cosmopolitanism* and 'Kant's Changing Cosmopolitanism':

Kant argues, [in *Idea for a Universal History*], that the way in which states leave the international state of nature to join into a state-like federation is structurally

---

10] In the larger project I work out how the JS and the Kingdom of Ends are connected and what exactly one should understand by 'moral whole.' For now, it is sufficient that we are aware that there is a moral whole toward which unsocial sociability is meant to help us strive. Kleingeld and I agree that such a moral whole exists, though we may disagree as to what constitutes it.

11] By IaG, Kleingeld means to refer to UH. By GTP, she means to refer to 'On the Proverb: That may be true in Theory, but is of No Practical Use.'

12] She is getting this from St. Pierre's *Projet pour rendre la Paix Perpétuelle en Europe*. (1713)

13] Rousseau wrote 'Extrait' and 'Judgement' around the same time, but 'Judgement' was published posthumously in 1782. (Kleingeld 45).

14] I am here presenting Kleingeld's position; not my own. Obviously, she is assuming her position on international government to be correct in making this assertion; a position with which I explicitly disagree.

15] I will respond to each part of what I perceive to be Kleingeld's position, but only after first presenting all the lynch pins of that position.

16] I just mean that it is not clear what role the Kingdom of Ends plays with this talk of a moral whole. I resolve this in the larger project, of which this paper is a small part.

similar to the way individuals leave the state of nature to join into a state. [...] *Individuals* unite into a state 'in which *freedom under external laws* can be encountered combined, in the greatest possible degree, with irresistible power' (IaG 8:22). Similarly, Kant claims, *states* will ultimately be forced, by the hardship resulting from the rivalry and wars between them, to exit the state of nature and enter a juridical condition. States exhibit 'the same unsociability' as individuals; they experience 'precisely the ills that pressured individual human beings and compelled them to enter into a lawful civil condition,' and thus states too will come to see the advantages of joining a federation with common laws and law enforcement (IaG 8:24). (Kleingeld 2012, 45-46; 2009, 177-78)

In other words, on Kleingeld's reading, Kant has argued that individuals in the state of nature have a kind of unsocial sociability that can only have the teleology of serving to help us create a moral whole. Furthermore, this moral whole is understood (by both myself and Kleingeld) to include not just the JS at the local level, but also at the global level (i.e., not merely locally, but internationally as well).<sup>17</sup> This is apparent, she argues, from the fifth, sixth, and seventh theses in UH. The seventh thesis says essentially that the civil condition requires solving the problem of international relations between nations: "a perfect civil condition depends on the problem of law-governed external relations among nations and cannot be solved unless the latter is [solved]" (UH, 8:24). It is within the discussion of this seventh thesis that Kleingeld finds the textual evidence that states, like individuals, must enter into a union with an irresistible constitution over and above them. That is: since states exhibit the same unsocial sociability, they must need to enter into the very same type of civil condition as individuals.

Finally, Kleingeld further supplements her argument for her position through a semantic point about the chosen German word for what we see in the English as 'federation': the word '*Bund*.' She points out that '*Bund*' "is itself neutral as to whether or not the institution has the power to enforce its laws" (Kleingeld 2009, 178). Her argument here is essentially that since '*Bund*' can be understood either as (1) a coercive federal government over and above the states that are unified within it or as (2) a non-coercive voluntary congress (as I interpret the concept), it clearly can mean the former in UH and then change in meaning to the latter in *TPP*. Again, relying on the historical point made earlier about Saint-Pierre, Kleingeld claims that the semantically neutral term allows her to make this claim about Kant's position in UH.

### III. A REVIEW OF THE HISTORICAL CONTEXT OF KANT'S *BUND*

Before I show what seems odd to me about this argument, there are some historical considerations that have already been discussed by Caranti: he points out that while Kleingeld is correct to claim that Kant is writing in a tradition, she is wrong to claim that

---

[17] Of course, we disagree primarily on what Kant means in UH by the global conditions of that JS.

Kant agreed with St. Pierre as a result.<sup>18</sup> There were two different traditions from which Kant drew: “(a) the natural law tradition, which includes the ancient *ius gentium*, and (b) preceding peace projects crafted and forcefully proposed by authors with whom Kant was certainly familiar.” But the second tradition, which is surely that to which Kleingeld was appealing, is not a clear case of simply carrying on the same view from thinker to thinker. There are few interesting and connected reasons to think this: First of all, Kant was living during the French Revolution, and there were very few democracies at the time. “France was perceived by the European autocracies not ‘merely’ as a dangerous model for domestic rebels but also as a new power ready to export its principles through some sort of republican crusade” (Caranti 2017, 108). Kant himself was living under a monarch when he wrote the *MM* and *TPP*, and thus his writing of them constitutes a break from what many in his country must have felt about the rabble in France. Thus, Kant’s view of the just IGO, while certainly influenced by St. Pierre (one can hardly deny the seventh thesis’s mention of St. Pierre by name), was more likely the *culmination and evolution* of a peace project going on in the 1700s. (Caranti 2017, 111-12)

According to Caranti, Kant argues against *ius gentium*, putting his view at odds with St. Pierre’s (thus simultaneously also denying the Rousseauvian claim that this federation would be created and enforced through the use of force). (2017, 110-13)<sup>19</sup> “[...] [*Ius gentium* was mainly an attempt to identify [...] the conditions that make war between states just, thus taking for granted that some wars were legitimate. [...] [Kant] simply declares all wars as illegitimate” (Caranti 2017, 110-11).<sup>20</sup> The evidence for Caranti’s claim comes from *TPP*:

[...] [F]or Hugo Grotius, Pufendorf, Vattel, and the like (only sorry comforters) – although their code, couched philosophically or diplomatically, has not the slightest *lawful* force and cannot even have such force (since states as such are not subject to a common external constraint) – are always duly cited in *justification* of an offensive war [...]. (*TPP*, 8:355)

---

18] While Caranti never actually writes that Kleingeld is wrong, his assertions show that she could not possibly be right from this historical context. Caranti admits that “Kant’s ‘philosophical project’ does not arise from a void” (Caranti 2017, 108). But, as the reader will see, this already does a lot to undermine Kleingeld’s position.

19] I share some views in common with Caranti, but we do differ in some respects. In this case, I do think that Kant was really more interested in talking about peace than what justifies war, and except for rare cases of defensive wars where they are arguably legitimate, wars *are* generally illegitimate on my reading of Kant. Caranti’s claim that they are *always* illegitimate seems a bit strong to me.

20] Of course, it is a matter of debate whether Kant really thought *all* wars to be illegitimate. He seems to defend defensive wars, for example. “As for *active violations* which give a *right to go to war*, these include *acts of retaliation* [...], a state’s taking it upon itself to obtain satisfaction for an offense committed against its people by the people of another state, instead of seeking compensation (by peaceful methods) from the other state” (*MM*, 6:346). This undermines Caranti’s claim in my view. But, in agreement with Caranti, it does seem to be the case that Kant was aiming for the conditions that might sustain a peace, rather than the conditions that would make it ok to go to war.

Kant uses this kind of argument against both the justification to go to war and the justification for rebellion. Hence this represents a departure from one fundamental aspect of St. Pierre's *coercive* IGO: namely the idea that it can *justly* use war to enforce that countries act 'rightly' with each other. Even if, therefore, we grant that Kleingeld is right that Kant was not aware of 'Jugement sur la Paix Perpétuelle,' it seems odd to have Kant espouse a view that directly contradicts this position. Kleingeld, of course, ultimately agrees that Kant's 'Bund' was to be non-coercive by *MM* and *TPP*. It is her claim that Kant changed his mind that is at issue. But I just want to point out that Caranti is not, here, merely saying that she should have noticed this from his latter works. Rather, his claim is stronger: from a historical point of view Kant likely was trying to juggle many different aspects of a longer peace project than merely one book by L'Abbé de St. Pierre, and perhaps Kant's works should be interpreted to reflect that juggling act.

In parallel with this, Caranti notes that the political conversation, by the time of *TPP*, had developed some complexity; this naturally leads Kant to a more nuanced and evolved response to the peace project. Published only months before *TPP*, an anonymously published pamphlet, entitled *Épître du Vieux Cosmopolite Syrach à la Convention Nationale de France*, debated the same issues that Kant had debated in *TPP*: 1) there is no right of interference between states, in line with the 5<sup>th</sup> preliminary article; 2) the world republic would not be created by force like its individual nation counterparts; and 3) the 'world federation' is attainable and thus a legitimate political goal. What is important about this, Caranti explains, is that it shows that *TPP* "appears as the answer [to these questions] by a world famous philosopher to the hottest political issues debated in the European context." Kant was struggling to balance a few things: 1) the need for valuing the "autonomy and sovereignty of each people" as against Enlightenment principles; 2) how hard it is to implement perpetual peace as against the moral duty to strive for it; 3) admitting that violence can promote positive change as against the notion that there is no right to rebellion; 4) the practicality of provisional rules as against the necessity of permanent rules; and finally 5) that while we have a duty to strive for perpetual peace, and while it is perhaps guaranteed anyway (as in UH), it is not also the case that there is no need for that moral obligation as a result (Caranti 2017, 113-14). It seems that, on Caranti's reading, Kleingeld has not offered us the whole picture of that to which Kant was responding, and perhaps the historical leg of her argument is broken.

#### IV. KLEINGELD'S SEEMINGLY QUESTION-BEGGING STRATEGY

In any case, the historical claims, along with what Kleingeld argues about the word 'Bund', seem to me to beg the question. Perhaps it would be fairer for me to say that Kleingeld depends on her main argument concerning unsocial sociability to be cogent. Certainly, if she is right about what Kant says in UH, and if Kant's hope was to be in agreement with Saint-Pierre, then it would stand to reason that Kant might have either ignored or not read Rousseau's counter proposals in 'Jugement sur la Paix perpétuelle.' But she ought not

to claim the reverse, as she seems to: that Kant's having ignored Rousseau's latter work, along with his agreeing with St. Pierre, further supports her claims that Kant meant for the federation to be coercive. Furthermore, the ambiguity in the term '*Bund*' does not support her claims unless she has succeeded already at convincing us that Kant meant for the federation to be coercive due to the unsocial sociability of states. Nevertheless, Kleingeld presents the ambiguity in the term as a kind of support for her historical point about Kant's changing his view on what a 'federation' is supposed to be. Since we really cannot do much more here until we tackle what she says there about unsocial sociability, I will deal with that now.

Kleingeld does use the text effectively to argue for her position, but it seems to me that textual support from elsewhere (notably, in UH itself) re-contextualizes what Kant says in discussing the seventh thesis (and there is, thus, no need to build in the assumption that Kant has changed his mind to explain what he says there). Consider, for example, the very next, the eighth, thesis (and some of the discussion of this thesis):

*One can regard the history of the human species [...] as the realization of a hidden plan of nature to bring about an internally, and for this purpose, also an externally perfect national constitution, as the sole state in which all of humanity's natural capacities can be developed. [...] One sees that philosophy also has its chiliastic vision, but one whose occurrence can be promoted by its idea [...], though only from afar, and it is thus anything but fanciful. (UH, 8:27)*

The reference to a "chiliastic" view in the second sentence only supports the claim that Kant thought of the JS itself as a mere ideal; something that we cannot ultimately achieve, but for which we nonetheless have a duty to strive (*MM*, 6:350). Thus, though, we initially will attempt to go through a step similar to Saint-Pierre's coercive federal government, it is in service of getting toward an externally perfect national constitution (teleologically, we may try to form coercive federations, but ultimately, we will want to make a voluntary congress to form that 'externally perfect constitution').<sup>21</sup> We can therefore allow that the seventh thesis speaks of a universal *coercive* federation, without at the same time making the ideal of the JS (the chiliastic vision) synonymous with this coercive federation. On my reading, the eighth thesis refers to the *voluntary* federation to which Kant alludes in the latter works.

We should not confuse teleological claims about what is true about us with what ought to be true. Some of the teleological claims will coincide with what we ought to do, but some will not. For example, we have unsocial aspects. We are violent when we are in too close a proximity. *Teleologically*, we can find a use for this unsociability in terms of how it brings out our other natural capacities best, in the context of our also having sociability. But we should not add that we *ought* to be unsocial. Otherwise Kant would be guilty, in UH, of the naturalistic fallacy. Rather, we simply have a teleology that ultimately does help us to

---

<sup>21</sup> In a separate paper, I plan to make a positive argument for this claim. My presenting it here is meant *merely* to show that the text would seem to contextually support this simpler resolution to the dilemma Kleingeld is trying to resolve (simpler, that is, than the solution she espouses).

bring about what ought to be. Similar, then, to unsocial sociability (and the inherent violence that goes along with it when humans are forced in close proximity in the state of nature), the creation of a Saint-Pierre style federation can be part of what we will do, in service what we ought to do later: i.e., to create that voluntary congress. Thus, Saint Pierre's vision is not clearly what Kant has in mind as part of what we *ought* to bring about (but only, at most, what Kant thinks is part of our *teleology* on the way to making the world a moral whole).

## V. CONCLUSION

Since making sense of the teleological arguments in UH cannot stop at the seventh thesis, when Kant has nine theses there, one must endeavor to understand the others. The eighth thesis is located only a couple of pages after the relevant passages quoted by Kleingeld and actually does a lot to explain the disparity between the explanation of the seventh thesis and what is said in the *MM* and in *TPP*. An "externally perfect civil constitution" serving all of humanity seems most straightforwardly to describe the JS considered globally (and not, say, the Kingdom of Ends).<sup>22</sup> It seems to me that this global

---

22] It is noteworthy that another way to read the eighth thesis is that it refers to the Kingdom of Ends, and that the seventh thesis, and all of Kant's discussion there about St. Pierre, really does refer to the JS. In fact, the easiest way to make Kleingeld's picture fit with *Universal History* is to read the eighth thesis this way. Kleingeld seems to make this move when she writes: "The final end of history according to the *Idea* seems to be identical to the 'moral world' discussed in the *Critique of Pure Reason* under the name of the highest good. In the first *Critique*, this is the ideal [...] in which all agents act morally, and in which, as a consequence of their virtuous action, all are happy (A 809/B 837)" (2009, 185-86). The purpose of *this* paper, however, is merely to show that my way of reading the text keeps Kant consistent between his works on the topic of the JS considered globally.

I will just briefly say here that the Kingdom of Ends seems out of place in UH, since the work is about our teleology, and the Kingdom of Ends, *in no work* of Kant's, is talked about in terms of teleology. Rather, the Kingdom of Ends is talked about as a hope that we must have as a result of our being moral. The question of the final ends of our existence are asked in the context of determinism; namely the questions in UH are asked in terms of cause and effect and what we can expect will happen as a result of what has happened. "Whatever concept one may form of *freedom of the will* in a metaphysical context, its *appearances*, human actions, like all other natural events, are certainly determined [*bestimmt*] in conformity with universal natural laws. History – which concerns itself with providing a narrative of these appearances, regardless of how deeply hidden their cause may be – allows us to hope that if we examine *the play of the human will's freedom in the large*, we can discover its course to conform to rules [*regelmässig*] as well as to hope that what strikes us as complicated and unpredictable in the single individual may in the history of the entire species be discovered to be the steady progress and slow development of its original capacities" (UH, 8:17). Kant's position is hopeful, in that when we look at UH and at the moral works, we see agreement between that which we ought to do and that which eventually will happen. But there are two points that need to be addressed: Firstly, the Kingdom of Ends is not something we ought to bring about, but rather a *hoped for* consequence of our doing what we ought to do (determined, so to speak, by the categorical imperative). For example, in the "Canon of Pure Reason" in the first *Critique*, Kant distinguishes between that which we ought to do (which he says is not appropriate to discuss in there in the first *Critique*) and that for which we can hope (*CPR*, A 805/B 833). Second, the Kingdom of Ends is noumenal and nothing phenomenal can *cause* anything noumenal, for the noumenal is defined as being separate from the laws of cause and effect (*CPR*, A 252/B 308). Thus to include the Kingdom of Ends in UH is, in a strict sense, unreasonable, from a teleological point of view.

view of the JS must be a voluntary congress of states (even in UH), and this seems further supported by the following two interesting bits from *TPP*: (1) aside from the rights of people in a given state, and aside from the rights of states in their voluntary congress, there is also the concern of the right of people to have a kind of world citizenship in “a universal nation of men (*ius cosmopoliticum*).” Kant writes that this latter right is “necessary in relationship to the idea [...] of perpetual peace” (*TPP*, 8:349). Furthermore, (2) in Kant’s discussion of the second definitive article for perpetual peace – “The right of nations shall be based on a federation of free states” (*TPP*, 8:354) –, he writes concerning this federation:

This league [of peace] does not seek any power of the sort possessed by nations, but only the maintenance and security of each nation’s own freedom, as well as that of the other nations leagued with it, without their having thereby to subject themselves to civil laws and their constraints (as men in the state of nature must do) (*TPP*, 8:356).

Thus, it seems to me to be a bit myopic to view the seventh thesis as referring to the JS and to assume that Kant means for the eighth thesis to describe some other universal ideal (e.g., the Kingdom of Ends). We need not add the assumption that Kant has fundamentally changed his mind as to what constitutes the JS at the global level (unless we think he did this *in the course of only two pages!*).

jkatz9@tulane.edu

#### REFERENCES

- Babst, D. 1964. Elective Governments – A Force for Peace. *Wisconsin Sociologist* 3: 9-14.  
 ———. 1972. A Force for Peace. *Industrial Research* 14/2: 55-58.
- Caranti, L. 2017. *Kant’s Political Legacy: Human Rights, Peace, Progress*. Cardiff: University of Wales Press.
- Doyle, M. W. 1983a, 1983b. Kant, Liberal Legacies, and Foreign Affairs. Parts I and II, *Philosophy and Public Affairs* 12/3 and 12/4: 205-35, 323-53.
- Kant, Immanuel. 1983. Idea for a Universal History with a Cosmopolitan Intent. In Ted Humphrey (trans. and ed.), *Perpetual Peace and Other Essays: on Politic, History, and Morals*, 29-40. Indianapolis, IN: Hackett Publishing Company. [UH]
- . 1983. On the Proverb: That may be true in Theory, but is of No Practical Use. In Ted Humphrey (translated and edited), *Perpetual Peace and Other Essays: on Politic, History, and Morals*, 61-92. Indianapolis, IN: Hackett Publishing Company. [TT]
- . 1996. *Metaphysics of Morals*. Translated and edited by Mary J. Gregor. New York, NY: Cambridge University Press. [MM]
- . 1996. *Toward Perpetual Peace*. In Mary J. Gregor (edited and translated), *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy*, 311-51. New York, NY: Cambridge University Press. [TPP]
- . 2003. *Critique of Pure Reason*. Translated and edited by Norman Kemp Smith. New York: Plagrave Macmillan. [CPR]
- Kleingeld, Pauline. 2009. Kant’s Changing Cosmopolitanism. In Amélie Oksenberg Rorty and James Schmidt (eds.), *Kant’s ‘Idea for a Universal History with a Cosmopolitan Aim’: A critical guide*, 171-86. New York: Cambridge University Press.

- . 2012. *Kant and Cosmopolitanism: The Philosophical Ideal of World Citizenship*. New York: Cambridge University Press.
- Rousseau, Jean-Jacques. 1761. Extrait du Projet de Paix Perpétuelle de Monsieur l'Abbé de Saint Pierre. In Bernard Gagnebin and Marcel Raymond (eds.), *Oeuvres Complètes* vol. III, 563-89. Paris: Gallimard, 1964.
- . 1782. Jugement sur la Paix Perpétuelle. In Bernard Gagnebin and Marcel Raymond (eds.), *Oeuvres Complètes* vol. III, 591-600. Paris: Gallimard, 1964.
- Saint-Pierre, Charles-Irénéé Castel, Abbé de. 1713. *Projet pour rendre la Paix Perpétuelle en Europe*. Utrecht: Schouten.

# Mill's Political Perception of Liberty: Idiosyncratic, Perfectionist<sup>1</sup> but essentially Liberal

Leonidas Makris  
Aristotle University of Thessaloniki

**Abstract:** There is a dominant perception of liberty among most contemporary liberals. It is one close to empiricism's portrayal of freedom as a natural right of every person to advance her interests. According to this view, there are no demanding conditions under which people can be regarded as free agents but their unfettered behaviour from external inhibitions. It is widely thought that Mill's liberalism does not deviate considerably from this tradition. The present text suggests a different reading of the gist of Mill's political rationale. Highlighting the positive – in Berlin's sense – aspects of his politically demanding concept of liberty which successfully combines facets of different traditions serves the purpose to stress the alternative features of his approach. Underlining the active political intervention Mill is instructing by implicating the state to preserve the essence of freedom and induce it in society, manifests how his position diverges from that of mainstream neutralist liberals.

**Key words:** liberty, autonomy, individuality, perfectionism, Mill John Stuart.

Mill is often portrayed just as an unambiguous defender of the freedom to engage in any human behaviour not perceived harmful to others. It will be claimed that such a view of his moral and political philosophy is an unfair and incomplete one. His concept of liberty is overall consistent with his ideal of human flourishing and of a liberal society; a society prospering only when it promotes civic patterns and values transcending the picture of an individual concentrated exclusively in self-interested activity. This conclusion is not though as straightforward as it might sound. Mill's conception of freedom carries strong influences from different and, to a certain extent, antagonistic traditions. They range from empiricism's view of freedom as a natural right of every individual securing her self-interest to ancient Greek and romantic groundings of liberty in the capacity of humans to act in accordance to rational moral law. Contrary to many interpretations of Millian liberty as prioritizing empiricism's emphasis on the freedom to pursue one's private passions<sup>2</sup> I believe that the latter pole of thought had a more profound impact on his understanding of freedom. The German Romantics as well as a direct recovery of Greek ideas of self-development are active influences shaping Mill's ethical thought. "[His] liberal idea was

---

1] There is an ample range of interpretations attributed to 'perfectionism', e.g. Wall (2008), Haksar (1979), Finnis (1987), Gray (2000b), Hurka (1993), Rawls (1973). The way I use the term here follows more that of Wall. It does not specify the precise content of activities qualifying as perfectionist nor does it dictate political authorities to maximise them. Promoting ideals of human flourishing does not equal promoting excellence. Perfectionism here is compatible with the harm principle in advancing autonomy but in the way I see Mill as interpreting the notion of 'harm' i.e. promoting both negative and positive duties in order for people not to be harmed. It resembles also Hurka's perfectionism in the sense that it pursues as a worthwhile political aim (valuable) autonomous agency.

2] See Berlin (1969), Rees (1985), Ryan (1991), Fiss (2003), Elshtain (2003), etc.

[more] a romantic-hellenic idea of free self-development in every aspect of one's human power." (Skorupski 1999, 224-25)

Notwithstanding the above, Mill challenged the view that ancient classical and modern as well as German romantic and its alleged counterpart, British empiricist views of liberty are at irreconcilable odds. While Mill values empiricism's protection of individual rights, he also shares with the romantics the need for a more substantive view of freedom, one of a competent individual using freedom well. By embracing in his concept of liberty self-legislation and self-determination, he refutes justice as a mere adherence to rules assigning to it an ethical outlook. By encouraging the cultivation of human perfection and general moral development Mill attempts to challenge the view that there is an established antithesis between self-interest and universal obligations in the decision-making process. Challenged forcefully also by Krause (2002) this dichotomy is easily traceable in current political theory and philosophy. It consists, on the one hand, of empiricism's self-interests as expressed by some of its advocates<sup>3</sup> and, on the other hand, of moral idealism's autonomy as expressed for the most part by Kant and Rousseau. Freedom identified with self-interested activities, without a higher or lower content or a direct relation to the public good, contrasts autonomy as a higher form of freedom focusing on acting for the universal good and not for ourselves. Mill opposes this dichotomy by maintaining that forms of self-interested conduct can also be higher forms of freedom and contributors to the general good.

The current prevalent strand of liberalism, dominated by a neutralist approach towards the good, can certainly benefit from such an interpretation of Mill. The influence of such a classical and fervent supporter of freedom on the course of liberal thinking can be powerful since the interpretation of the core concept offered here is distinctive. Particularly when in opposition to what most contemporary liberals suggest, what I see as the gist of his perception of liberty, justifies an interventionist state expected to take active measures in order to be able to essentially defend it. The fact that I will be largely disregarding the utilitarian component in Mill's theory of freedom is not only related to the huge attention his utilitarianism has already attracted. It is also an outcome of the perfectionist analysis that I here suggest as cogent in order to grasp the gist of Millian liberalism. According to this analysis, Mill's thought is permeated by a particular type of perfectionism. Thus, even his concept of utility has a hierarchical and qualitative aspect

---

3] Devigne (2006) classifies as important empiricists promoting self-interested activities, among others, Hobbes, Locke and Hume. Indeed, Locke follows Hobbes in linking the idea of the good with human desires, pleasure and pain with good and evil (Locke 1975). However, Locke and Hobbes differ considerably since Locke believed in stringent divine limits in the self-interested action (Patten 2006). Thus, Devigne mistakenly classifies him as chiefly instigating self-interest. For Locke God created man and we are God's property (Uzgalis 2007). It follows for example that for him man 'has no liberty to destroy himself' or commit suicide (Locke 1966) restricting accordingly his self-directed activities.

(Utilitarianism CW x, 1985) which lies on a perfectionist basis and undermines the fundamental principles of utilitarianism.<sup>4</sup>

### I. USING DIFFERENT FREEDOMS

Leaning towards positive freedom as a self-directed moral existence yet not opposed to the classic liberal tradition as originally conceived, Mill uses indeed the term liberty in various ways. As he acknowledges when young he was more sympathetic to the empiricist thesis accepting the link between ‘unprocessed’ necessity and freedom. Somehow apologizing for such earlier ideas, in his *Autobiography* (CW) Mill criticizes the lack of a self-conscious state of mind in empiricism’s rationale where there is no difference between being “conscious of a feeling” (or desire) and “merely having the feeling.” (CW xxxi 1989, 138)

It is not only during his younger age that Mill uses liberty in a different way. A very thoughtful typology of the various uses that we can encounter throughout his work is offered. “Negative freedom”, “rational self-direction”, “autarchy” and “autonomy” (Gray 1996, 74). This last notion though transmits better the kernel of Mill’s liberalism and as such deserves more attention here. Mill should not be associated with libertarianism and a negative conception of autonomy that misrepresents his notion of self-development. It is unfair to Mill’s liberalism to use a primarily negative right of autonomy, which, when coming into play, describes solely Millian individuals with powers already partly developed. Such an approach is faithful neither to Mill nor to current accounts of autonomy. It attempts to distance Millian political philosophy from a more collectivist, social democratic version of liberalism, leading it towards its libertarian pole (Donner 1991). Despite getting right Mill’s overall evaluation, his ‘socially embedded’ concept of liberty and his cardinal commitment to help all people to lead meaningful lives, a two-folded comment should complement Donner’s position regarding Millian freedoms. Firstly, it is a fact that throughout his work Mill uses the term liberty in different ways. In *Logic* and *Hamilton’s Philosophy* for instance Mill argues that in essence only some individuals do in fact self-amend. But overall Mill also confirms Donner’s egalitarian description postulating that all individuals should have options to self-amend their character (Devigne 2006, 70). In any case, we should stress that it is not justified to argue that Mill was a restrictive negative libertarian. Even Gray, who is accused for doing that (Donner 1991), actually sees in Mill’s work a “positive state action to benefit” people and “a large range of desirable state

---

4] Bentham’s utilitarianism (1996) entitles humans to pursue as an end pleasure and the biggest amount of happiness they can get. This is regardless of how each individual indulges in pleasure since this can be very subjective. His utilitarianism evaluates actions simply based upon their consequences and in particular the overall happiness created for everyone affected by the action. Contrary to this, Mill’s utilitarianism (CW x 1985, 212-14) is dominated by the verdict of ‘competent judges’ who are better equipped to decide the best among pleasures and modes of existence. The ‘judges’ are people who manage to experience both higher and lower pleasures, and if they have to choose, they opt for the former. This is yet another strong evidence of Mill’s perfectionism, an aspect which deserves our attention and dominates not only his liberalism but also his concept of utility.

activities having nothing to do with harm prevention”, adding that the Millian principle of liberty is not at all violated by such interventions providing they are not authoritative (Gray 1996, 61-63). Secondly, it is true that despite being blamed for identifying always autonomy with negative freedom, Gray clearly distinguishes between the two of them (1996, 74, 77). He also finds in *On Liberty* “unmistakable traces of a Kantian conception of autonomy, absorbed by Mill [...] from Humboldt” (1996, 78), a romantic view that clearly criticized the empiricist negative conception of liberty.

To thinkers like Rousseau and Kant liberty as self-determination is not just the unfettered pursuit of someone's empirical desires. If people are to be *really* free they must be autonomous managing and regulating their lives in a mode presuming the distinction between the environment and self (Rousseau 1987; Kant 1996). This is exactly the autonomy attributed to Mill when he says that, on top of exercising rational capacities in objective choice-conditions, an autonomous agent should be to some extent disentangled from the conventions of his social environment and from other people's influence. Such an ideal of personal autonomy is among Mill's cardinal commitments. Millian autonomy should by no means be identified with a passive, negative, libertarian and individualistic tradition. Mill argues for liberty not because its protection reassures a society of free men; Mill seeks to *promote* a society of autonomous people whose actions express principles, fruits of a process of critical reflection. More openly than it is with 'autarchy', Millian 'autonomous agency' is something to be achieved and should not be regarded as a natural endowment or inheritance. (Gray 1996)

Gray's typology with the different nuances of freedom is therefore apt to describe not only the strictly negative 'self-regarding area' but also the positive notion of 'self-development' supported strongly by a robust concept of liberty as autonomy, both encountered in Mill's work. They are respectively described by Mill himself: “[Negative freedom as] a sphere of action in which society [...] has if any, only an indirect interest; comprehending all that portion of a person's life and conduct which affects [mainly] himself”. Simultaneously, Mill relates the positive perception of freedom with people's interconnectedness. Since the conduct of any society member affects others, the positive encouragement of her development can potentially prevent harm to others. “No person is an entirely isolated being; it is impossible for a person to do anything seriously or permanently hurtful to himself, without mischief reaching at least to his near connections, and often far beyond them” (CW xviii 1977, 225, 280). Mill indeed uses the term liberty for two complementary, but distinct, conceptions of freedom. The first concentrates on limiting the individual's external coercion by the state and society and the second on cultivating developed human beings capable of forming their own decisions and desires. The co-existence of different concepts of freedom does not constitute an inability from Mill's part to define and distinguish adequately self-regarding from other-regarding spheres or the concept of 'effects' from 'interests'. Niggling about such differences seems to be more related to the inappropriateness of an account to accommodate Millian autonomy expressed as perfectionist individuality, supported positively as self-mastery

and linked with a particular type of human flourishing (On Liberty CW, chap.3). Rees (1985) generally subscribes to the view of Mill as a leading exponent of the negative idea of liberty as plainly the absence of restraint. Due to the absence in his work of a methodologically cogent typology of liberty, in reaching conclusions of the utmost importance for the essence of freedom, Mill indeed sometimes uses indiscriminately the term liberty to convey its normative (positive) as well as its more neutral (negative) meaning. This surely explains the agreement between many<sup>5</sup> that his ‘one very simple principle’ of liberty is anything but simple.

## II. PREVAILING AUTONOMY

Despite the ramifications of a process needed to elicit the terminology<sup>6</sup> and the evaluation of the different Millian freedoms, an overall conclusion about the moral and political core of Mill’s celebrated principle can be quite effectively deduced. Notwithstanding the ambiguity that a distinction between negative and positive understanding of liberty posits to interpret Mill’s spirit as establishing a negative thesis concerning freedom is mistaken (Berger 1984). Berlin commits this mistake when he interprets Mill as primarily focusing on a limited area of personal freedom which should by no means be violated (1969). As Devigne (2006) observes, nearly all contemporary commentators portraying Mill as the archetypical theorist of negative liberty do it by disregarding his concern for wisdom.<sup>7</sup> Construing chiefly Mill’s theory as ruling out strictly interference with the freedom of others and forbidding social control, just because Mill said that society may interfere with individual conduct only “to prevent harm to others” (CW xviii 1977, 223), oversimplifies his notion of ‘harm’ as well as his account of freedom overall. Thus seen, the essence of Mill’s work is directed at establishing a negative thesis of freedom. Berger is right to find this misleading since it greatly underestimates the most distinctive features of his work. His liberalism, as Berger puts it, is clearly a powerful, innovative and positive doctrine. “This is the doctrine of the importance to human well-being of individual self-development, or, as I prefer to call it, autonomy.” (1984, 229)

Berger’s view affirms that such a notion of autonomy – interconnected well with his concept of happiness of the competent judges – is in accord with Mill’s perfectionist notion of self-development. It is such a concept of autonomy that can express better the essential spirit of his freedom and his liberalism by combining – as mentioned above – two seemingly different traditions. On the one hand, Mill stresses intellectual development as the core of rational and critical reflective skills indispensable to achieve autonomy, he emphasizes the importance of liberty of choice and of self-determination and he combines individuality and authenticity. On the other hand, he articulates clearly his view of the

---

5] For example, Gray (1996) and Rees (1991).

6] Mill never uses for example the term autonomy to describe his freedom.

7] Devigne (2006) regards as notable exceptions Thompson (1976) and Berkowitz (1998).

ideal person – one who has achieved balance and harmony between moral, intellectual and affective development – of freedom and individuality with sociality, attachment and caring for others. He enunciates a view of human flourishing postulating no inconsistency in the need to combine in a happy human life these sides of self-development. Such an enriched ideal concept of autonomy combining self-mastery and exertion of social solidarity approximates the essence of Mill's account of liberty and resonates in Donner's explication of it. This is because the latter is aware of the overlapping between such treatments of autonomy and parts of Mill's self-development i.e. what Donner sees as his quintessence of liberty. (Donner 1991)

Such a concept of autonomy as self-development expresses better Mill's gist of liberty and his view of social feelings; in addition, it is also in accordance with his notion of happiness. Mill sees liberty as a prerequisite of happiness for specific reasons. For him human development – a prerequisite of elevated happiness – is feasible only when people are free. An objectively sound ideal is necessary to achieve genuine happiness with altruistic life being such an ideal. While it cannot be imposed as a moral obligation – a condition for the ideal of altruism is its spontaneity – when people embrace it voluntarily it becomes a great source of self-realising happiness. For Mill self-development is linked with ideals of living and forms part of an overall argument connecting happiness and freedom; thereby it has a prominent place in his theory (Skorupski 2006). Mill's liberty does not rest on "the idea of abstract right, as a thing independent of utility"; a utility based on human flourishing depicted in the ideal decisions of competent judges and conceived in order to promote human development, that is, "in the largest sense, grounded on the permanent interests of man as a progressive being" (CW xviii 1977, 224). Such 'permanent interests' in Mill stem from people's potential for free self-development incorporated deeply in the core of his ethical and political outlook; an essence of human good as something dynamic, developmental and individual (Skorupski 2006). In order to achieve specific human potentialities that Mill clearly favoured, what idealists of the nineteenth century called 'self-realisation', or what he calls 'moral freedom' (see *Logic* and *Hamilton's Philosophy*, CW), he supported good social institutions enabling the flourishing of free self-development. The perfect compatibility between Mill's liberal ideal of self-culture and his greatest happiness principle is evident. Only free self-culture combined with rules protecting society can lead to full self-development, and solely by completing the self-development of people's potential we can obtain high forms of happiness. Once again, we can see that Mill is in favour of self-realisation or autonomy expressed in Aristotelian fashion but possibly via a rationale of romanticism. The teachings of Coleridge, Kant and others had an impact on Mill's initiative to emphasize the capacity for individual self-mastery and the exertion of wilfulness. "A person feels morally free who feels that his habits or his temptations are not his masters, but he theirs: who even in yielding to them knows that he could resist" (CW viii 1974, 841). Mill refers to an advanced quality of rational will as 'moral freedom' and like Kant, identifies it with reliable virtue. This is evident in the following: "[W]e must feel that our wish, if not strong enough to alter

our character, is strong enough to conquer our character when the two are brought into conflict in any particular case of conduct. And hence it is said with truth, that none but a person of confirmed virtue is completely free.” (CW viii 1974, 841)

Mill evidently distinguishes between better and worse ways of life and he associates this differentiation closely with freedom. He ranks freedom through self-mastery higher than servility to custom or to pressing physical needs. His freedom resonates the stated Aristotelian view for the priorities of human soul, with reason guiding the mortal clay's passions toward virtue (Aristotle 1985). Unlike Aristotle, though, Mill does not identify in detail the particular life a self-directed individual should lead or the exact choice-worthy goods and virtues he should favour, widening thus the range of life-styles within which any person may hope to attain his excellence. It is what Gray calls Mill's affinity for pluralism (1996) which merges smoothly with the culmination of his freedom, that is, the ability to desire “for its own sake, the conformity of [one's] own character to [a] standard of excellence, without hope of good or fear of evil from other source than [one's] own inward consciousness” (CW x 1985, 95). It is clear Mill promotes an ideal of a certain type of individual as the capable one to attain ‘complete freedom’ (Logic, CW). This stems also from his discussion in *Utilitarianism* (CW x 1985, chap. 2) where he favours a developed mind forming a type of character that evolves into a good in itself. And this preference can be attributed, among other things, to “love of liberty and personal independence” (CW x 1985, 212). Moral freedom implicates the reassurance of the opportunity for the development of character based on the cultivation of mental faculties and a level of self-consciousness that permits someone to reflect upon his own state of mind. It is a kind of character re-evaluation and self-amendment. “[The] feeling of our being able to modify our own character if we wish, is itself the feeling of moral freedom which we are conscious of” (CW viii 1974, 841). Again, we can observe Mill's emphasis on romantic aspects of individual and liberty. This is because he opposes “the supposed [empirical] alternative of admitting human actions to be necessary” and inevitable, i.e. a result of an excessively deterministic process, “inconsistent with [...] instinctive consciousness, as well as humiliating to the pride and even degrading to the moral nature of man.” (CW viii 1974, 836)

### III. INDIVIDUALITY AS AUTONOMY

It is not only the concept of moral freedom or liberty that occasionally conveys the message of what I called autonomy but also Mill's notion of individuality: “It is desirable [...] that in things that do not primarily concern others, individuality should assert itself. Where [...] customs of other[s] [...] are the rule of conduct, there is wanting one of the principal ingredients of human happiness and quite the chief ingredient of individual and social progress” (CW xviii 1977, 261). Crisp too draws a parallel between Mill's individuality and a notion of autonomy defined in a perfectionist way. Apart from seeing Millian individuality as a minimum requirement to run our own life and not merely rely on social custom, his elucidation approximates the present one. “We might call this

autonomy, though that term is not found in Mill" (Crisp 1997, 196). Based on a simple analysis of the word's etymology, Crisp attributes to Mill a notion not merely envisaged as a capacity adding to one's welfare but as exertion of that capacity in self-government. Combining the indispensable role of rationality, the value of intellectual development for good self-government and components of individuality – all of foremost importance for Mill's 'true liberty' – Crisp fuses these elements in his notion of autonomy. While involving spontaneity, Crisp's autonomy is not just that. As a constituent of individuality and so of welfare, autonomy necessarily implicates the development of people's own potentialities. Pointing to the ideal-regarding aspect of Millian freedom he draws the parallel between reflective arrival at true belief and the exercise of autonomy as consisting in the cultivation and use of intellectual capacities. (Crisp 1997, 196)

Crisp's assertion finds abundant support in Mill's work. Following Humboldt (1993), Mill ascertains that "individuality of power and development", the "end of man", has two prerequisites, "freedom and variety of situations". Individuality of development and freedom are fused in autonomy because "the human faculties of perception, judgement, discriminative feeling, mental activity, and even moral preference, are exercised only in making a choice". And the constituents of individuality – 'freedom and variety' – through their union give rise to "individual vigour and manifold diversity, which combine themselves in 'originality'". It is of a pre-eminent significance for Mill to stress that "the faculties are called into no exercise by doing a thing merely because others do it, no more than by believing a thing because others believe it". "He who does anything because it is the custom, makes no choice" while the one who follows his "own reason", "his own feelings and character" is the one "who employs all his faculties" and therefore "chooses his plan for himself" (CW xviii 1977, 261-62). Subsequently, individuality expressed with originality portrayed in one's own strong feelings, impulses and will – filtered with their appropriate cultivation – is outspokenly linked with a particular ideal of character and grounded in Mill's view of human nature. "To say that one person's desires and feelings are stronger and more various than those of another, is merely to say that he has more of the raw material of human nature and is therefore capable [...] of more good". Construed like this, as plentiful of "the sternest self-control", individuality is delineated as the source of "love of virtue" and "energetic character" (CW xviii 1977, 263-64).

A person whose desires and impulses are his own – are the expression of his own nature, as it has been developed and modified by his own culture – is said to have a character. One whose desires and impulses are not his own, has no character, no more than a steam-engine has a character. If, in addition to being his own, his impulses are strong, and are under the government of a strong will, he has an energetic character. (CW xviii 1977, 264)

Summing up the arguments unfolded here we could claim that vital for the understanding of Mill's liberalism is to recognize that loss of freedom is not identified with coercion by others. Lack of self-development of character also entails loss of liberty.

As we can see human perfection for Mill consists not solely of the application of rationality and of an active attitude towards life. He makes it conspicuous that it also

demands the elevation of the will. He notices the positive role impulse can play to render the individual capable of gaining self-command. Consistent with his analysis of the ancient Greek spirit, Mill links powerful desires with strong wills and postulates that a stronger will facilitates the path to an autonomous and ingenious existence. “There is no natural connexion between strong impulses and a weak conscience. [...] Desires and impulses are as much a part of a perfect human being as beliefs and restraints”, providing the former ones are “properly balanced” (CW xviii 1977, 263). Hence, Mill contends that when guided internally by will, justice, and reason, desire “contributes to human perfection”; if a society neglects the role of strong desires it impedes progress and it undermines the general good (Devigne 2006, 167). The authenticity therefore of a developed individuality, which among other things presupposes a will forged around strong desires, evokes the picture of human perfection which, in order to be complete, includes promoting the public good.

The gradual unravelling of the Millian autonomy seems to be disclosing a very rich and detailed vision about human flourishing. Genuine individualism is decisively supported by reason, will, strong desires, dignity and duty to oneself; only when such individuality is approximated real social and political progress becomes attainable. The thorough and unfeigned conception of freedom, linked with the ability to overcome barriers like a dominant public opinion and personal impediments like unbridled desires, is in Mill closely tied to self-development. If one wants to pursue an active self-development and determination – in turn linked with Mill’s view of human nature and excellence – he “must use observation to see, reasoning and judgement to foresee, activity to gather materials for decision, discrimination to decide, and when he has decided firmness and self-control to hold to his deliberate decision”. He definitely requires all these qualities and their exercise to be employed precisely by “his own judgment and feelings” (CW xviii 1977, 263). Without them, the whole merit of human existence is challenged: “What will be his comparative worth as a human being? It really is of importance, not only what men do, but also what manner of men they are who do it”. And in a direct link between self-development and the underlying basis of forming an admirable human essence, he adds that the task of self-development is to exemplify this kind of man. “Among the works of man, which human life is rightly employed in perfecting and beautifying, the first in importance is surely man himself”. Yet seeing development as multifaceted, Mill stresses that human nature is not a machine to be programmed according to a detailed prescription. It should be treated like “a tree, which requires to grow and develop itself on all sides, according to the tendency of the inward forces which make it a living thing.” (CW xviii 1977, 263)

Mill derives his concept of individuality from an explanation of human well-being which takes account of our developmental and ‘progressive’ nature. Human nature, well-being and individuality are interwoven in a two-fold argument. With complete development of potential, we can reach highest forms of well-being. Also, his liberal ideal of full personal development aims at people’s wholeness by stressing both education of feeling and education of reason and will. While Mill’s individuality aspires to touch philosophic truth as such, it does so via an innovatively synthesized

perspective combining different elements. His individuality weaves together romantic ideas of authenticity, revealing the unrepeatable and ingenious parts of a person, and the classical “perfectionist emphasis on development” aiming at advancing the higher powers of human nature (Muirhead 2004, 116). Mill does indeed combine several elements of different perspectives in his enriched concept of individuality discerned as autonomy. In addition to the Humboldtian and Kantian perception of autonomy – “the ground of the dignity of human nature and of every rational nature” (Kant 1993, 41) – Mill uses for the development of his individual ingredients from various analyses. His affinity to cultivate reason and promote strong will and desires comes from the ancient Greek tradition; his picturing of human perfection combining creativity and concern for the public good used religious and aesthetic culture as instruments of inspiration; his idea of human excellence builds on modernity’s fidelity to universal authority while praising Humboldt’s view that variety of situations is a requisite to individuality (Devigne 2006). The extended array of different strands of thought that have exerted an influence on Mill’s multifaceted concept of liberty could provoke, justifiably at a first glance, an objection about the coherence of such a notion. In addition, if it is to be interpreted as such, a growing scepticism could arise about its compatibility with liberalism as it is commonly perceived. Reconciling Mill’s views on reason and the will and identify how they contribute to the individual’s freedom is not an easy task.

#### IV. THE CORE OF MILL’S FREEDOM IS CONSISTENTLY LIBERAL

It is true that Mill faces a challenge when he attempts to reconcile forms of individuality implicating higher thoughtfulness and the habitual pursuit of desires. By attempting to integrate in his thought the reformed platonic dialectic, Coleridge’s synthetic dialectic and the morality of German Romantics, he formulates a conception of liberty that combines empiricism’s causality and the romantic conception of free will aiming at overcoming the common oscillations in political theory and philosophy. Oscillations between allegedly antagonistic conceptions of liberty, empiricist versus romantic, ancient versus modern. Galston’s interpretation of Mill suggests that liberal tradition has space for a conception of an intrinsic individual excellence intertwining freedom with these diversified components: influenced by romanticism, Mill devises a liberal conception of individual excellence as the full flowering of individuality; it innovatively combines the classical Greek impetus to develop human powers through activity with the modern realization of the idiosyncrasy of each individual blending such powers (Galston 1988). Given though Mill’s stance to introduce an idea of human excellence, the question if it can be effectively merged with freedom and variety of situations as requisites for human individuality remains. Does Mill’s morality of freedom fail to meet the challenge to combine exalted individuality and habitual desires? Can the habitual and the conscious comprise at the same time key features of the free individual? How can the personified expression and strenuous identity and will of his liberal self be

reconciled with his ideal of self-development? To reply to such queries we should recall that while Mill's liberalism values self-mastery and advances its preconditions, the latter do not presuppose only objective rational and emotional skills conceived independently of our personal inclinations to want certain things. Using our faculties efficiently involves partly the resolute pursuing of the desires we desire. Conscious volition is juxtaposed beside cultivated reason and fortitude which through continuing practices make the free individual approximate "the case of the person of confirmed virtue." (CW x 1985, 238)

The rapprochement of seemingly antithetical components in Mill's autonomy is attained by his commitment to human liberation implying an ideal of the person which suggests a conception of the good. There is a profound connection between good life as autonomous and life performed necessarily by a particular ideal type of human being, the autonomous agent. On the one hand, this ideal excludes heteronomous existence dominated by unrestrained emotions or by a tradition which people accept without any challenge. On the other, as Mill postulates the free approach to customs fosters a certain type of person. "A different type of human excellence [...] a conception of humanity as having its nature bestowed on it for other purposes than merely to be abnegated"; more of a "Pericles" than "John Knox" or "Alcibiades" (CW xviii 1977, 265-66). Pericles was indeed Mill's greatest hero of antiquity (Bain 1882). Mill compares these historical figures to illustrate his preferred notion of human perfection; a fusion of certain qualities leading to 'the highest possible good'. Aiming at that he consistently puts forward Pericles as an exemplar of human excellence (Devigne 2006). Only a personality of such calibre, and no other with possible concessions in his individual skills, seems to have made it to comply with the demanding requirements of Mill inspiring his ideal of a strong autonomy; an ideal dominating his liberal apprehension of the good. Mill's formulation of a distinctively liberal conception of the good associates the best polity with that which secures that good, the flourishing of the individual conceived as "strongly autonomous". "On this reading, the liberalism of [...] Mill is thus 'perfectionist' [even] in Rawls's sense; [...] it effectively asserts and enforces a particular conception of the good life." (Crowder 2002, 36)

It is evident that the Millian ideal of autonomy thus conceived could be included in a species of narrow perfectionism<sup>8</sup> as Hurka (1993) defines it, that is, with strong and exclusive foundations in human flourishing. Despite being one of the focal points of Mill's political message and the one that characterizes the distinctive nature of his liberalism, as mentioned above, it cohabits with different, of minor importance for this matter, exegeses of liberty. Whilst his ideals of autonomy and happiness convey the perfectionist weight of Mill's morality, their coexistence with freedom defined – following Gray's typology – as a negative concept<sup>9</sup> confuse Hurka to the extent that he attributes to the latter an

---

8] The terms 'broad' and 'narrow' perfectionism here signify respectively accounts where autonomous life is presented either in combination with other basic principles (e.g. utility and rights) or as the cardinal one against which the claims of a morality are weighted. I believe Mill's liberalism is described better by the first type of perfectionism. Hurka too uses the above criterion to distinguish 'narrow' from 'broad' perfectionism.

9] Negative freedom is of secondary importance behind Mill's autonomy as exemplified in his

absolute weight. Thus, while I agree with Hurka that Millian autonomy as an intrinsic good can cohabit with another intrinsic good like utility – and in that sense being of broad nature – they are both in turn linked with a particular perception that Mill has about human flourishing. To achieve that Mill does not attribute absolute weight to *any* free choice negatively defined, as Hurka implies, but to autonomous choices expressing his perfectionist aim for individual development. Yet, Mill retains a certain commitment to a negatively defined liberty, meaning that he fosters restrictions on what others can do to the individual by the exercise of their wills.

Thus far, negative liberty is an essential condition for the individual's freedom but freedom as autonomy is not realised merely because one of its conditions has been met. Mill's central aim remains forming individuals capable of exercising choices skilfully, and autonomously. That is exactly the autonomy which Hurka incorporates in his Aristotelian perfectionism and calls "deliberated autonomy". The formerly mentioned Millian test of value as expressed by 'competent judges' presumes some negotiation of liberty. But this 'transfer' of liberty is permissible only to a lower 'negative level', through consulting non-coercive means "because there is much non-coercive promotion of the good that perfectionism approves" which is consistent with the liberal ideal (Hurka 1993, 151, 159). And this negotiation of liberty can only take place to the extent that it contributes to the formation of an autonomous character. Mill clearly supports such a developed character-individual which can result only by 'directing' liberty to such an ideal result. His liberal ideal could "never gain widespread acceptance until most develop the type of personal character requisite to its implementation" (Riley 1998, 157). Hence, his notion of negative liberty independently from his autonomy – proving here the aptness of Gray's (1996) terminology – is clearly not absolute but only of an instrumental role in a wider plan that leads to a perfectionist understanding of liberalism. For Mill a free and potentially happy individual is expected to express her good and competent nature and developed character. Therefore, he establishes a link between liberty as autonomy and perfection.

## V. THE ROLE OF THE STATE

While Hurka is decisive in defending autonomy from a perfectionist standpoint and imputing to Mill a similar defence of it, he is ambiguous about Mill's freedom negatively perceived. This specific flaw in Hurka's superficial analysis of Mill's freedom, overly interpreting it as an absolute principle, is demonstrated in the ambiguity of his view on the Millian state. Firstly, he suggests that Mill never wanted the state to interfere with citizens' lives. Then, confirming Mill's perfectionism, he verifies that "neutrality is not a traditional liberal ideal, for it is rejected by Mill: He thinks a person's choosing badly, although no reason to coerce her, does justify 'remonstrating' and 'reasoning' with her". And Hurka

---

ideas of individuality and self-development. Baum – like I do here – calls the latter Mill's 'freedom as autonomy' (Terchek 2002).

(1993) uses this Millian argument precisely to stress why the state should be using these means actively to support the liberal ideal and why therefore his perfectionism is against state neutrality. Indeed, for Mill, self-development and genuine liberty have certain specific requisites – mental material, institutional – so meaningful development cannot take place under just any conditions. Favourable conditions do exist for Mill (Valls 1999) and under them the human potentiality for autonomous agency must be developed. (Baum 1998)

As it is the case with his happiness, he is in favour of the state and society being actively involved in promoting his ideal of individuality and autonomy. He actually does not see the need why a good state should be a power independent from a society where individual interest and autonomy can flourish. “What was now wanted was that the rulers should be identified with the people, that the interest and will was the interest and will of the nation. There was no fear of its tyrannizing over itself” (CW xviii 1977, 218). A good government should be representing every citizen and hence no one should be afraid of its influence and policies. Mill’s view is affected by Coleridge’s stance that there is a need for the institutions to help create a national culture which can morally help to develop the citizenry (Coleridge 1983). Mill appears sceptical towards the incentives of many in England who insist in supporting state neutrality claiming that in an opposite case its action would be inimical to the public and private interest. “In England [...] there is a considerable jealousy of direct interference by the legislative or executive power with private conduct, not so much from any just regard for the independence of the individual as from the still subsisting habit of looking on the government as representing an opposite interest to the public” (CW xviii 1977, 222-23). Mill is critical of the obsessive focus to restrain the government’s ability to confine liberty of action because it disregards whether the agent’s desires and motives are his own or not. Although he often argues forcefully against the state’s direct and intrusive interference in private affairs, Mill also maintains that there is enough space for society – within which a functional state operates in accordance to its directives – to mould the “goodness and wisdom” of its individuals. “If society lets any considerable number of its members grow up mere children, incapable of being acted on by rational consideration of distant motives, society has itself to blame for the consequences”. As it can be inferred, the state should assume an active role in trying to prevent such an event not only by participating in the shaping of “all the powers of education” but in influencing positively with its policies “the ascendancy which the authority of a received opinion always exercises over the minds who are least fitted to judge for themselves.” (CW xviii 1977, 282)

The qualities required for full self-development and autonomy are “self-regarding virtues” as well as “social” ones. “It is equally the business of education to cultivate both” (CW xviii 1977, 277). Mill endorses an activist state which contributes to the material and institutional prerequisites for self-development. There is a moral obligation in Mill’s society to help each other cultivate self-regarding virtues. Failing to comply with such duty legitimizes society to censure people or raise taxes guaranteeing state education in self-regarding duties. All this is compatible with Mill’s liberty principle and whether

society actually employs such methods is a question of efficient policy, not a matter touching on liberty. "So Mill is not an ethical neutralist about the state" (Skorupski 2006, 49-50). The enforcement of universal education, the aid to help educate the poor, the duty of the state to supervise the educational system (On Liberty CW, chap. 5), are not the only means the state should use to promote the best conditions for an autonomous existence. This is because education should not be perceived as strictly related only to traditional teaching; rather "knowledge and culture, which have no obvious tendency to better the fortunes of the possessor, but solely to enlarge and exalt his moral and intellectual nature, shall be [...] obtruded upon the public" (CW vi 1982, 259). Opposing the libertarian wing of liberalism Mill also promotes a legally enforceable taxation for purely redistributive purposes. In addition, he closely relates taxation with an underlying concept about what is *good* for people and how they can acquire more knowledge about it. "It is hence the duty of the state to consider, in the imposition of taxes, what commodities the consumers can best spare" and select "those of which it deems the use [...] to be positively injurious". Thus, the state should "indirectly discourage conduct which it deems contrary to the best interests of the agent." (CW xviii 1977, 297-98)

In a rare direct intervention of a moralistic sense Mill is even willing to relinquish to the state the power to "forbid marriage unless the parties can show that they have the means of supporting a family" (CW xviii 1977, 304). Mill usually *suggests* solutions for similar issues and does not resort to morally objectionable imposing measures like this one. The latter though is yet another indication of the significance he attributes to the ideal conditions for mental advancement and consequently for self-development. Mental cultivation is such a laudable goal for Mill that can even entail restricting some individual liberties to ensure a good level of education linked with the well-being of the families. Generating conditions for high forms of individuality is for Mill as significant as establishing equal rights for all. Against the 'free-marketeers' of the time Mill is also in favour of legislative interventions ameliorating the context within which individual choices are made. The legislation to restrict the working week is an example (Skorupski 2006). The imposed limits to free trade and the rules enforced on employers by increasing the amount of public control to prevent fraud or to ensure sanitary conditions and protect the workforce are other examples. "Such questions involve considerations of liberty, only in so far leaving people to themselves is always better, *caeteris paribus*, than controlling them: but that they may be legitimately controlled for these ends, is in principle undeniable" (CW xviii 1977, 293). Promoting self-realisation, self-mastery and self-development, ingredients of Millian autonomy, is a task with which the state should comply. Actively seeking to improve people, the most important feature of good government is "the virtue and intelligence of the human beings composing the community" (CW xix 1977, 390), something which is certainly not an infringement of legitimate liberty. As Skorupski puts it, "a liberal state can legitimately promote conceptions of the good" and "it is not a principle of Millian liberalism that the state should be ethically or aesthetically neutral" or that it "should not have a conception of the good among its core allegiance-inspiring values". A society with a duty to educate its

members about better ways of living should employ the state too for its objective: It breaks no Millian principle to do that through all public institutions and activity funded by a democratic vote of the citizens. (Skorupski 2006, 103-104)

## VI. CONCLUSION

Most of the above-mentioned arguments as well as the perfectionist grounds on which Mill is defended here clearly separate him from many contemporary liberals.<sup>10</sup> Most of them, following the dominant current of today's liberalism, think that the state should not promote any conception of the good. Mill consistently focuses on the problem of reconciling wisdom and liberty under his concepts of individual exertion and development as autonomy; to accomplish this he employs the state as an additional help for people's moral education. He criticizes a state sterile and neutral towards its citizens' mental expansion, questioning the value of an administration of justice perfecting its operating machinery while ignoring the task of moral education (Devigne 2006). The government should actively seek to aid and stimulate people's exertion and development. "The worth of the State", Mill asserts, "is the worth of the individuals composing it; and a State which postpones the interests of their mental expansion and elevation [...], a State which dwarfs its men [...] will find that with small men no great thing can really be accomplished" (CW xviii 1977, 310). A good polity with a functional government should aim at promoting the health of the individual's character, leading it to flourish in both the public and private domains. Mill contends that whether people become or not autonomous is contingent upon factors – educational, political, economic and psychological – which can advance their capacities for autonomy. And the Millian state plays an active role in ameliorating all the autonomy-generating conditions. Hence, Mill's conception of freedom as autonomy, presupposing the implication of means and the availability of opportunities for self-development and self-government, refutes the ostensibly oppositional relationship between freedom and power; this is because the state's active intervention in favour of this freedom-autonomy often personifies power. The misfortune is that the same negative liberal tradition that assumes this permanent antithesis mistakenly perceives an active Millian state as inimical to freedom. (Baum 1998)

Recapitulating the role of Mill's state in contributing to the active promotion of *liberty as autonomy* we could claim that it stems from the same perfectionist basis inspiring the conception of the notion itself. Hurka's propounded model for the liberal perfectionist state verifies this. Mill's state complies with all the criteria which the Hurkian perfectionist state puts forward. This is because its intervention comes into play in order to defend a particular concept of liberty which comprises the gist of Mill's political thinking. Respecting citizens' autonomy by promoting non-coercively the good, the state favours education not only in its strict sense but also as universal mental

---

[10] E.g. Rawls, Ackerman, Larmore, Nagel, R. Dworkin, etc.

cultivation, provided through taxation and subsidization. While human propensity to follow the good materializes under favourable conditions, people also have other desires which presuppose help to resist or to accomplish. As Mill proves and as Hurka (1993) concludes, it is therefore fitting that politically we can favour liberty but reject the ideal of state neutrality.

*lmakris2@yahoo.com*

#### REFERENCES

- Aristotle, 1985. *Nicomachean Ethics*. Translated by T. Irwin. Indianapolis: Hackett.
- Bain, Alexander. 1882. *John Stuart Mill: A Criticism with Personal Recollections*. London: Longmans.
- Baum, Bruce. 1998. J. S. Mill on Freedom and Power. *Polity* 31(2): 187-216.
- Bentham, Jeremy. 1996. *An Introduction to the Principles of Morals and Legislation*. Edited by J. H. Burns and H. L. A. Hart. Oxford: Clarendon Press.
- Berger, Fred. 1984. *Happiness, Justice, and Freedom: The Moral and Political Philosophy of John Stuart Mill*. Berkeley and London: University of California Press.
- Berkowitz, Peter. 1998. *Virtue and the Making of Modern Liberalism*. Princeton: Princeton University Press.
- Berlin, Isaiah. 1969. Two concepts of liberty. In *Four Essays on Liberty*. Oxford: Oxford University Press.
- Coleridge, Samuel. 1983. *On the Constitution of Church and State in The Collected Works of Samuel Taylor Coleridge*, edited by K. Coburn. Princeton: Princeton University Press.
- Crisp, Roger. 1997. *Routledge Philosophy Guidebook to Mill on Utilitarianism*. London: Routledge.
- Crowder, George. 2002. *Liberalism and Value Pluralism*. London and New York: Continuum.
- Devigne, Robert. 2006. *Reforming Liberalism: J. S. Mill's Use of Ancient, Religious, Liberal, and Romantic Moralities*. New Haven and London: Yale University Press.
- Donner, Wendy. 1991. *The Liberal Self: John Stuart Mill's Moral and Political Philosophy*. Ithaca and London: Cornell University Press.
- Elshtain, Jean Bethke. 2003. Liberty and the Problem of Authority. In *On Liberty-John Stuart Mill*, edited by David Bromwich and George Kateb, 208-33. New Haven: Yale University Press.
- Finnis, John. 1987. The Legal Enforcement of 'Duties to Oneself': Kant vs Neo-Kantians. *Columbia Law Review*, 87(3): 433-56.
- Fiss, Owen. 2003. A Freedom both Personal and Political. In *On Liberty-John Stuart Mill*, edited by David Bromwich and George Kateb, 179-96. New Haven: Yale University Press.
- Galston, William. 1988. Liberal Virtues. *The American Political Science Review* 82 (4): 1277-1290.
- Gray, John. 1996. *Mill on Liberty: A Defence*. London and New York: Routledge.
- . 2000b. Mill's Liberalism and Liberalism's Posterity. *The Journal of Ethics*, 4(1-2): 137-65.
- Haksar, Vinit. 1979. *Equality, Liberty, and Perfectionism*. Oxford: Oxford University Press.
- Humboldt, von Wilhelm. 1993 [1854]. *The Limits of State Action. (The Sphere and Duties of Government)*, edited by John Barrow. Indianapolis: Liberty Press.
- Hurka, Thomas. 1993. *Perfectionism*. New York: Oxford University Press.
- Kant, Immanuel. 1993. *Grounding for the Metaphysics of Morals*. Indianapolis: Hackett Publishing Company.

- . 1996. *Practical Philosophy*. Edited by Mary Gregor. Cambridge: Cambridge University Press.
- Krause, Sharon. 2002. *Liberalism with Honor*. Cambridge, MA: Harvard University Press.
- Locke, John. 1966. *Two Treatises of Government*. Edited by Peter Laslett. Cambridge: Cambridge University Press.
- . 1975. *An Essay Concerning Human Understanding*. Edited by Peter Nidditch. Oxford: Oxford University Press.
- Mill, John Stuart. 1963-1991. *The Collected Works of J. S. Mill*, 33 vols. [abbreviated as 'CW']. Edited by John Robson. Toronto and London: University of Toronto Press and Routledge & Kegan Paul.
- . 1974. Volume VII: *A System of Logic Ratiocinative and Inductive*, [abbreviated as 'Logic'], Part I.
- . 1974. Volume VIII: *A System of Logic Ratiocinative and Inductive*, [abbreviated as 'Logic'], Part II.
- . 1977. Volume XIX: *Essays on Politics and Society*, Part II (Considerations on Representative Government, 371-753) [abbreviated as 'Repr. Govern.'].  
 ———. 1977. Volume XVIII: *Essays on Politics and Society*, Part I (On Liberty, 213-310).
- . 1979. Volume IX: *An Examination of Sir William Hamilton's Philosophy*.
- . 1982. Volume VI: *Essays on England, Ireland, and the Empire* (Notes On the Newspapers, 149-280).
- . 1985. Volume X: *Essays on Ethics, Religion, and Society* (Mill's Utilitarianism by D. P. Dryer, lxiii-cxiv; Remarks on Bentham's Philosophy, 3-18; Bentham, 75-116; Coleridge, 117-64; Whewell on Moral Philosophy, 165-202; Utilitarianism, 203-60).
- . 1998. Volume XXXI: *Miscellaneous Writings* (Jeremy Bentham and James Mill, 1-256).
- Muirhead, Russell. 2004. *Just Work*. Cambridge, MA: Harvard University Press.
- Patten, Alan. 2006. Review of R. Harrison's *Hobbes, Locke, and Confusion's Masterpiece: An Examination of Seventeenth-Century Political Philosophy*. *Philosophical Books* 47 (4): 352-55.
- Rawls, John. 1973. *A Theory of Justice*. Oxford: Oxford University Press.
- Rees, John. 1985. *John Stuart Mill's 'On Liberty'*. Edited by Geraint Williams. Oxford: Clarendon Press.
- . 1991. A Re-Reading of Mill On Liberty. In *J. S. Mill On Liberty In Focus*, edited by John Gray and G.W. Smith, 169-89. London and New York: Routledge.
- Riley, Jonathan. 1998. *Routledge Philosophy Guidebook to Mill on Liberty*. London and New York: Routledge.
- Rousseau, Jean-Jacques. 1987. Discourse on the Origins of Inequality. In *The Basic Political Writings*, edited by D. A. Cress, 45-81. Indianapolis: Hackett Publishing Company.
- Ryan, Alan. 1991. John Stuart Mill's Art of Living. In *J. S. Mill On Liberty in Focus*, edited by John Gray and G. W. Smith. London: Routledge.
- Skorupski, John. 1999. The Ethical Content of Liberal Law. In *Ethical Explorations*. Oxford: Oxford University Press.
- . 2006. *Why Read Mill Today?* London and New York: Routledge.
- Terchek, Ronald. 2002. Review of B. Baum's *Rereading Power and Freedom in J. S. Mill*. *The American Political Science Review*, 96(2): 397.
- Thompson, Dennis. 1976. *John Stuart Mill and Representative Government*. Princeton: Princeton University Press.

- Uzgalis, William. 2007. John Locke. In *The Stanford Encyclopedia of Philosophy (Summer Edition)*, edited by Edward N. Zalta, <http://plato.stanford.edu/archives/sum2007/entries/locke/> (accessed 24 May, 2018).
- Valls, Andrew. 1999. Self-Development and the Liberal State: The Cases of John Stuart Mill and Wilhelm von Humboldt. *The Review of Politics*, 61(2): 251-74.
- Wall, Steven. 2008. Perfectionism in Moral and Political Philosophy. In *The Stanford Encyclopedia of Philosophy (Fall Edition)*, edited by Edward N. Zalta, <http://plato.stanford.edu/archives/fall2008/entries/perfectionism-moral/>. (accessed 24 June, 2018).

# Rawlsian Stability and the Hazards of Envy

Alexandros Manolatos

National and Kapodistrian University of Athens

**Abstract:** This paper explores the role of envy in the third part of *A Theory of Justice* and challenges a wide-spread game-theoretic view of stability. The proponents of this view see Rawls's account of stability as an attempt to solve a collective action problem. I claim that Rawls treats the development of envious feelings as a distinct source of instability which is not part of a collective action problem and has to be addressed separately. My thesis entails that we shouldn't read the congruence between the right and the good as the culmination of Rawls's overall argument for stability. This reading is supported by the revised account of stability in *Justice as Fairness: A Restatement* and leads to a better understanding of Rawls's political turn.

**Key words:** Rawls, stability, envy, game-theory, Weithman, congruence.

The problem of stability is according to Rawls fundamental to political philosophy. A political conception of justice must be utopian and at the same time realistic. It must contain principles and ideals that can be seen as possible, given the circumstances of justice and the laws and tendencies of the social world.

The central role of stability in Rawls's political theory is reflected in the structure of *A Theory of Justice*.<sup>1</sup> The third part of the book is devoted to an examination of the questions of stability and congruence. As Rawls (1971, 55) writes in *TJ*, however attractive a conception of justice might be, it is defective if it fails to engender in human beings a desire to act upon it. So the main aim of the third part of Rawls's magnum opus is to show that the principles selected in the first part are going to generate their own support and are more stable than other principles. This is shown by the assessment of the relative strength of the sense of justice they cultivate and of the opposite temptations to act unjustly.

The significant role of stability in Rawls's political philosophy is reflected also by his turn to *Political Liberalism*.<sup>2</sup> In the introduction to *PL* he informs the readers that he decided to revise his theory because the account of stability he presented in *TJ* was not realistic enough. These revisions were necessary because "the problem of stability is fundamental to political philosophy and an inconsistency there is bound to require basic readjustments" (Rawls 2005, xvii).

Surprisingly, even after the publication of *PL*, Rawls's account of stability remained initially at the margins of the overall academic discussion on his work. Freeman (2007) has remarked accurately that from all the voluminous commentary on Rawls's work, very little of significance has been written on his argument for stability in the third part of *TJ*.<sup>3</sup>

---

1] All references to *A Theory of Justice* are to the original edition, unless otherwise stated. The book is cited hereafter as *TJ*.

2] All references to *Political Liberalism* are to the expanded edition. The book is cited hereafter as *PL*.

3] Freeman made that remark almost ten years after the publication of *PL*. He was one of the first who stressed the importance of stability in Rawls's work.

One possible explanation is that most critics saw Rawls's political turn as a response to the critique of communitarianism.<sup>4</sup>

The first significant exclusive commentary on Rawls's argument for stability was Paul Weithman's book *Why Political Liberalism? On John Rawls's Political Turn*. Weithman offers a thorough and novel analysis of the third part of *TJ*, which reveals the role of ideals in Rawls's argument for stability. The publication of his book revamped the debate on Rawls's political turn and sparked an interest in the role of stability in his political philosophy. At the same time, Weithman's book established a game-theoretic view of Rawls's account of stability in *TJ*.

In this article my main goal is to oppose this game-theoretic view and to stress the importance of the special psychologies in Rawls's overall stability argument.<sup>5</sup> I will claim that the game-theoretic view distorts how Rawls conceives the problem of stability in *TJ*, since it focuses solely on the congruence argument and the challenges of mutual assurance and isolation. I will first outline the game-theoretic view of Rawlsian stability and then I will propose a different reading that gives equal weight to the problem of envy. This reading reveals the inherent connection of stability with the issue of distributive justice and social status. Envy is a serious challenge for the justification of justice as fairness because the inequalities sanctioned by the difference principle are theoretically unlimited and they could arouse hostile feelings that threaten social unity. A large part of Rawls's account of stability aims to show that this challenge can be met by the reciprocal and egalitarian character of the two principles. If we see the problem of stability only through the lens of game theory we lose this important aspect of Rawls's view.

In the final section of the article, I claim that if we abandon the game-theoretic view we can have a better understanding of Rawls's political turn.<sup>6</sup>

## I. THE GAME THEORETIC VIEW OF RAWLSIAN STABILITY

In *TJ* the problem of stability is examined in the third part of the book, which comprises of three chapters. In the first chapter Rawls presents his theory of the good and the Aristotelian principle. In the second chapter he discusses how the citizens of a well-ordered society acquire a sense of justice and he compares the stability of the two principles of justice with other conceptions. The final chapter examines whether the sense

---

4] Freeman notes that there is a widespread perception that the revisions Rawls has made to *TJ* leading up to *PL* have come largely in response to communitarian criticisms (2007, 175).

5] In *TJ* Rawls presents envy and spite as special psychologies or special attitudes that limit rational behavior. The will for domination and submission and a peculiar aversion to risk and uncertainty are the other special psychologies he cites in *TJ*.

6] In a series of papers between 1980 and 1987 Rawls started modifying his theory and revising some of his arguments. These revisions led to the publication of *PL* and the recasting of justice as fairness as a political and not as a comprehensive conception of justice. In the section "The Revised Account of Stability" I will discuss in more detail the main differences between *TJ* and *PL*.

of justice coheres with our good. The three chapters are interlinked, but the way the overall argument is structured is not so straightforward and has to be reconstructed.

One possible reconstruction is based on the assumption that Rawls conceives the problem of stability in game-theoretic terms.<sup>7</sup> The proponents of this game-theoretic view interpret the problem of stability in *TJ* as a collective action problem that arises from two sources of instability which threaten a well-ordered society.<sup>8</sup> The first source of instability arises from the fact that citizens of a well-ordered society will respect the two principles of justice only if they have the assurance that others will act in the same way. This is called by Rawls “the assurance problem” (1971, 270). The second source of instability arises from the fact that if they know that others comply with their duties, they will be tempted to ignore their duties and act as free-riders.

Those who support this view hold that Rawls’s argument for stability is targeted exclusively at solving a generalized prisoner’s dilemma and the related problem of assurance. They claim that the goal of the third part of *TJ* is to show that a well-ordered society would not suffer from these two sources of instability and that this is achieved by the congruence between the right and the good.<sup>9</sup> The congruence argument is presented in the section “The Good of the Sense of Justice”. In this section Rawls argues that citizens of a well-ordered society would affirm their sense of justice because in that way they could satisfy four desires. The desire to express themselves as free and equal rational beings, take part in forms of social life that call forth their own and others talents, avoid the psychological cost of hypocrisy and have ties of friendship. The advocates of the game-theoretic view see this

---

7] The influence of game theory in *TJ* is not evident only in the argument for stability. Rawls uses a general analogy between society and games which helps him emphasize the idea of fairness (Galisanka 2017). In *TJ* Rawls refers to this analogy when he discusses the rule of law (1971, 235), the idea of pure procedural justice (1971, 85), the idea of social union (1971, 525) and when he compares institutions with games (1971, 55). There is also an input of game theory in the argument for the selection of principles from the original position, where Rawls introduces the maximin rule (1971, 152). I would like to thank an anonymous referee for constructive comments regarding the overall use of game theory by Rawls.

8] This view is expressed by McLennen (1989), Weithman (2010), Thrasher and Vallier (2015), Quong (2014) and Freeman (2003, 2007). Freeman and Weithman have contributed decisively in illuminating Rawls’s distinctive understanding of stability. Freeman (2007) reveals how the congruence argument is connected with the Kantian interpretation of justice as fairness. In his seminal book on Rawls’s political turn, Weithman (2010) highlights the role of ideals of conduct, friendship and association in the justification of the congruence argument. I find these reflections convincing and I do not intend to challenge them. My objection centers on the fact that they both see the congruence argument as a culmination of the overall argument for stability.

9] John Thrasher and Kevin Vallier (2015) contend that in *TJ* Rawls attempts to solve the assurance problem by showing that citizens have reason to endorse their sense of justice as part of their good, namely that the right and the good are congruent. This congruence ensures that “each person in the well-ordered society will be motivated by a sense of justice to comply with the public conception of justice and will know that everyone else is motivated in the same way” and this “common knowledge of compliance” preserves mutual assurance between citizens (Thrasher and Vallier 2015, 937).

section as the culmination of Rawls's reasoning on stability, where all the elements of his theory of the good and his moral psychology are combined and tied together.<sup>10</sup>

Their view has definitely some textual support. Rawls presents the problem of stability in game-theoretic terms in two sections of *TJ*. In the section "Economic Systems" he presents two problems that arise in distributive systems, the problems of isolation and assurance. He compares the problem of isolation with "the general case of the prisoner's dilemma of which Hobbes's state of nature is the classical example", which arises "whenever the outcome of the many individuals decisions made in isolation is worse for everyone than some other course of action, even though, taking the conduct of the others as given, each person's decision is perfectly rational" (Rawls 1971, 269). The assurance problem in contrast is associated with the fact that "each person's willingness to contribute is contingent upon the contribution of the others" (Rawls 1971, 270).<sup>11</sup>

In another section of *TJ* Rawls writes that there are two tendencies leading to instability. The first one, arises from the fact that some citizens may be tempted to avoid fulfilling their duties, if they believe that they will still benefit from the distribution of public goods. The second one, arises from the fact that citizens may stop contributing their share if they believe or suspect that others are not contributing theirs. Rawls notes that this instability is likely to be strong, when the risk of sticking to the rules when others are not, is too high (1971, 336).

There is also textual support that the congruence argument aims to solve the free-rider and mutual assurance problems. In the end of the section "The Good of the Sense of Justice" Rawls claims that "the hazards of the generalized prisoner's dilemma are removed by the match between the right and the good" (1971, 577). The hazards of the generalized prisoner's dilemma refer to the instability that threatens a well-ordered society by the forces that prompt its citizens to skip their duties. As I mentioned above, Rawls identifies these forces as the temptation to adopt a free-riding behavior and the lack of assurance that others will comply with their duties. The match between the right and the good, namely their congruence, restricts the temptation to act as a free-rider because citizens in a well-ordered society would judge that they are better off by being just persons than

---

10] This claim is supported by the fact that at the beginning of the section Rawls writes that "now that all the parts of the theory of justice are before us, the argument for congruence can be completed" (Rawls 1971, 567). Freeman (2003, 277) for example writes that the congruence argument "begins in Part III of *Theory of Justice (TJ)*, is developed for over 200 pages, and culminates (in Section 86) at the end of a very long book".

11] It is important to note that Rawls makes clear that his view on stability is different from that of Hobbes. Rawls aims to ensure a moral stability or a "stability for the right reasons" and not a mere *modus vivendi*. Unlike Hobbes, the aim is not to achieve the obedience to law by the existence of some external mechanism that delivers sanctions. Instead, Rawls purports to show that the two principles of justice can be internalized by the citizens of a well-ordered society and become a part of their good.

they would be if they took advantage of others. And the fact that it is publicly known that everyone is motivated in the same way,<sup>12</sup> solves also the problem of mutual assurance.<sup>13</sup>

## II. THE SPECIAL PSYCHOLOGIES AND THE HAZARDS OF ENVY

The game-theoretic view seems well justified if one takes into account the excerpts mentioned above. Yet there are other parts of this very long book that do not confirm this view and raise doubts as to whether Rawls was focused solely on the free-rider and mutual assurance problems. A large part of the chapter “The Good of Justice” seems to be completely independent of the overall game-theoretic thinking. More specifically, in the three sections where he discusses the problem of envy (80-82), Rawls seems to regard envy as a separate source of instability that threatens a well-ordered society. There are a number of passages that support this claim.<sup>14</sup>

In the section “The Problem of Envy” Rawls presents envy as a collectively disadvantageous propensity which is dangerous and worsens the situation of all the parties involved. He writes that after a conception of justice is selected in the original position, we must check if it is going to arouse envy at such an extent that “the social system becomes unworkable and incompatible with the human good” (Rawls 1971, 531). He then notes that regarding justice as fairness, we have to assess if the inequality allowed by the difference principle is so acute that it generates destructive feelings of envy (Rawls 1971, 532). Consequently, Rawls explains the reasons why envy is so dangerous. First, when we envy other persons we no longer value what we have and this loss arouses hostile feelings. Second, we are willing to deprive them of their benefits even if we have to

---

12] The public knowledge of this fact could mean for example the data published by tax authorities and other public institutions.

13] This solution to the assurance problem is briefly mentioned by Rawls (1971, 336) in another part of *TJ*, where he says that the assurance problem “is to maintain stability by removing temptations of the first kind, and since this is done by public institutions, those of the second kind also disappear, at least in a well-ordered society”.

14] Rawls uses some aspects of game theory to clarify the special psychology of envy. He notes for example that envy can be collectively disadvantageous. “When others are aware of our envy, they may become jealous of their better circumstances and anxious to take precautions against the hostile acts to which our envy makes us prone. So understood envy is collectively disadvantageous: the individual who envies another is prepared to do things that make them both worse off, if only the discrepancy between them is sufficiently reduced”. Rawls (1971, 532). At the same time envy can also prompt us to excel, which can be collectively advantageous. “A somewhat different case is that of emulative envy which leads us to try to achieve what others have. The sight of their greater good moves us to strive in socially beneficial ways for similar things for ourselves” (1971, 533). In *TJ* Rawls examines how the first kind of envy can threaten the stability of justice as fairness. This threat is triggered by psychological and social conditions that are separate from the ones that are related with the problems of mutual assurance and free-riding behavior. The problem with envy is not that we don’t have the assurance that others are doing their share, is that we judge ourselves “happy or unhappy only by comparison with others” (Kant 2009, 27). I would like to thank an anonymous referee for comments regarding the relation between envy and game theory in *TJ*.

give up something ourselves. Third, when others know that we envy them they may take preventive measures against possible hostilities (Rawls 1971, 532).

The problem of envy is mentioned by Rawls also in the chapter “The Original Position”, where he presents stability as one of the criteria that the parties have to take into account when they choose a conception of justice. In this chapter Rawls relates the problem of envy with that of stability, while he doesn’t make any reference to the mutual assurance problem. In particular, he notes that in the last part of *TJ* he will try to show that the selected principles lead to a well-ordered society where envy and other destructive feelings are not likely to be strong and that the conception of justice undermines the circumstances that trigger disruptive attitudes (Rawls 1971, 144).

The above passages from *TJ* indicate that Rawls treated the problem of envy as a distinct destabilizing power which is not part of a collective action problem and a prisoner’s dilemma. They also indicate that Rawls regarded the problem of envy as an important test for the two principles of justice.

One possible reply by the supporters of the game-theoretic view is that the examination of the problem of envy is not distinct from the overall reasoning on the assurance and isolation problems and the congruence between the right and the good. For example, Weithman (2010, 141) believes that the conclusions Rawls draws about the mitigation of envy by the two principles of justice are integrated in the final congruence argument. He thinks that these conclusions provide an important supplement on Rawls’s argument that citizens of a well-ordered society regulated by the two principles of justice would have a desire to avoid the psychological cost of hypocrisy and the desire to have ties of friendship. These two desires, along with the desire to express themselves as free and equal beings and the desire to participate in forms of life that call forth their own and others talents, form the grounds of congruence between the right and the good.<sup>15</sup>

Regarding the first desire, Weithman (2010, 141) points out that it is because of “the argument in 82” that a typical member of a well-ordered society “would regard the costs of hypocrisy and deception as high relative to the benefits of the wealth she could get above her fair share”.<sup>16</sup> What is the “argument in 82” according to Weithman? It is the argument that members of a well-ordered society would not feel envy for those who have more wealth, because their status is not defined by their relative position in the distribution of wealth but by the position of equal citizenship. If they know that others respect them as equal citizens, they would not be moved to seek wealth and economic status as a means to self-respect. So, the fundamental assertion in Rawls’s argument is

---

15] The final congruence argument in *TJ* is that persons who have these four desires would also have a desire to act justly, because by doing so they promote their good.

16] Weithman argues that the idea that members of a well-ordered society would regard as very high the cost of hypocrisy is not obvious and it needs to be explained. For Weithman, when Rawls says that the cost of hypocrisy would be higher in a well-ordered society he does not mean it in absolute terms. Instead, he implies that they would be higher relative to what citizens could get by paying them, for example, the greater wealth they would enjoy by cheating on their taxes.

that “the position of equal citizenship answers to the need we might have thought people had for economic status” (Weithman 2010, 142). The same holds for the second desire, the desire to have ties of friendship. The members of a well-ordered society would judge that the potential costs to those with whom they have ties of friendship give them strong reasons to treat their sense of justice as supremely regulative, since they don’t need more wealth for self-respect.

Freeman (2007, 153) believes that the argument from the absence of envy is just one part of the “peculiar array of arguments in chapter 9 of *Theory*” that shows that the sense of justice is not “in many respects irrational and injurious to our good”.<sup>17</sup> In particular, he maintains that in the final chapter of *TJ*, Rawls argues that the sense of justice is not arbitrary, entirely conventional and self-destructive, is not grounded in a self-debasing submission to authority, it accounts for the good of community and it doesn’t mask a lack of self-worth and a sense of failure and weakness. This array of arguments, in which the absence from envy is included, is presented by Freeman as part of Rawls’s main argument for the good of justice, the congruence argument.

Even if Weithman and Freeman were right<sup>18</sup> and the absence of envy plays a role in the congruence argument, this would not be a proof that Rawls treats the problem of envy as an intrinsic part of a collective action problem. There are numerous passages in *TJ* where Rawls presents envy as a serious threat to the stability of a well-ordered society, a threat which is distinct from the hazards of a generalized prisoner’s dilemma and calls for a distinct solution. The destructive character of envy introduces a difficulty that goes beyond the generalized prisoner’s dilemma. It is not only that we don’t have the assurance that others are doing their share, but it is also the sidelong glance we cast at one another, taking an interest in others relative position in the distribution of wealth, income and social status.<sup>19</sup>

In the sections 80 to 82 Rawls demonstrates why the two principles of justice alleviate this threat, at least more than the principle of utility. The congruence argument on the other hand does not deal with the disruptive effects of envy. The argument that the sense of justice is compatible with a person’s good can be an answer to the problems of mutual assurance and isolation but is irrelevant to the problem of special psychologies.

---

17] Freeman responds to Brian Barry’s claim (1995) that the acquirement of a sense of justice is sufficient to prove stability within Rawls’s framework.

18] I believe that the absence of envy is one of the reasons that citizens see their political society as a good, but I don’t think that it functions as a “supplement” argument in the way presented by Weithman. My claim that we have to treat the absence of envy as a distinct argument for the stability of a well-ordered society doesn’t mean that it doesn’t have a supportive role in the congruence argument. The mitigation of envious feelings contributes to the good of the citizens of a well-ordered society. But this shouldn’t overshadow the fact that Rawls treats envy as a distinct source of instability which is inextricably linked with distributive justice.

19] This point is stressed by William A. Edmundson (2017, 105) who claims that envy and the special psychologies introduce a new kind of difficulty.

### III. THE ARGUMENT FROM THE ABSENCE OF ENVY

I have argued that the game-theoretic view underestimates the importance of envy in Rawls's account of stability by not recognizing it as a distinct source of instability. I have also argued that the proponents of this view treat the congruence argument as the culmination of Rawls's overall argument for stability.

My thesis is that we can have a better understanding of Rawls's reasoning on stability if we see the congruence argument as only one part of a larger argument. Another important part is the argument from the absence of envy. So, on my interpretation, the overall stability argument is developed in three parts. In chapter 8, Rawls presents a moral psychology designed to show how people in a well-ordered society can acquire a sense of justice. In chapter 9, Rawls argues that a) members of a well-ordered society would not be moved by feelings of envy and b) that their sense of justice would be congruent with their good. The argument from the absence of envy aims to show how the two principles of justice mitigate the destabilizing feelings of envy. The congruence argument purports to remove the hazards of the generalized prisoner's dilemma.<sup>20</sup>

I will try to present a detailed reconstruction of the argument from the absence of envy and its connections with the bases of self-respect. I will emphasize how it supports the justification of the two principles of justice not only in comparison with the utilitarian principle but also with a more egalitarian conception. Consequently, I will look at the survival of this argument in *Justice as Fairness: A Restatement*.<sup>21</sup>

In *TJ* Rawls argues that self-respect is a primary good and "perhaps the most important" one (Rawls 1971, 396). This emphasis reveals the importance of self-respect in the choice of the two principles of justice from the original position. It is equally important though for the relative stability of justice as fairness as a conception of justice. In section 29 Rawls defines self-respect as the sense of one's own worth. It is the sense that the plan of life we have chosen is worth carrying out. He gives a more nuanced definition in the chapter "Goodness as Rationality". There he claims that self-respect or "self-esteem" has two aspects. The first aspect, is how we sense our own value, the conviction that our life plan is worth carrying out. The second one, is the confidence we have in our ability to execute this plan (Rawls 1971, 440).

---

20] In chapter 9 Rawls (1971, 513) discusses "various desiderata of a well-ordered society and the ways in which its just arrangements contribute to the good of its members". He argues that justice as fairness demonstrates the objectivity of judgments of justice, something that can assure us that our sense of justice is not arbitrary and conventional. He shows also that justice as fairness is not grounded in a self-debasing submission to authority, that it can be combined with the good of community and that it enables citizens to express their nature as free and equal human beings. These are also considerations in favor of the stability of justice as fairness. The main difference is that the argument from the absence of envy and the congruence argument correspond to distinct sources of instability.

21] Cited hereafter as *JFR*.

This is the reason that makes self-respect a primary good and perhaps the most important one. Without it “nothing may seem worth doing, or if some things have value for us, we lack the will to strive for them [...] and we sink into apathy and cynicism” (Rawls 1971, 440).

Rawls stresses that there are two circumstances that support the first aspect of self-esteem, the sense of our own worth. The first one is having a rational plan of life that satisfies the Aristotelian principle and the second one is finding that others appreciate who we are and what we do (Rawls 1971, 440). So there is a reciprocal feature on how we secure our self-respect. Our self-esteem is dependent on how we think our fellow citizens value us. Rawls’s strategy is to show that the citizens of a well-ordered society regulated by the two principles of justice would be immune to destabilizing feelings such as envy because they would have a strong sense of self-respect.<sup>22</sup> In section 81, Rawls claims that the conception of justice as fairness supports the self-respect of citizens, because they all have the same basic rights and are treated as equals. This is based on the assumption that in a well-ordered society liberties and rights are more important for the self-esteem of citizens than their income share.<sup>23</sup> Rawls (1971, 545) gives two reasons why liberties and rights should have priority for securing the primary good of self-respect. First, if their self-respect was anchored in their wealth and income, then the pursuit of status and self-esteem would be a zero sum game, where the improvement of one person’s position would lower that of someone else.<sup>24</sup> Second, it would be irrational to accept equal division of wealth to secure equal status and self-esteem, since there are compelling grounds for allowing an unequal distribution of wealth. Thus, it is more preferable to support the primary good of self-respect by giving priority to the basic liberties and defining the same status for all.

Rawls offers two additional points on why justice as fairness would secure self-respect. The first is that in a well-ordered society the less favored know that the greatest advantages others enjoy work also for their benefit. The difference principle permits deviations from strict equality only if this would make the less advantaged worst off. This reciprocal feature of the difference principle makes it easier for them to accept the disparities between themselves and others.<sup>25</sup> The second point is that justice as fairness does not associate the distribution of income and wealth with moral virtue and excellence. So the less fortunate

---

22] Rawls makes a straightforward connection between envy and self-respect by assuming that the main psychological root of the liability to envy is a lack of confidence in the worth of our plans of life and our ability to execute them. By contrast, someone with a robust self-esteem has no desire to level down the advantages of others.

23] “The basis for self-esteem in a just society is not then one’s income share but the publicly affirmed distribution of fundamental rights and liberties. And this distribution being equal, everyone has a similar and secure status when they meet to conduct the common affairs of the wider society” (Rawls 1971, 544).

24] This is another point where Rawls uses some aspects of game theory in his overall discussion of the problem of envy.

25] It is a principle which expresses the ideas of fraternity and reciprocity, since citizens feel that the others care for their good and they don’t try to take advantage of their bad luck.

have no reason to view themselves as inferior. In addition, Rawls (1971, 537) notes that given the background institutions of a well-ordered society, inequalities are not extreme and that the various social unions which flourish in a well-ordered society reduce the painful visibility of these inequalities, because citizens tend to compare their situation with others that hold positions similar to their aspirations. These features make it less possible for the less advantaged to experience humiliation.

For all the above reasons, Rawls concludes that justice as fairness is not challenged by outbreaks of envy and that it seems relatively stable. In the end of section 81 he states that so long “as the pattern of special psychologies elicited by society either supports its arrangements or can be reasonably accommodated by them, there is no need to reconsider the choice of a conception of justice” (Rawls 1971, 541). This could lead to the false conclusion that the argument from the absence of envy is applied by Rawls only to show that justice as fairness is a stable and feasible conception that would not be destabilized to a troublesome extent by destructive feelings of envy. There is no doubt that one of the main aims of Rawls is to demonstrate how justice as fairness removes the hazards of envy, in the same way that the congruence argument removes the hazards of the generalized prisoner’s dilemma. But the absence of envy is also an argument for the relative stability of justice as fairness. In the third part of *TJ* and in parallel with other considerations regarding the problem on stability, Rawls compares the relative stability of justice as fairness with other conceptions of justice. This comparison serves as an important confirmation of the arguments in favor of the two principles of justice made in the first part of the book. As Rawls (1971, 498) notes, “other things equal, the preferred conception of justice is the most stable one”.<sup>26</sup>

One of the reasons the two principles of justice are more stable from the utility principle or from a conception of strict equality, is that they lead to social structures in which envy is not likely to be strong (Rawls 1971, 144). In the third part of *TJ*, in the section “The problem of relative stability”, Rawls (1971, 499) argues that in a well-ordered society regulated by the two principles of justice, citizens know that others don’t take advantage of their bad luck and that they have an unconditional concern for their good and this caring strengthens their self-esteem. The utilitarian conception instead, is destructive of the self-esteem of those who rank low in

---

26] There is an ambiguity in the justificatory role of stability in *TJ*. There are passages where Rawls (1971, 498) claims that relative stability is crucial in the justification of the two principles: “There seems to be no doubt then that justice as fairness is a reasonably stable moral conception. But a decision in the original position depends on a comparison: other things equal, the preferred conception of justice is the most stable one. Ideally, we should compare the contract view with all its rivals in this respect, but as so often I shall only consider the principle of utility”. But a few pages later Rawls (1971, 504) writes: “These remarks are not intended as justifying reasons for the contract view. The main grounds for the principles of justice have been presented. At this point we are simply checking whether the conception already adopted is a feasible one and not so unstable that some other choice might have been better”.

the distribution of wealth and income. Given how crucial self-respect is for the psychological immunity against destructive feelings of envy, the two principles of justice are more stable than the principle of utility. The greater stability of the two principles is already implied in the first part of *TJ*, in the section “Main grounds for the two principles”, where Rawls (1971, 178) claims that one of the reasons for the choice of the two principles of justice in the original position is that they give greater support to self-respect and this increases the effectiveness of social cooperation. If they chose the utility principle they wouldn’t have the support that the reciprocal character of the two principles provides to their self-esteem.

The test of envy is also important for the comparison of the two principles of justice with a conception of strict equality. Although, such a conception is initially excluded in the original position, the parties would have reason to reconsider it if the inequalities allowed by the difference principle are so great that excite envious feelings to a dangerous extent. This could be the case for example “if how one is valued by others depends upon one’s relative place in the distribution of income and wealth” or if wealth, income and other goods are “fixed and cannot be enlarged by cooperation” (Rawls 1971, 545). Rawls examines this possible scenario in the section “The Grounds for the Priority of Liberty” and he concludes that strict equality should not be considered as a solution, since in a well-ordered society the self-respect of citizens is secured by their equal civic status.<sup>27</sup>

#### IV. THE REVISED ACCOUNT OF STABILITY

The case that the third part of *TJ* contains an argument from the absence of envy, apart from the congruence argument, is supported by Rawls’s account of stability in *JFR*.

In this book Rawls reformulates the presentation and defense of justice as fairness, integrating the changes he made in *PL* and distancing himself from the partially comprehensive view of the good that he presupposed in *TJ*. This reformulation shows how justice as fairness can be understood as a form of political liberalism and not as a comprehensive doctrine. A comprehensive doctrine is a set of beliefs that “covers the major religious, philosophical, and moral aspects of human life in a more or less consistent and coherent manner” (Rawls 2005, 58). It is comprehensive in the sense that it covers a wide range of values and is not restricted to political ideals. In *TJ* justice as fairness was presented as a comprehensive

---

27] It is worth mentioning that Rawls acknowledges more than once that a socialist regime could be considered as a possible solution to the problem of stability. In his paper “Fairness to Goodness”, he writes that “the principles of justice do not exclude certain forms of socialism and would in fact require them if the stability of a well-ordered society could be achieved in no other way” (Rawls 1975, 546). In *JFR*, Rawls raises the question if a liberal socialist regime would have a better chance of stably realizing justice as fairness than a property-owning democracy.

doctrine because it contained ideals of personal conduct.<sup>28</sup> Rawls recasts all the basic arguments and ideas of *TJ* and presents them in a way that is consistent with a political conception of justice. A political conception of justice is not derived from a single comprehensive doctrine but from ideas implicit in the public political culture of a democratic society and it is compatible with a wide range of different comprehensive doctrines.

One of the arguments he revises is of course the stability argument. In *PL* Rawls makes a radical change in his account of stability, introducing the idea of an overlapping consensus between different comprehensive doctrines and repudiating the idea of congruence between the right and the good. As expected, this change is preserved in *JFR*. Rawls (2001, 181) writes that one of the main questions regarding stability is whether in view of the general facts that characterize a democracy's political culture, and in particular the fact of reasonable pluralism, the political conception can be the focus of an overlapping consensus. The attempt to show congruence between the right and the good is abandoned, since it can't be assumed that citizens accept a particular comprehensive view to which justice as fairness belongs (Rawls 2001, 187). Along with the congruence argument, Rawls abandons in *JFR* the whole game-theoretic view. There is not any mention to the prisoner's dilemma or to any other kind of games.<sup>29</sup>

But although Rawls has removed the congruence argument in *JFR*, he still affirms that the two principles of justice should be tested against the destabilizing effects of envy and other special psychologies. In particular, Rawls (2001, 181) writes that the parties in the original position must check whether people who grow up in a well-ordered society acquire a strong sense of justice and are not moved to act by envy and spite. This is one of the two criteria for the stability of a conception of justice, along with the possibility of an overlapping consensus.<sup>30</sup>

If the test of envy was just a part of the congruence argument, one would expect that Rawls would stop using it after the revisions he made in *PL*. The fact that he

---

28] These ideals are based on the Aristotelian principle and the Kantian conception of the person. In *TJ* Rawls assumes that citizens of well-ordered society would form their plans of life by giving priority to the cultivation of their talents and to their autonomy.

29] There is a reference to the problem of mutual assurance, although it is not presented in a game-theoretic context and is viewed more as a necessary condition of stability and less as a distinct threat. Rawls (2001, 196) says that citizens are ready and willing to do their part in those arrangements provided they have sufficient assurance that others will also do theirs. This assurance is given by the existence of an overlapping consensus, since citizens know from past experience that others comply with liberal arrangements and come to have increasing trust in one another (Rawls 2001, 197).

30] Rawls notes in *JFR* that the discussion in sections 81 and 82 of *TJ* addresses sufficiently the problem of envy and doesn't need substantial changes. The only addition to the argument from envy is the general remark that the more citizens see their political society as good and the greater their appreciation of the political conception in securing the three essentials of a stable regime, the less they will be prompted by envy and other special attitudes (Rawls 2001, 202).

maintains this test as a criterion of stability in *JFR*, which was published in 2001, supports the view that Rawls treats the destabilizing effects of envious feelings as a distinct problem from the hazards of a generalized prisoner's dilemma. In *TJ* the argument from the absence of envy solves the first problem and the congruence argument the second.

## V. THE POLITICAL TURN

In *PL* Rawls abandons any attempt to prove that justice as fairness is a stable conception because it enables the congruence between the right and the good. He aims to show instead that justice as fairness can be the focus of an overlapping consensus between opposing conceptions of the good.

Despite of this sweeping change, the central elements of the account of stability presented in *TJ* remain intact in *PL*. Rawls still holds that we have a capacity for a sense of justice and that under the favorable conditions of a well-ordered society citizens will develop this capacity. Furthermore, he still views stability as a moral problem which is distinct from the Hobbesian concern for social order.

But does he still present the problems of assurance and isolation as possible sources of instability? In *Why Political Liberalism? On John Rawls's Political Turn* Weithman argues that in *PL* Rawls deploys the idea of overlapping consensus so that he can solve the mutual assurance problem and the prisoner's dilemma, this time taking into account the fact of reasonable pluralism. The problem now is that in a liberal democratic society there is a plurality of different comprehensive doctrines and each citizen needs an assurance that other citizens' conception of the good doesn't provide sufficient reason to act against the demands of justice. The existence and public knowledge of an overlapping consensus solves according to Weithman (2010, 297) this mutual assurance problem.

In *PL* however there is not a single mention on the prisoner's dilemma or any other collective action problem. One would expect that if Rawls had in *PL* a game-theoretic approach to stability he would have presented the nature of the problem he tries to solve in a game-theoretic way.

Rawls mentions though something else in the introduction of the second edition of *PL* which sheds more light on the problem of stability he tries to solve in *PL*. He notes that *TJ* and *PL* deal with two different sources of conflict (Rawls 2005, lviii):

There are three main kinds of conflicts: those deriving from citizens' conflicting comprehensive doctrines; those from their different status, class position, and occupation, or from their ethnicity, gender and race; and finally, those resulting from the burdens of judgments. Political Liberalism mitigates but cannot eliminate the first kind of conflict, since comprehensive doctrines are politically speaking, unreconcilable and remain inconsistent with one another. However, the principles of justice of a reasonably just constitutional regime can reconcile us to the second kind of conflict. For once we accept principles of justice or recognize them as at least reasonable (even though not as the most reasonable) and know that our political and

social institutions conform to them, the second kind of conflict need no longer arise, or arise so forcefully. I believe that these sources of conflict can be largely removed by a reasonably just constitutional regime whose principles of political justice satisfy the criterion of reciprocity. *PL* does not take up these conflicts, leaving them to be settled by justice as fairness (as in *Theory*) or by some other reasonable conception of justice. Conflicts arising from the burdens of judgments always, however, remain and limit the extent of possible agreement.

It is my contention that we can have a better understanding of the political turn if we see it as an attempt to address a source of conflict that didn't concern Rawls in the third part of *TJ*. In *PL*, Rawls doesn't mention income inequalities, the alienation of less privileged citizens and the problems that arise from the division of labor, not because he doesn't believe that they pose a threat to the stability of a well-ordered society, but because he focuses on the conflicts that arise from inconsistent and unreconcilable comprehensive doctrines and especially between religious and liberal comprehensive doctrines. This is the source of instability he is concerned with in *PL*. The extensive revisions he made to *TJ* serve the purpose of presenting a liberal conception of justice which can address this source of conflict. If we see these revisions as a more realistic attempt to solve the generalized prisoner's dilemma presented in *TJ* we don't grasp that *PL* marks a shift of focus in Rawls's project.

It is of course possible to describe the conflicts between religious and liberal comprehensive doctrines in game-theoretic terms. In a society with a plurality of comprehensive doctrines citizens can be uncertain as to whether one doctrine may seek to achieve national preeminence. They need an assurance that others are going to respect the demands of justice. If not, the mistrust between the different comprehensive doctrines might weaken citizen's commitment to their sense of justice. But this description fails to capture the real character of the political turn. It undermines the change of focus in Rawls's thinking. In *TJ* the problem of reasonable pluralism is ignored and there is not any discussion on how the values of a liberal democratic society can be compatible with the values of religious comprehensive doctrines. If we describe the conflicts between different comprehensive doctrines as a special case of a collective action problem, then we risk trivializing their depth.

## VI. CONCLUSION

A fundamental question that Rawls addresses in *TJ* is that of political and social stability. Institutions must not only be just but they must generate their own support over time. If a just conception of justice fails to achieve this, then another one must be considered.

Rawls's argument for the stability of justice as fairness is developed in two stages. First, he aims to show that a well-ordered society regulated by the two principles of justice would bring about in its members an effective sense of justice. Second, he claims that their sense of justice would be strong enough to outweigh propensities to act otherwise.

This claim is based on the idea of congruence. The members of a just society would have a strong sense of justice because they would affirm that it is part of their good.

According to a game-theoretic view of Rawls's account of stability the congruence argument is the culmination of his overall argument for the stability of justice as fairness. This view is based on the assumption that Rawls's aim in the third part of *TJ* is to present a solution to a generalized prisoner's dilemma and the related problem of assurance. It finds textual support in section 86 of *TJ* where Rawls writes that the hazards of the generalized prisoner's dilemma are removed by the match between the right and the good. But other passages from *TJ* indicate that Rawls treats envy as a distinct source of instability which is not part of a collective action problem. In the sections 80-82 of *TJ* he shows that the two principles of justice would not generate feelings of envy to a socially dangerous extent.

My thesis is that the overall stability argument in *TJ* is developed in three parts. In the first part Rawls claims that people in a well-ordered society would acquire a sense of justice. He then argues that members of a well-ordered society would not be moved by feelings of envy and that their sense of justice would be congruent with their good. The last two parts of the argument respond to two different tendencies that might prompt citizens to act against their sense of justice.

This reading of the problem of stability in *TJ* can deepen our understanding of Rawls's political turn and is aligned with the revised account of stability in *JFR*.

*manolatosalex@yahoo.gr*

#### REFERENCES

- Barry, Brian. 1995. John Rawls and the Search for Stability. *Ethics* 105 (4): 874-915.
- Edmundson, William. A. 2017. *John Rawls: Reticent Socialist*. Cambridge: Cambridge University Press.
- Freeman, Samuel. 2007. *Justice and the Social Contract. Essays on Rawlsian Political Philosophy*. New York: Oxford University Press.
- . 2003. Congruence and the Good of Justice. In *Cambridge Companion to Rawls*, edited by Samuel Freeman, 277-316. New York: Cambridge University Press.
- Galisanka, Andrius. 2017. Just Society as a Fair Game: John Rawls and Game Theory in the 1950s. *Journal of the History of Ideas* 78 (2): 299-308.
- Kant, Immanuel. *Religion within the Bounds of Bare Reason*. Translated by Werner S. Pluhar. Indianapolis: Hackett Publishing Co, 2009.
- McClennen, Edward F. 1989. Stability and the sense of Justice. *Philosophy & Public Affairs* 18 (1): 3-30.
- Quong, Jonathan. 2014. On the Idea of Public Reason. In *A Companion to Rawls*, edited by David A. Reidy, and Jon Mandle, 265-81. Chichester: Wiley Blackwell.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge MA: Harvard University Press.
- . 1975. Fairness to Goodness. *The Philosophical Review* 84 (4): 536-54.
- . 2005 [1993]. *Political Liberalism*. New York: Columbia University Press.

- . 2001. *Justice as Fairness: A Restatement*. Cambridge MA: Harvard University Press.
- Thrasher and Vallier. 2015. The Fragility of Consensus: Public Reason, Diversity and Stability. *European Journal of Philosophy* 23 (4): 933-54.
- Weithman, Paul. 2010. *Why Political Liberalism? On John Rawls's Political Turn*. New York: Oxford University Press.

# Capabilitarianism without Paternalism?

Jens Peter Brune  
University of Greifswald

**Abstract:** Despite the exceptional growth of interest in the Capability Approach (CA) in specific fields of application, the approach has yet to clarify certain basic concepts, such as freedom, functioning, and, in particular, its key concept of capability. (1.) Taking two basic tenets of CA as my starting point, I would first of all like to (2.) contribute to the clarification of these central concepts. It thereby becomes apparent that “capability” is, on the one hand, an essentially hybrid concept with both an internal and external aspect. On the other hand, it addresses the freedom of choice between various lifestyles as a second-order competency. (3.) Against this background, this paper offers a suggestion to correct the terminology used by Martha C. Nussbaum in her concept of capability, and subsequently it will outline a basic model of the transformation of resources into forms of being and activities with reference to Ingrid Robeyns. In conclusion I will couple this rather static perspective with two critical points: (4.) First of all, a more dynamic view reveals hitherto barely considered aspects of freedom in everyday human lives. (5.) Secondly, the idea of capabilities opposes the attempt at a separation of capabilities from functions in order to avoid paternalism, as suggested by Nussbaum.

**Key words:** capabilities, functionings, substantial freedom, 2<sup>nd</sup> order competence, paternalism.

## I. TWO CENTRAL IDEAS OF CA

Over the past 30 years, the Capability Approach (CA) has evolved into a reference framework for diverse and interdisciplinary areas of research. It has been applied in the fields of welfare economics, developmental policy, gender studies, and educational research. It has proved especially fertile in the area of poverty research and has been put to use in various countries and regions.<sup>1</sup> In terms of philosophical discourse, CA rivals other theories of justice, is acknowledged in specific spheres of ethics, such as economic and environmental ethics, and has also been raised the debate on animal ethics (Nussbaum 2004, 2006; Wissenburg 2011; Melin and Kronlid 2016). At the same time, its advocates, who I shall henceforth refer to as “capabilitarians”, adopting the term used by Ingrid Robeyns (2016) in the title of one of her essays, admit that despite the conceptual resources at its disposal, CA has yet to develop a comprehensive theory of justice. This has resulted in a longstanding debate as to whether it is even a uniform approach and, if so, what are its basic defining elements (cf. Crocker 2008, 55; Nussbaum 2011a, 18-19; Robeyns 2016 [2011]).

---

[1] In addition to Amartya Sen’s early studies (Sen 1977, 1981) and his studies on India published together with Jean Drèze (Drèze und Sen 1997, 2002), here are just a few examples: Volkert (ed. 2005) and Arndt et al. (2006) on Germany, Krishnakumar and Ballon (2008) on Bolivia, Panzironi and Gelber (2012) on the Asia-Pacific region, and Vollmer (2013) for Mozambique. Additional case studies can be found in Deneulin, Nebel, and Sagovsky (2006), Ibrahim and Tiwari (2014), as well as Otto, Walker, and Ziegler (2018).

Nevertheless, we can discern two aspects that, when formulated in a sufficiently generalized manner, clearly apply for all normative (ethical) sophisticated variants of CA. The *first* aspect focuses on the supposedly simple matter of what people actually do with their lives and who they can be.<sup>2</sup> The primary aim of this question is to analyze the conditions under which people have the opportunity to realize their particular conception of a good and flourishing lifestyle or individual well-being. The second key aspect is to frame the normative requirements, the actual social, political, and economic conditions, in such a way that all individuals have an *equal opportunity* to realize their conception of well-being or a good lifestyle, at least to a certain threshold level.<sup>3</sup> This is the crucial challenge faced by CA in establishing a theory of justice.

## II. FUNCTIONINGS, CAPABILITIES AND FREEDOM

What individuals actually do and who they are capable of becoming is generally expressed in CA with the help of three central concepts: capability, functioning, and freedom. In his 1993 essay, *Capability and Well-Being*, Amartya Sen provides a preliminary answer to the question of what is meant by “functionings” and “capability,” as well as how they are related. His answer is meant to serve as the starting point for a step-by-step expansion of the approach with the goal of achieving a basic model:

*Functionings* represent parts of the state of a person – in particular the various things that he or she manages to do or be in leading a life. The *capability* of a person reflects the alternative combinations of functionings the person can achieve, and from which he or she can choose one collection. (1993, 31)

Here, the term “functioning” stands for everything that a person *actually* does and has accomplished in their life. It stands for *realized* activities (“doings”) and states of affairs (“beings”). This last category includes such heterogeneous conditions as having health insurance or not, being a citizen or being stateless, and being law abiding or a criminal. Sen designates those activities and states of affairs that *could be* realized by a person at a particular point in time as an individual’s “capability” (in the singular form). It is comprised of the individual pool of attainment *opportunities*, thus an individual power or potential to realize various combinations of activities and states of affairs (cf. Robeyns 2017, 91-92). As the “ability to achieve” (Sen 1987, 36), capabilities are likewise regarded as *positive freedoms*:

---

2] “What are people actually able to do and to be? What real opportunities are available to them?” (Nussbaum 2011a, x; see Robeyns 2005, 94).

3] There is certainly no consensus among capabiltarians that the approach should focus on a threshold level, thereby accepting a sufficiency rule as a principle of distribution. This position is advocated to one degree or another by Anderson (1999, 2010) and Nussbaum (2006). Others, such as Arneson (2006), instead tend to support an egalitarian solution, while Robeyns (2016) does not exclude a prioritarian distribution principle along the lines of Rawls’ Difference Principle.

Functionings are, in a sense, more directly related to living conditions, since they are different aspects of living conditions. Capabilities, in contrast, are notions of freedom, in the positive sense: what real opportunities you have regarding the life you may lead. (Sen 1987, 36; cf. 1999, 74-75)

Nussbaum similarly places freedom on the side of capabilities and, in this respect, concurs with Sen's position:

In contrasting capabilities with functionings, we should bear in mind that capability means opportunity to select. The notion of freedom to choose is thus built into the notion of capability. (Nussbaum 2011a, 25)

In essence, the step of actualizing a realization option takes the form of a "choice" of activities and states of affairs that a person "has reason to value" in light of their idea of a good life (Sen 1999, 14, 18, 73, 293, *passim*; 2009, 231). As far as I can see, all normatively discerning versions of CA designate this evaluative moment of individual selection to realize or not to realize being and activity possibilities as the paradigmatic location of *freedom*.<sup>4</sup> However, one can only refer to a substantive freedom of a person to transform their existing capabilities into concrete forms of being and activities if certain *additional* and various requirements – each depending on the nature of the situation – are met. For example, a woman possessing a good physical constitution may nonetheless be extremely limited in her mobility if the laws of her country or local traditions prevent her from leaving the house without a male chaperon. Someone else might find themselves unable to obtain a desired job despite having the best qualifications, only because they are a member of a stigmatized minority. At this point, of significance are two references concerning terminology and one reference to the priority given to the evaluative moment as a moment of freedom:

(1) *Ambivalent use of the term "capability"*: It is impossible to find a consistent use of the term "capability" in the relevant CA literature. Instead, its use varies between two main meanings and their derivatives. The main meanings can be seen to concur with Aristotle's differentiation of two *modalities of possibility*: In the *Metaphysics* (1019b34), Aristotle differentiates between a possibility and a potency. In the first case, the possible state of affairs *has no reference* to a capacity and is therefore merely logically non-contradictory (a possibility), while the second case is a real possibility that *has reference* a capacity (potency).<sup>5</sup> In fact, capabilityarians regularly employ the term "capability" in accordance with either only one or a combination of the two following basic meanings<sup>6</sup>:

---

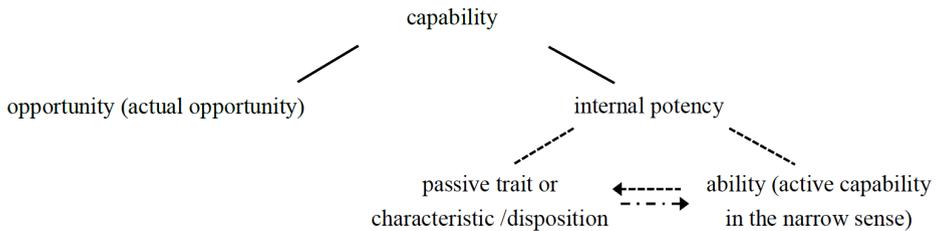
4] In addition to this concept of freedom that is tailored to individual life opportunities, Sen recognizes a second aspect of freedom that is related to (political) decision-making processes (cf. Sen 2002, 585).

5] What I refer to here as 'potency' corresponds to what Alan Gewirth characterizes as 'capacity,' which he distinguishes from 'capability': While capability is "primary active," capacity can "be passive as well as active; it is an ability not only to develop but also to be developed in certain ways" (1998, 63, n. 5).

6] In this respect, compare the analysis given by Crocker (2008, 171-77), who investigates the use of terminology by Sen and Nussbaum. Crocker comes to the conclusion that for Sen "a person's capability (for a particular functioning) is a possibility, option, freedom, or opportunity 'facing' the person. But this

- that of a “possibility” in the sense of a *logically* possible option, which *actually* ensues as a contingent “opportunity”;
- as an “ability” in the sense of a *potency* inherent in an individual, which, in turn, can be differentiated in terms of being active or passive, as well as poorly or strongly developed.

We refer to a “potency” in an active sense when we mean a capacity, a certain know-how, or even a skill. I will call this a *capability in the narrow sense* of the term (cf. Sedmak 2011, 32). Inasmuch as capabiltarians also include physical and psychological traits or characteristics of individuals as capabilities (see below), it appears that what is meant in this case are passive rather than active capacities. Regarding the transformation of such traits and characteristics into functionalities, one could again turn to Aristotle and speak of *dispositions* (attitudes). A capability in the narrow, active sense can then be regarded as the individual potency that can exert influence on external circumstances or internal states (feelings, thoughts, dispositions). The fundamental distinctions are summarized in Figure 1.



**Figure 1.** Internal differentiation within the concept of capability

It is quite obvious that one’s activities can exert an influence upon one’s feelings and dispositions. Conversely, there is no denying that feelings and dispositions can have an enabling, encouraging, or even an inhibitory effect upon our active abilities. As a case in point, James J. Heckman and Chase O. Corbin have demonstrated how various active capabilities reciprocally enhanced each other during ontogenetic skill formation within a dynamic model (2016, 344-48).

At first glance, the possibilities offered by a combination of an actual *opportunity* (O), either arising or not from external circumstances, and a capability (C), understood in the narrow sense of an ‘internal’ potency, are rather straightforward:  $C \wedge O$ ,  $C \wedge \neg O$ ,  $\neg C \wedge O$ , and  $\neg C \wedge \neg O$ . On second glance, however, opportunities as well as capabilities – whether active or passive in nature – can be quantifiably assessed: Something can be more or less advantageous, something else can be more or less defined. Although I have easy access

---

freedom may be due to a variety of internal factors, including abilities and other personal traits, as well as external factors” (2008, 172).

to sturdy building materials, for instance, my capability in the narrow sense (know-how, skills) to ensure that I am “well-sheltered” (Sen 1992, 45) is not notably developed. Conversely, a skilled but impoverished bricklayer or carpenter in a Manila slum cannot properly employ his skills if he only has plastic sheeting and corrugated sheet metal at his disposal. This *bifurcation* of the concept of capabilities is likewise reflected in the concept of freedom.

(2) *Equating freedom and capability*: There is a tendency, if not an explicit attempt, in many capability texts to use the terms “substantial” or “real” freedom and capability as semantic equivalents. To mention but a few examples: Sen makes reference to “substantive freedoms – the capabilities – to choose a life one has” (1999, 74) and to “the capabilities that a person has, that is the substantive freedoms he or she enjoys to lead the kind of life he or she has reason to value” (1999, 87; emphasis added by the author). Nussbaum speaks of “capability, in the sense not of coerced functioning but of substantial freedom to choose and act” (2011a, 24). Clearly intended as an explanation, Robeyns adds “capabilities” in brackets when referring to the term “substantive freedoms”: “substantive freedoms (capabilities)” (2005, 93, 111). Keleher mirrors “[o]ne’s capability set, including un-chosen options” with “an individual’s freedom to engage the world and make significant decisions about what she will be and do in her life” (2014, 62). Asserting such an equivalence, however, is by no means unproblematic, neither intensionally nor extensionally.

This is an apt moment to take a look at the intensional aspect of this equivalence: When we think of capabilities as substantial freedoms, then it clearly follows that at a particular point in time “t” we possess the *greatest* freedom of choice of activities (doings) and states of affairs (beings) when ‘C’ and ‘O’ are fully combined as ‘ $C \wedge O$ ’ in t, such that an individual could uncompromisingly realize their conception of a good life at that particular point in time. Viewed in this light, the normative objective of CA must be to at least achieve the threshold level of  $C \wedge O$  permanently and to the greatest possible extent for all cases of  $C \wedge \neg O$ ,  $\neg C \wedge O$ , and  $\neg C \wedge \neg O$ . Yet, since C is an internal aspect and O is an external aspect that to a certain extent<sup>7</sup> can be quantifiably gauged *independently of each other*, there must be a threshold value for C and one for O. As such, capability as a substantial freedom is basically a *hybrid* concept that always includes both an internal and an external aspect. Irrespective of whether the scope of justice is applied according to the distribution norms of egalitarianism, prioritarianism, or sufficiency theory (Kaufmann 2006; Davis and Wells 2016; Nielsen and Axelsen 2017), a uniform standard for the hybrid concept of substantial freedom cannot suffice.

(3) *Significance of the evaluative moment*: As a welfare and development economist, Sen leaves the evaluative moment and its reasons completely up to individuals themselves, and, in contrast to conventional welfare economics, treats this as part of the expanded data base of his theory. Sen would prefer to leave any discussion of relevant

---

7] The restriction “to a certain extent” is due to the possibility that capabilities in the narrower sense can be lost if, in the long term, there is insufficient external opportunity to exercise them.

value considerations and the assessment of their underlying reasons to the democratic decision-making process (Sen 1999, chap. 6; cf. Crocker 2008). By contrast, Nussbaum, in adopting certain Aristotelian arguments, initially understood her theory as a suggestion as to *what* elements constitute a good life – even though substantially divergent views on what the good life entails has since led her to adopt a political liberalism along the lines of Rawls (see below).

Despite this arguably greatest difference between the two, both Sen and Nussbaum have introduced their approaches as an alternative to the various versions of resourcism: inasmuch as John Rawls, Charles Beitz, and Thomas Pogge, on the one hand, and Ronald Dworkin, on the other, direct their focus on the fair distribution of more or less primary goods and services, they are thereby subject to the initial criticism of merely dwelling on the *means* to an end and not on the end itself:

An important problem arises from the fact that primary goods are not constitutive of freedom as such, but are best seen as means to freedom [...]. (Sen 1992, 80)

By referring to “freedom” as the actual, final end, Sen disavows any welfare theoretical and utilitarian determination of goals in the form of well-being, welfare, or utility (happiness, desire fulfilment, or preference satisfaction). As all the varieties of CA consider that primary goods, including negative freedoms (freedom *from* y, that is, from external despotism and repression), serve as a means to realize *positive freedoms* (cf. Crocker 2008, 121-22), the question immediately rises as to what kind of *freedom* (freedom to do what?) is actually meant here. Insofar as positive freedoms are understood as realization opportunities that can be embodied in capabilities, one is left to pose the analogue question in terms of capabilities: What kind of capabilities are ends in themselves and not merely further means to some higher-level goal?

Opinions differ sharply with respect to the answer to this question. The *first*, liberal, and materially parsimonious answer provided by Sen is that the purpose of freedom is to allow every individual to decide *for themselves* what they want to do and to become. Robeyns formulates this point as follows:

What is ultimately important is that people have the freedoms or valuable opportunities (capabilities) to lead the kind of lives they want to lead, to do what they want to do and be the person they want to be. (2005, 95)

Simply put, those who are *in the position* “to lead the kind of lives they want to lead” have the substantial freedom (capability) to transform those capabilities into functionings that, from their point of view, they have “reason to value.” In this sense, Sen also designates in *The Idea of Justice* (2009) freedom as an end: “primary goods are merely means to other things, in particular freedom [...]” (2009, 234). How should this be understood? For illustrative purposes, let us imagine someone who would like to lead a simple life consisting of optionally eating, sleeping, reading a book, riding a bike, or spending time with friends. Let us further assume that at a particular point in time, this person has

the capability of freely choosing between these five options, as well as combining and organizing them at her own discretion with the aim of realizing their idea of the good life. This person would therefore possess the relevant internal potencies (abilities = capabilities in the narrow sense) and the external opportunities “to lead the kind of life she wants to.” From this perspective, freedom of choice here is clearly related to the performance of a *decision-making competency* (cf. Leßmann 2011), namely a second-order capacity that pertains to the available specific first-order capacities (to eat, sleep, read a book, etc.):

Capability reflects a person’s freedom to choose between alternative lives (functioning combinations), and its valuation need not presuppose unanimity regarding one specific set of objectives (or, as Rawls calls it, ‘a particular comprehensive doctrine’). (Sen 1992, 83)

According to this interpretation, the determination of (final) ends lies in the hands of every individual – or in the hands of a democratically constituted society. The *liberal* version of CA thereby finds itself in a state of latent tension with respect to alternative, yet potentially more controversial options. These alternatives maintain that the actual end must *somehow entail a specified concept of the good life or a successful lifestyle*, as Nussbaum originally did with her “thick vague conception of the good” (1990, 217 ff.). There are a variety of starting points to arrive at such specifications. One could turn to sources in the history of civilization in order to attain an idea of humanity<sup>8</sup>, which could subsequently be broadened to include a concept of a good or flourishing life. This would then permit us to inquire as to which capabilities are embodied in this lifestyle (Nussbaum 1987, 1990, 1992, 1995). One could equally begin with an analysis of human abilities, identify criteria for particularly excellent capabilities (i.e. key capabilities, Sedmak 2011), and then make explicit the very nature of humanity embedded within these capabilities. Ultimately, both of these approaches lead to a quarrel about a consensual view of a successful lifestyle, which Sen would *preferably eschew* in order to create room for second-order capabilities. Since any lifestyle pre-determined to be good or flourishing certainly cannot be based on *arbitrary* capabilities chosen at the discretion of individuals or of democratically elected legislators. According to Nussbaum, it should be the task of philosophy to distinguish a qualified and therefore invariably *limited* spectrum of worthy capabilities, and in this way ensure

the philosophical underpinning for an account of basic constitutional principles that should be respected and implemented by the governments of all nations, as a bare minimum of what respect for human dignity requires. (2000a, 5, cf. 51, 116)

---

8] “What are the features of our common humanity, features that lead us to recognize certain others, however distant their location and their forms of life, as humans and, on the other hand, to decide that certain other beings who resemble us superficially could not possibly be human?” (Nussbaum 1990, 219).

It should be noted that the tremendous aspiration to specify capabilities as the object of generally respected and implemented constitutional principles concerning the exercise of political power would inevitably *restrict* the second-order freedoms of individuals:

The basic choice that Nussbaum leaves to individuals and communities is how to specify and implement the ideal of human flourishing that *she* – the philosopher – offers as the morals basis for constitutional principles. (Crocker 2008, 162)

Then again, a normatively ambitious CA, which demands the collective promotion of capabilities for all mankind, either through the state or international organizations, could not seriously uphold a *strict neutrality* with respect to the diversity of human capabilities. At the very least, CA should provide selection and evaluation *criteria* in order to allow for debate. Three brief comments should suffice at this point: If we make provisions for general conditions of shortage, the *first* question would concern the relevance of any capability with respect to the opportunity of a good or even only a decent lifestyle (Nussbaum 2000a, 13, 76; 2000b, 222-23; 2006, 44, 53; 2011a, 15, 29-42). For example, an extensive system of primary schools should be implemented before setting up schools for ventriloquists. *Secondly*, there are all kinds of functionings (such as shelter, medical care, and mobility) that can be realized through various capabilities and therefore could be evaluated in terms of *efficiency and sustainability*. Lastly, we would have to consider the *moral* quality of capabilities: “Many capacities inherent in our nature are bad (e.g. the capacity for cruelty)” (Nussbaum 2011b, 25; cf. 2000a, 83; 2011a, 61; Robeyns 2016, 406). A closer inspection of the dark side of human nature (Claassen und Düwell 2013, 496-97) reveals that under the surface of derogatory designations such as “cruelty”, there exists a broad range of capabilities with highly ambivalent potential: the expert cut into human flesh can equally serve to heal (surgery) or to torture (cruel infliction of harm). Many competencies that distinguish a good diplomat engaged in a difficult mission are those also shared by notorious swindlers. And the expert with a solid training in computer sciences can secure data from hostile attacks or alternatively engage in such attacks. The list of dual-use capacities is optionally expandable and can be easily expressed in the terminology of freedom. CA should therefore provide criteria that would aid in distinguishing between morally good, morally neutral, and evil realizations of capabilities.

### III. NUSSBAUM’S CONCEPT OF CAPABILITIES AND A SIMPLE BASIC MODEL OF CA

In terms of capabilities, Nussbaum’s variant of CA essentially differentiates “basic capabilities,” “higher-level capabilities,” with which she clarifies the internal differentiation between “internal” and “combined capabilities,” and “central capabilities.”<sup>9</sup> As, in my opinion, the ongoing discussion concerning Nussbaum’s list of central capabilities does not contribute anything essentially new towards the clarification of a basic model of

<sup>9</sup> Nussbaum employs this terminology (at the latest) with her essay “Capabilities and Human Rights” (Nussbaum 1997, 289-90).

capabilities, I prefer instead to concentrate here on explanations and offer corrective suggestions on terminology concerning (1) basic and (2) higher-level capabilities in order to, along with Ingrid Robeyns, establish the connection between the central concepts of CA (3).

(1) Keeping with the spirit of early poverty research (Charles Booth, Benjamin S. Rowntree; cf. Gillie 1996), Sen still understood “basic capabilities” as those that correspond to more or less essential functionings needed for survival in situations of extreme deprivation, such as “the ability to be well-nourished and well-sheltered, the capability of escaping avoidable morbidity and premature mortality, and so forth” (Sen 1992, 45; cf. 1999, 36). His language maintains the characteristic style of approaches focusing on those basic needs and goods that primarily affect discernible physical living conditions.

In Nussbaum’s writings, the terminology shifts to a more pronounced *ontogenetic* perspective, which is directed towards the individual development of humans in a psychophysical unity. As representative of “basic capabilities,” she lists “the capability for speech and language, the capability to love and gratitude, the capability for practical reason, the capability to work” (Nussbaum 2000a, 84). Assuming that these can already be at the disposition even of an infant, Nussbaum understands the term “capabilities” here in the sense of very generally delineated innate equipment or innate powers (Nussbaum 2011a, 23). It should then be possible to differentiate and scale such powers, such as whether they require many years of practice or, to a certain degree, are ready to function from the very start, such as the capabilities of seeing and hearing (Nussbaum 2000a, 84). Without a doubt, the capability for practical reason and the capability to work are among those particularly requiring intensive practice. If we still wish to include these in the category of basic capabilities, the term “capability” seems to refer to a *predisposition* to a capability, thereby understood as a kind of potential to a potential. It is likely to be the case that basic capabilities as well as predispositions to capabilities do not simply develop autonomously but require an appropriate environment. Evidence for this can be seen in the widespread phenomenon of *stunting*, particularly among infants and young children in developing countries. Malnutrition and undernutrition can lead to developmental disorders even in the fetus, which can have an impact on the later training of motor and cognitive capabilities (cf. Handa and Peterman 2006; Grantham-McGregor et al. 2007; Groppo 2015).

(2) In terms of higher-level “trained or developed traits and abilities”, such as the capability “of participating in politics” or “to think critically or speak publicly” (Nussbaum 2011a, 22), Nussbaum recommends distinguishing between internal and combined capabilities: “I use the term combined capabilities for the combination of trained capacities with suitable circumstances for their exercise” (2008, 357; cf. 2000a, 84-85; 2011a, 22). Accordingly, the internal capacities include being able to think critically, be politically active, to speak in public, and to organize, which are all essential for the external guarantees of the freedom of speech, assembly, and association. Furthermore,

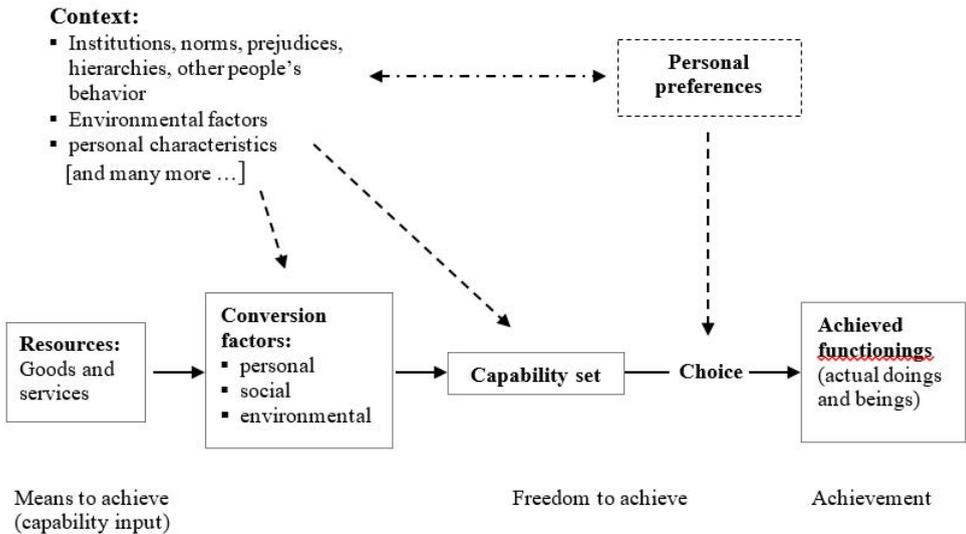
Nussbaum includes individual characteristics, such as “personality traits, intellectual and emotional capacities, states of bodily fitness and health, internalized learning, skills of perception and movement” (2011a, 21) to her list of internal capabilities. These are dynamic phenomena, and in order to flourish as *combined* capabilities, their training and maintenance require accommodating social, economic, and political framework conditions. Yet, even those capabilities classified by Nussbaum as “basic” require suitable *external* conditions in order to flourish (recall stunting), and, as substantial capabilities, would therefore also be “combined.” As such, I would prefer not to use Nussbaum’s classification of internal capabilities as combined, but instead distinguish them from “basic” capabilities with the designation of capabilities developed to a *higher level*. The second step leaves terminological *generalists* free to interpret the general possibility of participating in social life as a “capability” (Brighouse and Unterhalter 2010, 69), whereas terminological *specialists* can, for example, speak of someone possessing the professional competence to serve as a judge if that individual has passed the second legal state examination (bar exam). Of course, the borders between a generalized and a specialized perspective on capabilities are fluid, and the complex interaction between internal potential and external circumstances at all levels of development and specialization explain to no small measure the difficulties in measuring and comparing capabilities (Comim 2008; Nussbaum 2011a, 61). Nevertheless, for the sake of clarity, the generalists should be allowed to speak of capability sets for their purposes, while the specialists can refer to the specifying singular.

(3) In order to clarify the interaction between internal potencies and external conditions in the constitution of capabilities as substantial freedoms, I would recommend making use of what Sen referred to as “conversion factors” (1992, 79-87), namely those factors that co-determine if and how a capability or a set of capabilities can be transformed “into freedom of choice over alternative combinations of functionings [...]” (1992, 81) at a particular point in time with the help of actually available resources. Robeyns broke down these conversion factors according to personal, societal, and environmental factors with regard to the capability set of an individual and situated these capabilities in a social context (cf. 2005, 98-9; 2016). She has thereby arrived at a non-dynamic basic model of the transformation of resources into achieved functions by means of individual capability sets. Figure 2 adopts her representation, although it simplifies two aspects of the model.<sup>10</sup>

---

10] The first point is that on the input side of her representation, Robeyns also breaks down the sources through which people can obtain or finance goods and services: non-market production, market production, net income, as well as transfer in kind. The second point is that element that I have labeled as “personal preferences” is differentiated by Robeyns according to “preference formation mechanism” and “social influences on decision making” on the one hand, and “personal history and psychology” on the other (see 2005, 98). As these sources and differentiations are irrelevant for my considerations here, I have discarded them.

*Personal* conversion factors include individual characteristics, such as intelligence, talent and skills, physical condition (state of health), gender, and basic biochemical parameters,



**Figure 2.** Stylised non-dynamic representation of a person’s capability set and her social and personal context. Modified version adapted from Robeyns (2005, 98).

such as metabolism rate. When it comes to *social conversion factors*, a whole range of heterogeneous social characteristics and conditions come into play, ranging from laws, conventions, hierarchies, and cultural practices and traditions (including discriminatory attitudes based on gender or ethnicity) to public goods, such as infrastructure, educational system, and so forth. Also included here are those external guarantees of freedom of speech and assembly invoked by Nussbaum, which complement the internal capabilities of critical thinking and being politically active, enabling them to become combined capabilities and therefore de facto operative capabilities (freedoms). Environmental conversion factors include not only long-term climatic conditions and geographical locations, but also catastrophes occurring in the relative short-term, such as earthquakes, severe weather, hurricanes and floods, and drought.

#### IV. UNDERESTIMATED MOMENTS OF FREEDOM

Having regard this basic model of CA I would offer one brief explanatory comment and two lengthier immanent-critical remarks. The keywords here are (1) interdependence, (2) variability and (3) freedom.

(1) *Interdependence*: Numerous causal and epistemic relations, as well as correlates of responsibility, can clearly exist between the various types of conversion factors.

For example, *environmental catastrophes* may be causally attributable to collectively interconnected activities, including social factors. And the severity of the consequences of such environmental events for persons affected depends to no small measure on the extent and quality of the prognostic, preventative, and response resources deployed. In terms of *personal* factors, there is sufficient evidence to show that individual access to education as well as a balanced diet are dependent on social parameters, such as ethnicity, class, and gender. And with respect to recurrent political disputes on international trade agreements, the antagonists are nonetheless united in their acceptance of the premise that *institutional* factors exert a considerable influence on the work and living conditions of the affected populations – in either a positive or negative sense.

(2) *Variability*: Secondly, it is significant that, typically, the possibilities of realizing capabilities or even sets of capabilities are dependent on personal conversion factors. Under other circumstances, in a varied instantiation of the schema, these capabilities or sets could be similarly subsumed under personal conversion factors or even under achieved functionings. A vivid example deals with what Nussbaum classifies as an essential capability: “Being able to participate effectively in political choices that govern one’s life” (Nussbaum 1997, 288; 2003, 42; 2006, 77; 2011a, 34, etc.). It encompasses, as a combined capability, the fundamental and actual guarantee of political rights, including a sub-type of social conversion factors, which can, for example, be implemented in local environmental political activities (functionings), i.e. activists seeking to prevent the destruction of ecologically valuable marshlands as a result of road construction. Should the protest to maintain the biotope prove successful, the activists would thereby also prevent a local restriction upon their exercise of a further capability, similarly classified by Nussbaum as being essential, namely, “being able to live with concern for and in relation to animals, plants, and the world of nature” (1997, 288; 2003, 42; 2006, 77; 2011a, 34, etc.). Inasmuch as combined capabilities concern substantial freedoms, the activists could have devoted their “substantial” political freedom to uphold ecological *preservation* to another substantial freedom – the freedom to be able to develop a relationship to other living beings and to nature as a whole. The freedom/capability to engage in political activities is retained as long as the relevant parameters (legal frameworks, passive and active dispositions, conversion factors) do not change for the worse. In addition, the protagonists have gained the freedom to experience nature, which, presumably, would have been lost had they not intervened. According to this interpretation, we are dealing with an increase in freedom *relative* to a conceivable situation where there is a *lack of political freedom*. Looking back, the original political capability (in the full sense of a substantial freedom) could thus also be interpreted as a conversion factor for the capability to develop a relationship with nature. At the same time, this example of a variable instantiation illustrates a synergy effect, which, according to Heckman and Corbin (2016), can produce an interaction of various capabilities. This synergy effect can be formulated both in terms of capability as well as *in terms of freedom*.

(3) *Freedom*: As we have seen, CA situates the freedom of a person in their individual choice between alternative combinations of doings and beings against the background of existing capability sets. These capabilities embody, as it were, that what a person is in a position to achieve (freedom to achieve) at a particular point in time and under the given contextual conditions. The contextual conditions, in turn, can be analyzed on the one hand in terms of accessible resources as the means to pursue a particular way of life, and on the other hand in terms of institutional frameworks, the value systems and norms of the local community, natural environmental conditions, personal traits and characteristics, and so on. In particular, the extent, orientation, and scope of these possibilities for doing and being appear to be the determining factors in establishing what a person may achieve at the moment of decision. They *condition* the latitude of freedom intrinsic to the capabilities yet appear to be peculiarly detached from any access by the individual. What thereby emerges is a depiction of persons whose ways of life tends to be reduced at any time to a given inventory of realization opportunities in conflict with internal and external experiences. Capabilitarians' professionally widespread preoccupation with and concentration on the poor and poorest of this world may have assisted in the promotion of such a depiction. Nonetheless, this could advance the de-differentiation of the two moments of meaning in the guiding term "capability" as external opportunity and internal potency.

Without any further explanation, Robeyns designates her schema of a "person's capability set" as a non-dynamic depiction. A more dynamic view of the connection between functionings, capabilities, and freedoms emerges when we take into account the previously discussed *interdependence* of conversion factors and the *variability* in the attribution of capabilities under capability sets, functionings, and conversion factors. Speaking from a first-person perspective, through the decision to realize lower-level capabilities into actual functionings, I frequently develop, under ontogenetic auspices, the potential to other or higher-level capabilities and functionings. The moment of freedom that arises in the first choice is simultaneously passed on to the next level, which allows for the development of further and higher-level capabilities and functionings. Finally, this is reflected in personal conversion factors or even institutional norms (laws)<sup>11</sup>, which, in the form of framework conditions, will then allow me (or not) to employ a *then* given set of capabilities in the functionings I prefer.

While this potential for freedom is systematically underestimated in Robeyns non-dynamic model, it nonetheless appears to me that the *freedom to decide* on capabilities to be realized, seen as the paradigmatic location of freedom, is prone to misunderstanding and overestimation. One view, let us say one that is semantically naïve and empirically uninformed, could imply that between a given capability set, on the one side, and optional functionings (doings and beings), on the other, there existed a clearly defined

---

11] Correspondingly, Sen (1999, 36-40) differentiates between a "constitutive role," which, for instance, political freedoms have on development, and an instrumental function of freedom that "concerns the way different kinds of rights, opportunities, and entitlements contribute to the expansion of human freedom in general" (1999, 37).

and genuine leeway space, which would provide individual freedom to choose a solid and straightforward position to protect from paternalistic encroachments. CA would reserve this leeway space for every individual so that they could realize their preferred functionings in accordance with their own conception of the good life. Such a conception is certainly evoked by the previously mentioned semantic *equivalence* of capabilities with freedoms. In conclusion, we will examine one of Nussbaum's ideas to demonstrate that this simplified picture simply does not hold true.

#### V. HOW AN OBJECTION TO PATERNALISM CANNOT BE AVOIDED

Initially, Nussbaum recommended that political incentive measures should exclusively concentrate on capabilities (1997, 287). She later conceded that such a restriction was simply mistaken when considering certain groups of people. Basically, Nussbaum wants to adopt a politically liberal standpoint along the lines of Rawls, namely one that allows for a plurality of conceptions of a good life, as she recognized that the promotion of functions in the case of 'normal' adults "were precluding many choices that citizens may make in accordance with their own conception of the good, and perhaps violating their rights" (2000a, 87; cf. 2006, 171-72). This affirmation of political liberalism that perceives itself as being *anti-paternalistic* admits certain exceptions:

( $\alpha$ ) In the case "of people with severe mental impairments," who do not, for example, have the capacity to decide upon a designated medical treatment and therefore cannot be 'informed' in order to provide a possible 'consent,' they should rather be supported in their actual functionings (state of health) instead of capabilities (Nussbaum 2006, 172-73).

( $\beta$ ) Supporting functionings also remains a priority when this facilitates the development of certain adult capabilities in children. This presupposes the promotion of certain functionings in childhood, which in themselves serve as prerequisites for later capabilities: "exercising a function in childhood is frequently necessary to produce a mature adult capability" (Nussbaum 2000a, 90). Such a desired paternalism is reflected in compulsory education (Nussbaum 2006, 172, 377, 395; 2011a, 26, 148, 156).

( $\gamma$ ) Finally, Nussbaum is willing to allow for exceptions to the ban on paternalism when functionings are of key importance for the development or maintenance of other capabilities. Above all, she includes health, security, and environmental protection among the relevant spheres of life (2000a, 91; 2006, 172, 294).

At this juncture, I am not prepared to enter into the wide-ranging debate concerning various forms of paternalism (cf. Leßmann 2005; Nelson 2008; Claassen 2014). Instead, in two steps, I would prefer to show why Nussbaum's theory can hardly *avoid* the objection of paternalism, whether justified or not, because her primary focus is on capabilities and only in supposedly exceptional circumstances on functionings.

(1) The difficulties begin with the identification of capabilities and functions and their categorization along the lines of “basic” and “internal” capabilities. The characteristic of being “basic” can be understood along with Sen and early poverty research in the sense of being *urgent* for survival or the avoidance of “absolute deprivation” (Sen 1981, 17): “Basic capabilities” would then be those capabilities that ensure survival or, as the case may be, physical health within one’s environment (see above: 3. (1)). From the viewpoint of a developmental theory approach, however, basic could also be understood as a *Prius* [previous]. Accordingly, a capability would then be labeled basic in relationship to higher-level and/or more specific capabilities that build upon it (see above: III. (2)). Taken from this point of view, Nussbaum’s “capability to work” could now be seen ontogenetically as higher-level, even when perceived in a rather broad sense. At the same time, it could also be classified as “basic” in the sense of being “urgent,” to the extent that we assign normal adults as being in a “mature condition of readiness.” By contrast, in the case of small children, people with certain mental disabilities, or very old persons, we do not presuppose a mature capability to work nor – *a fortiori* – its urgency.

From a philosophical point of view, we could certainly leave the difficulties of the intensional and extensional delineation of a capability concept to relevant specialist disciplines, such as evolutionary psychology and learning theory. One thing, however, seems perfectly clear to me: The partitioning of capabilities according to types and levels is not something that is plainly apparent from the nature of the subject matter. Rather, it is dependent on theoretical positions and is thereby subject to particular interests. In particular, it provides absolutely no response to the *follow-up question*, which is of vital interest to any resource allocating entity, as to whether a person has developed a particular capability to a certain threshold level, without the person having even activated this capability: *The kind and degree of development of capabilities cannot at all be recognized, irrespective of relevant functionings.*

(2) I thereby come to the actual crucial issue: the differentiation between capabilities and functionings. The simple picture that clearly delineates capabilities from functionings is just not applicable for all those capabilities or even bundles of capabilities the acquisition of which could fit under the motto of “learning by doing.” Here, one could easily invoke examples of very heterogeneous capabilities of various difficulty and aggregation levels: *being able to* read, write, perform arithmetical calculations, make music, dance, swim, ride a bicycle, maintain social relationships, experience sexual satisfaction, as well as program computers, solve legal cases, or draft political programs. All of these capabilities require for their development that we actually exercise the relevant functionings (activities) to a particular degree – even if initially we are inept or amateurish – because it is only in this way that we can acquire the “tacit components” (Polanyi 1962) of the corresponding knowledge. As previously noted, Nussbaum does concede that “exercising a function in childhood is frequently necessary to produce a mature adult capability” (2000a, 90). Yet, in this case, it could appear as if the close interconnections of capabilities and functionings involved in learning were merely a specific feature of childhood. This would be misleading,

however, as empirical evidence makes amply clear: Widespread functional illiteracy among *adults* is not only explained by the circumstance of never having learned to read and write. Quite frequently, already acquired skills in reading and writing can be lost as a result of a protracted lack of practice (Buddeberg 2012, 206 f.; Sturm and Ziegler 2014). As such, it appears to me that a much closer interconnection between capabilities and functionings exists, namely one that spans the whole life cycle of an individual into old age. This also explains the basic thesis of CA as to why those capabilities falling under the heading of “learning by doing” as an *internal potency or disposition* will still always depend on suitable *external conditions*. Then, without actual opportunities for practice in the form of functionings, these capabilities can neither be developed nor be maintained at a previously achieved level. The idea of understanding capabilities as substantial freedoms requires exactly this interplay between external opportunities and internal potency (or rather capabilities *in the narrow sense*) and is thereby opposed to Nussbaum’s separation of capabilities from functionings in the name of avoiding paternalism.

In conclusion, if it is true that capability, understood as substantial freedom, is a hybrid concept that invariably includes an inner as well as outer aspect, then those versions of CA that base their normative ambitions more on *external opportunities* for the realization of internal potencies, tend to conform to views of Rawls. Therefore, along with “rights, liberties, and opportunities, and income and wealth” (Rawls 1999, 79), they also take into consideration a broad spectrum of basic goods. By contrast, those capabilitarians who place greater focus on supporting the internal aspect of capabilities by advancing a kind of “educational approach of justice” (Andresen, Otto and Ziegler 2010, 188), will thereby always take into account and promote functionings. As such, they, in particular, will be obliged to observe the sort of neutrality necessary to avoid the objection (legitimate or not) of paternalism.

*brune@zedat.du-berlin.de*

#### REFERENCES

- Anderson, Elizabeth. 1999. What is the Point of Equality? *Ethics* 109 (2): 287–337.
- . 2010. Justifying the Capabilities Approach to Justice. In *Measuring Justice. Primary Goods and Capabilities*, edited by Harry Brighouse and Ingrid Robeyns, Cambridge: Cambridge University Press: 81–100.
- Andresen, Sabine, Hans-Uwe Otto, and Holger Ziegler. 2010. Bildung as Human Development: An educational view on the Capabilities Approach. In *Capabilities – Handlungsbefähigung und Verwirklichungschancen in der Erziehungswissenschaft*, edited by Hans-Uwe Otto and Otto Ziegler, 165–97. 2. Ed. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Arndt, Christian, and Jürgen Volkert. 2006. Amartya Sens Capability-Approach – Ein neues Konzept der deutschen Armuts- und Reichtumsberichterstattung. *Vierteljahrshefte zur Wirtschaftsforschung* 75 (1): 7–29.

- Arneson, Richard J. 2006. Distributive Justice and Basic Capability Equality: 'Good Enough' Is Not Good Enough. In: *Capabilities Equality. Basic Issues and Problems*, edited by Alexander Kaufman, 17-39. New York and London: Routledge.
- Brighouse, Harry, and Elaine Unterhalter. 2010. Primary Goods versus Capabilities: Considering the Debate in Relation to Equalities in Education. In *Capabilities – Handlungsbefähigung und Verwirklichungschancen in der Erziehungswissenschaft*, edited by Hans-Uwe Otto and Otto Ziegler, 69-81. 2. Ed. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Buddeberg, Klaus. 2012. Literalität, Alter und Geschlecht. In *Funktionaler Analphabetismus in Deutschland. Ergebnisse der ersten leo – Level-One Studie*, edited by Anke Grotlüschen and Wibke Riekman, 187-209. Münster: Waxmann.
- Claassen, Rutger. 2014. Capability Paternalism. *Economics and Philosophy*, 30: 57–73. doi:10.1017/S0266267114000042
- Claassen, Rutger, and Marcus Düwell. 2013. The Foundations of Capability Theory: Comparing Nussbaum and Gewirth. *Ethic Theory and Moral Practice* 16: 493–510. doi: 10.1007/s10677-012-9361-8
- Comim, Flavio. 2008. Measuring Justice. In *The Capability Approach. Concepts, Measures and Applications*, edited by Flavio Comim, Mozaffar Qizilbash, and Sabina Alkire, 157-200. Cambridge: Cambridge University Press.
- Crocker, David A. 2008. *Ethics of Global Development. Agency, Capability, and Deliberative Democracy*, Cambridge: Cambridge University Press.
- Davis, John B., and Thomas R. Wells. 2016. Transformation without Paternalism. *Journal of Human Development and Capabilities* 17 (3): 360–76. doi: 10.1080/19452829.2016.1145198.
- Deneulin, Séverine, Mathias Nebel, and Nicholas Sagovsky. 2006. *Transforming unjust structures the capability approach*. Dordrecht: Springer.
- Gewirth, Alan. 1998. *Self-Fulfillment*. Princeton, NJ: Princeton University Press.
- Gillie, Alan. 1996. The Origin of the Poverty Line. *Economic History Review* XLIX (4): 715–30.
- Grantham-McGregor, Sally, Yin Bun Cheung, Santiago Cueto, Paul Glewwe, Linda Richter, and Barbara Strupp. 2007. Developmental potential in the first 5 years for children in developing countries. *Lancet*, 369 (9555): 60–70. doi: 10.1016/S0140-6736(07)60032-4
- Grosso, Valeria. 2015. Health consequences of childhood and adolescence shocks: Is there a 'critical period'? DIW Roundup [http://diw.de/de/diw\\_01.c.504995.de/presse/diw\\_roundup/health\\_consequences\\_of\\_childhood\\_and\\_adolescence\\_shocks\\_is\\_there\\_a\\_critical\\_period.htm](http://diw.de/de/diw_01.c.504995.de/presse/diw_roundup/health_consequences_of_childhood_and_adolescence_shocks_is_there_a_critical_period.htm) (accessed May 15, 2018).
- Handa, Sudhanshu, and Amber Peterman. 2006. Child Health and School enrollment. A Replication. *The Journal of Human Resources* XLII (4): 863-80.
- Heckman, James J., and Chase O. Corbin. 2016. Capabilities and Skills. *Journal of Human Development and Capabilities*, 17 (3): 342–59. doi: 10.1080/19452829.2016.1200541
- Ibrahim, Solava and Meera Tiwari, eds. 2014. *The Capability Approach: From Theory to Practice*. Palgrave Macmillan.
- Kaufman, Alexander. 2006. A sufficientarian approach? A note. In *Capability Equality. Basic Issues and Problems*, edited by A. Kaufman, 71–76. New York and London: Routledge.
- Keleher, Lori. 2014. Sen and Nussbaum: Agency and Capability-Expansion. *Éthique et Économique/Ethics and Economics* 11 (2): 54–70.
- Krishnakumar, Jaya, and Paola Ballon. 2008. Estimating Basic Capabilities: A Structural Equation Model Applied to Bolivia. *World Development* 36 (6): 992–1010.

- Leßmann, Ortrud. 2005. Der Capability-Ansatz von Sen als Gerüst für eine Neuinterpretation des Lebenslage-Ansatzes von Weisser. In *Armut und Reichtum an Verwirklichungschancen. Amartya Sens Capability-Konzept als Grundlage der Armuts-und Reichtumsberichterstattung*, edited by Jürgen Volkert, 149–69. Wiesbaden: VS – Verlag für Sozialwissenschaften.
- . 2011. Verwirklichungschancen und Entscheidungskompetenz. In *Der Capability-approach in sozialwissenschaftlichen Kontexten. Überlegungen zur Anschlussfähigkeit eines entwicklungspolitischen Konzepts*, edited by Clemens Sedmak, Bernhard Babis, Reinhold Bauer, and Christian Posch, 53–74. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Melin, Anders, and David Kronlid. 2016. Should We Ascribe Capabilities to Sentient Animals? A Critical Analysis of the Extension of Nussbaum's Capabilities Approach. *De Ethica* 3 (2): 53–63.
- Nelson, Eric. 2008. From Primary Goods to Capabilities: Distributive Justice and the Problem of Neutrality. *Political Theory* 36: 93–122.
- Nielsen, Lasse, and David V. Axelsen, 2017. Capabilitarian Sufficiency: Capabilities and Social Justice. *Journal of Human Development and Capabilities* 18 (1): 46–59. doi: 10.1080/19452829.2016.1145632
- Nussbaum, Martha C. 1987. Nature, Function, and Capability: Aristotle on Political Distribution. *Wider Working Papers* 31, December 1987. <https://www.wider.unu.edu/sites/default/files/WP31.pdf>
- . 1990. Aristotelian Social Democracy. In *Liberalism and the Good*, edited by B. B. Douglass, R. Mara, and H. Richardson, 203–52. New York: Routledge.
- . 1992. Human functioning and Social Justice: In Defense of Aristotelian Essentialism. *Political Theory* 20 (2): 202–46.
- . 1995. Aristotle on Human Nature and the Foundation of Ethics. In *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*, edited by J. E. J. Altham and Harrison Ross, 86–131. Cambridge: Cambridge University Press.
- . 1997. Capabilities and Human Rights. *Fordham Law Review* 66 (2): 273–300. <http://ir.lawnet.fordham.edu/flr/vol66/iss2/2>
- . 2000a. *Women and Human Development. The Capabilities Approach*. Cambridge: Cambridge University Press.
- . 2000b. Women's Capabilities and Social Justice. *Journal of Human Development* 1 (2), 219–247. doi: 10.1080/14649880020008749
- . 2003. Capabilities as Fundamental Entitlements: Sen and Social Justice. *Feminist Economics* 9 (2-3), 33–59.
- . 2004. 'Beyond Compassion and Humanity': Justice for Non-Human Animals. In *Animal Rights: Current Debates and New Directions*, edited by Cass R. Sunstein and Martha C. Nussbaum, 299–320. Oxford: Oxford University Press.
- . 2006. *Frontiers of Justice. Disability, Nationality, Species Membership*. Cambridge, Mass.: Harvard University Press.
- . 2008. Human Dignity and Political Entitlements. In *Human Dignity and Bioethics, Essays Commissioned by the President's Council on Bioethics*, Washington, D.C.: 351–80.
- . 2011a. *Creating Capabilities*. Cambridge, Mass.: The Belknap Press of Harvard University Press.
- . 2011b. Capabilities, Entitlements, Rights: Supplementation and Critique. *Journal of Human Development and Capabilities* 12 (1): 23–37. doi: 10.1080/19452829.2011.541731
- Panzironie, Francesca, and Katharine Gelber. 2012. *The Capability Approach. Development Practice and Public Policy in the Asia-Pacific Region*. London/New York: Routledge.

- Otto, Hans-Uwe, Melanie Walker, and Holger Ziegler, eds. 2018. *Capability-Promoting Policies: Enhancing Individual and Social Development*. Bristol: Polity Press. doi: 10.2307/j.ctt1zrvhn8
- Polanyi, Michael. 1962. *Personal Knowledge. Towards a Post-Critical Philosophy*. 2<sup>nd</sup> ed. London: Routledge.
- Rawls, John. 1999. *A Theory of Justice. Revised Edition*. Cambridge, Mass.: Harvard University Press.
- Robeyns, Ingrid. 2005. The Capability Approach: A Theoretical Survey. *Journal of Human Development* 6 (1): 93–114. doi: 10.1080/146498805200034266
- . 2016 [2011]. The Capability Approach. In *Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), edited by Edward N. Zalta, <https://plato.stanford.edu/entries/capability-approach/> (accessed May 15, 2018).
- . 2016. Capabilitarianism. *Journal of Human Development and Capabilities* 17 (3): 397–414. doi: 10.1080/19452829.2016.1145631
- . 2017. *Wellbeing, Freedom and Social Justice. The Capability Approach Re-Examined*. Cambridge, UK: Open Book Publishers. <http://dx.doi.org/10.11647/OBP.0130>
- Sedmak, Clemens. 2011. Fähigkeiten und Fundamentalfähigkeiten. In *Der Capability-approach in sozialwissenschaftlichen Kontexten. Überlegungen zur Anschlussfähigkeit eines entwicklungspolitischen Konzepts*, edited by Clemens Sedmak, Bernhard Babis, Reinhold Bauer, and Christian Posch, 29–52. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Sen, Amartya K. 1977. Starvation and Exchange Entitlements: A General Approach and Its Application to the Bengal Famine. *Cambridge Journal of Economics* 1: 33–59.
- . 1981. *Poverty and Famines. An Essay on Entitlements and Deprivation*. Oxford: Clarendon Press.
- . 1987. *The Standard of Living*. Cambridge: Cambridge University Press.
- . 1992. *Inequality Reexamined*. Oxford: Clarendon Press.
- . 1993. Capability and Well-being. In *The Quality of Life*, edited by M. C. Nussbaum and A.K. Sen, 30–53. Oxford: Clarendon Press.
- . 1999. *Development as Freedom*, New York: Anchor Books.
- . 2002. *Freedom and Rationality*. Cambridge, Mass/London: The Belknap Press of Harvard University Press.
- . 2009. *The Idea of Justice*. Cambridge, Mass.: Belknap Press of Harvard University Press.
- Sturm, René, and Petra Ziegler. 2014. Erwerbslosigkeit und Kompetenzerosion. Zu einer differenzierten Betrachtungsweise der PIAAC-Ergebnisse von 'erwerbslosen' Personen in Österreich. In *Magazin Erwachsenenbildung* 23, 09: 1–10. <http://www.erwachsenenbildung.at/magazin/14-23/meb14-23.pdf> (accessed May 15, 2018).
- Volkert, Jürgen, ed. 2005. *Armut und Reichtum an Verwirklichungschancen. Amartya Sens Capability-Concept als Grundlage der Armuts- und Reichtumsberichterstattung*. Wiesbaden: VS – Verla g für Sozialwissenschaften.
- Vollmer, Frank. 2013. *Poverty Revisited. The Capability Approach Operationalized in Mozambique*. Baden Baden: Nomos.
- Wissenburg, Marcel. 2011. The Lion and the Lamb: Ecological Implications of Martha Nussbaum's Animal Ethics. *Environmental Politics* 20 (3): 391–409.



# The Problem of Explanation and Reason-Giving Account of *pro tanto* Duties in the Rossian Ethical Framework

Hossein Dabbagh

Doha Institute for Graduate Studies & Institute for Cognitive Science Studies

**Abstract:** Critics often argue that Ross's metaphysical and epistemological accounts of all-things-considered duties suffer from the problem of explanation. For Ross did not give us any clear explanation of the combination of *pro tanto* duties, i.e. how principles of *pro tanto* duties can combine. Following from this, he did not explain how we could arrive at overall justified moral judgements. In this paper, I will argue that the problem of explanation is not compelling. First of all, it is based on the classical account of *pro tanto* duties. Principles of *pro tanto* duties can be understood in another way, i.e. in terms of reason-giving account that might be of help to provide a response to the critics. Furthermore, critics fail to see some evidence in Ross about how we can arrive at moral judgements.

**Key words:** W. D. Ross, *pro tanto* duty, all-things-considered duty, the problem of explanation, reason-giving account.

W. D. Ross's contribution idea of *prima facie* duty has been admired by A. C. Ewing as "one of the most important discoveries of the century in moral philosophy" (1959, 126). Ross believes that we should distinguish between *prima facie* duties and actual duties or a duty proper. He sometimes calls a duty proper a "toti-resultant attribute" of an action and a *prima facie* duty a "parti-resultant attribute" (1930; 2002, 28). Whether an act is a *prima facie* duty depends on whether it has a morally-relevant non-moral property. However, Ross believes that "[w]hether an act is a duty proper or an actual duty depends on all the morally significant kinds it is an instance of." (1930; 2002, 19-20)

Ross does not use the terminology of "*prima facie* duty" to refer to a duty that seems to be a duty *at first sight*, though it may turn out to be illusory on further reflection. Rather, the phrase "*prima facie* duty" is supposed to refer to a duty that comes from *part* of the nature of the moral situation. This duty that comes from *part* of the nature of the moral situation must be distinguished from overall duty that arises from the *whole* nature of the situation.

To illustrate this, suppose I promise my wife that I will go out with her tonight; however, my mother is sick, and I should stay with her. I have then two *prima facie* duties of fidelity and beneficence. To find out which duty is my actual duty, we have to use our judgement or perception in an "all things considered" way. We should consider the *whole* nature of the situation to reach our verdict. Of course, we may make a mistake in identifying something as a *prima facie* duty, but we may also make the same mistake in identifying something as an actual duty. Some of what we take to be actual duties may turn out to be illusory as well.

Moreover, it is not true that the contrast between *prima facie* and actual duties is a contrast between an *apparent* and a *real* duty. *Prima facie* duties are real moral considerations, not ones that seem to be duties at first sight. This leads us to think that a

*prima facie* duty, in fact, is *not a type of duty*. But it is *related to duty in a particular way*. Ross explicitly writes in this regard that:

[The *prima facie* duties] suggest that what we are speaking of is a certain kind of duty, whereas it is in fact not a duty, but something related in a special way to duty. (1930; 2002, 20)

So, Ross's distinction between *prima facie* and actual duties is not the distinction between apparent and real duties. I think the terminology "*pro tanto* duties" for *prima facie* duties and "all-things-considered duties" for actual duties is less misleading. Hooker, for instance, writes:

[T]he term "pro tanto" is less misleading than "prima facie". For the idea is that a duty or consideration is overridable, not that it can be seen at first glance but on closer inspection may prove to be an illusion. (1996, 534, fn. 6)

The idea behind *pro tanto* duties is that certain properties or features of situations count morally in favour or against action. Of these properties, some are more important in some case, but others are more important in other cases. For example, suppose that A borrows a gun from his friend B promising to give it back as soon as B asks. After one week, B asks for his gun. But A knows that B wants his gun to kill someone. On the one hand, A should keep his promise and return the gun to B. On the other hand, A has a *pro tanto* duty of non-maleficence. So relevant here are A's two *pro tanto* duties of fidelity and non-maleficence. We come to believe on reflection that of course A's all-things-considered duty is not to keep the promise.

## I. THE PROBLEM OF EXPLANATION

We can distinguish at least three claims in the Rossian ethical framework, i.e. two metaphysical claims and one epistemological one (Stratton-Lake 2002, x-xii; Audi 2004, 21). According to the:

1. *first metaphysical claim*: There are several (five) *pro tanto* duties or principles which are general. These principles determine the moral status of actions and contribute to the moral evaluation of concrete ethical situations.

And according to the:

2. *second metaphysical claim*: There is no hierarchy for these *pro tanto* duties, i.e. it is not the case that some *pro tanto* duties automatically outweigh any other in cases where they conflict. *Pro tanto* duties combine together and contribute to the deontic evaluation of different cases in such a way that we cannot know what would constitute the ultimate outcome (all-things-considered duty) in advance or even afterwards.

The second claim is controversial and must be explained. The issue is whether there is some way of thinking of *pro tanto* duties as a flexible hierarchy. For example, Dancy, as a commentator, believes that there is no structure to Ross's five basic *pro tanto* duties. Dancy thinks no sense can be given to the claim that promises are more important than,

say, good production. So, in his view, it is not true that one duty is more important than another. This leads Dancy to hold that there is no hierarchical order between *pro tanto* duties. Instead, we have a formless list of duties. Dancy writes:

There is no general ranking of the different types of *prima facie* duty. [...] There is just a shapeless list of them [...]. (1991, 221)

However, Dancy's interpretation might be wrong. In *The Right and the Good*, Ross lists five duties. He thought that these five basic duties or principles are the minimal number of such duties or principles (Ross 1930; 2002, 21, 27). Ross says that some *pro tanto* duties are more stringent than others, and in fact he offers a flexible hierarchy for them. He states that the duties of non-maleficence, fidelity and reparation are weightier than beneficence (Ross 1930; 2002, 22, 30). Elsewhere, in the *Foundations of Ethics*, he says that the *pro tanto* duty to keep a promise can be cancelled under certain conditions, and in those circumstances the fact that one has made a promise carries no deontic weight (Ross 1939, 188). Yet, the duties of non-maleficence, fidelity and reparation do not *always* outweigh beneficence. All Ross can say is that reparation, for instance, can be overridden only by a significant amount of well-being for others. However, an account of "significant" can vary from case to case (Stratton-Lake 2002, xxxvi-xxxvii).

Thus, the bottom line is that, on the basis of Ross's position, we cannot produce a higher-order principle to establish which *pro tanto* duties are more important than others. In fact, "the importance" or weight of each duty turns out differently in different contexts. So, we can restate the second claim this way. According to the:

2a. *second metaphysical claim-revised*: There are no rigid hierarchies (such as a lexical ordering) between *pro tanto* duties according to which one is always more important.<sup>1</sup> Although there is no such thing as a rigid hierarchy for *pro tanto* duties, there might be a flexible hierarchy for them.

One can think of a sort of flexible hierarchy for *pro tanto* duties as in many cases; for example, we can imagine that non-maleficence, fidelity and reparation *generally* outweigh beneficence. But it is not the case that we have a rigid hierarchy in that for example fidelity *always* outweighs beneficence. What is supposed to be the difference between the original version of the second metaphysical claim and the revised version is that although neither version accepts a rigid hierarchy, the second metaphysical claim states that there is no hierarchy for *pro tanto* duties but the revised version states that there might be a flexible hierarchy for them.

Finally, according to the

3. *epistemological claim*: The basic principles of *pro tanto* duties are self-evident in the sense that they *can* be known directly though sometimes they need *reflection*.<sup>2</sup> We *often* are

---

1] Audi (2004, 28) prefers to read Ross based on the second metaphysical claim; however, Stratton-Lake (2002, xxxvi-xxxvii) prefers to read Ross according to the revised version of second metaphysical claim.

2] Although Ross took all the *pro tanto* duties to be self-evident, one can ask why we should assume that they all are. However, since every system of morality takes some propositions to be self-evident, for the

justified in believing these moral principles non-inferentially.<sup>3</sup> In order to arrive at a moral judgement about our all-things-considered duty in a concrete moral situation, we have to use our *rational judgement*; however, these judgements are not self-evident.<sup>4</sup>

Taking some principles as self-evident does not entail that they are obvious and that there is no debate about them (Ross 1939, 188). We conclude what the ultimate outcome of the combination of principles of *pro tanto* duty is by looking at the case and reflecting on it in more detail in order to arrive at a moral judgement. Note that this is not a matter of non-inferential knowledge. The reason is that there is no such thing as a higher set of moral rules that can guide us to arrive at moral judgements (Ross 1930; 2002, 29).<sup>5</sup>

However, based on these claims, critics such as John Searle (1987, 81-90) object that Ross's metaphysical and epistemological accounts of all-things-considered duties, and the way in which different *pro tanto* duties are combined together and contribute to the moral evaluation and judgement in *moral conflict* cases, are vague and unclear. Let us call this:

*The Problem of Explanation:* Ross did not give us clear explanation of the combination of *pro tanto* duties, i.e. *how* principles of *pro tanto* duties can combine. Following from this, he did not explain *how* we could arrive at overall justified moral judgements.

This, however, needs to be qualified. As I have shown above, in the Rossian framework we cannot determine in the abstract what would be the ultimate outcome of a case in which, for example, fidelity is combined with another *pro tanto* duty, say gratitude, because the ultimate outcome strongly depends on the details of the case. Yet, as there is no rigid hierarchy for these principles of *pro tanto* duty, we have to judge how they combine together and contribute to the moral evaluation of the case. In order to do that, we pick out the outcome of the combination of several principles of *pro tanto* duty when we look at the case carefully and through further reflection on more details. So, according to Ross, in order to arrive at an all-things-considered judgement in a concrete ethical situation, one has to consider all the principles of *pro tanto* duties and think about the whole case and its characteristics insofar as one can. Let me elaborate this with an example of:

*The Manager Case:* Suppose that a manager of a company arranges an appointment that is crucial for the company, financially speaking. However, just half an hour before the appointment, he learns that his mother has been hospitalised following a car accident.

What is his all-things-considered duty? Would he be justified in cancelling his appointment to go and see his mother in hospital or should he fulfil his promise, leave his mother on her own and ask somebody else to see her? According to Ross, the manager

---

sake of argument, I assume that Ross is right and will not argue about this.

3] I say "often" because Ross holds in one of his papers that we *can* arrive at intuitive judgements even inferentially. See Ross (1927, at 121).

4] Ross thinks that there are no principles for all-things-considered duties.

5] Elsewhere, I explained what non-inferentiality entails in moral intuitionist framework, see Dabbagh (2017 & 2018).

has two different *pro tanto* duties, fidelity and gratitude (or beneficence), which come into conflict in this case. On the one hand, fidelity, which rests on the previous action of the manager, requires him to fulfil his promise. On the other hand, gratitude, which rests on the previous actions of his mother, requires him to stay with her in such an important situation. If the manager asks Ross how he can arrive at a justified moral judgement in this situation, what would Ross say?

The manager has to use his judgement to determine the outcome of this conflict of *pro tanto* duties. He must look at the case and reflect on it in more detail and consider everything that is morally relevant. So, one can understand how different *pro tanto* duties are combined and how we can arrive at a tenable moral judgement by looking at the case and reflecting on it in more detail. In fact, there is no other thing here to which we can appeal. Just exercise more reflection on the case, use your *judgement* and decide.

Critics, however, are not persuaded by the above Rossian story, because they are confused how “reflection” or “looking” is an explanation of anything. Consider again the manager case provided above. Later on, if the manager says that he still cannot see the point, what should he do to grasp the weightier principles of *pro tanto* duties? According to above story, the manager has to look at the case again, look again, and use his judgement. That is all. He cannot say anything else. So, the manager has to look at the case over time to arrive at a justified moral judgement. Nevertheless, the manager can say, “I am still perplexed, and I do not know straight off what to do”.

Critics believe that the manager is not, in effect, offered an account that tells him which is the most important of the conflicting *pro tanto* duties. Although the Rossian metaphysical account of the combination of principles of *pro tanto* duties in moral conflict cases and their contribution to the moral evaluation of different cases is understandable, the epistemological Rossian account is obscure. Thus, it seems that the Rossian account of *how* we could arrive at justified moral judgements in moral conflict cases has to be revised.

The problem of explanation is not, I think, compelling. First of all, it is based on the classical account of *pro tanto* duties, i.e. the *first metaphysical claim* (Dancy 1981). Principles of *pro tanto* duties can be understood in another way, i.e. in terms of reason that might be of help to provide a response to the critics (Scanlon 1998, 1-13; Stratton-Lake 2013; 2011b, 178; Audi 2004, 23-24). Furthermore, critics fail to see some evidence in Ross about how we can arrive at moral judgements. Let me explain.

## II. REASON-GIVING ACCOUNT OF PRO-TANTO DUTIES

In his *Second Thoughts in Moral Philosophy*, Ewing talks about *pro tanto* duties and all-things-considered duties with reference to reason. He distinguishes between good and conclusive reason for action (Ewing 1959, 63, 110). Having a *pro tanto* duty to do an action provides a good reason to do (or not to do) it. To have an all-things-considered duty for doing an action provides a conclusive reason for doing or not doing that. Of course, it is not the case that wherever there is a good reason to do *X* there is also a *pro tanto* duty to do

X. That is, it is not the case that if one has good reasons to do things, then one has real duty of any kind to do them. For instance, I have a good reason to take a painkiller right now, but no duty to do so. Although one can think that duties (or obligations) provide reasons, it is not plausible to say that reasons provide duties (or obligations), and hence we cannot accept that there are duties iff there are reasons. As another example, imagine that I have been very generous to Oxfam. I nevertheless have a reason to send even more money to that charity, the reason being that even more will do even more good. However, if I have already made very big contributions at a large cost to my own good, I have no duty to make further contributions.<sup>6</sup>

Following Ewing, Urmson and Stratton-Lake, I believe that understanding *pro tanto* duties in terms of reasons, i.e. as providing moral reasons to do acts, is more persuasive than the classic understanding (Urmson 1975, 112-13; Stratton-Lake 2011b, 147-150).<sup>7</sup> Principles of *pro tanto* duties specify facts that provide moral reasons for certain actions and explain why certain acts ought or ought not to be done. This is a kind of reply that Ross can give with regard to how *pro tanto* duties are combined.

However, one might object that it is not helpful to convert talk of *pro tanto* duties into talk of reasons. This conversion makes us pay a theoretical cost. All *pro tanto* duties are universal, although can be outweighed. They are defined in terms of their force and have an element of *insistence* in them (Owens 2012a and 2012b). But reasons purely as such do not have such an element. As a response, to be sure, *some* reasons explain and justify but do not insist, and others do insist.<sup>8</sup> The morally important reasons are mainly the ones that insist. For the sake of argument, by “moral reasons” here I mean the insisting ones. Talking about insistence, again, lends some credence to the point I made above about duties providing reason, and not *vice versa*.

Principles of *pro tanto* duties specify moral reason-giving facts. These principles state which facts provide reasons. I am not claiming that duties themselves are facts which provide reason. The duty might be “to meet you at 1:30”. The fact that provides the reason is “that I promised to meet you at 1:30”. Hence, the *pro tanto* duty of fidelity entails that the fact that I have promised to  $\Phi$  gives me a moral reason to  $\Phi$ .

This understanding of *pro tanto* duties in terms of reasons allows us to think of the conflict of *pro tanto* duties as a conflict of *moral reasons*. Therefore, to put my understanding of principles of *pro tanto* duties in a general form: if F is the feature that is *pro tanto* right and contributes in making  $\Phi$  your all-things-considered duty, then:

---

[6] See the papers by Darwall, Wallace and Little in the festschrift for Dancy, edited by Hooker, Bakhurst and Little (2013).

[7] Audi (2004, 104) also believes that Ross’s basic *pro tanto* duties are not only self-evident, but also provide independent moral reasons.

[8] Talk of reasons can be classified into two categories as “insistent” (obligatory) and “good ground” reason. Likewise, Dancy (2004), for example, talks about a distinction between two categories, i.e. “enticing” and “peremptory” reasons, that capture different normative relations.

“If F explains why you ought to  $\Phi$ , then F gives you a reason to  $\Phi$ ” (Stratton-Lake 2011a, 369).

Thus, fidelity, for instance, is a right-making feature. In other words, there are facts that provide reasons arising from fidelity; fidelity has an invariant deontic valence; and it contributes to the moral evaluation of different cases in the same way (Stratton-Lake 2002, xliii). But has the explanation problem been solved?

It is true that Ross said that looking at a moral case and reflecting on it gives us an account according to which we can arrive at a tenable judgement. This is how Ross expresses the point:

When I am in a situation [...] in which more than one of these *prima facie* duties is incumbent on me, what I have to do is to study the situation as fully as I can until I form the considered opinion (it is never more) that in the circumstances one of them is more incumbent than any other; then I am bound to think that to do this *prima facie* duty is my duty *sans phrase* in the situation. (1930, 2002, 19)

However, this is not the whole story. In some complicated cases, Ross would suggest that where relevant, we should aim to get further information, which might (for example) involve asking people what they are up to. He said that reflecting upon other cases might help us. Consider the following quotes from Ross:

The general principles of duty are obviously not self-evident from the beginning of our lives. How do they come to be so? [...] we see the *prima facie* rightness of an act which would be the fulfilment of a particular promise, and of another which would be the fulfilment of another promise, and when we have reached sufficient maturity to think in general terms, we apprehend *prima facie* rightness to belong to the nature of any fulfilment of promise. What comes first in time is the apprehension of the self-evident *prima facie* rightness of an individual act of a particular type. From this we come by reflection to apprehend the self-evident general principle of *prima facie* duty. (1930; 2002, 32-33)

In my view, we can develop the Rossian account by using the notion of “seeing other similarities” and seeing things as *similar*. Although Ross did not use these terminologies, I believe they already have a place in the Rossian framework. So, we can say that in the Rossian framework, *looking away* at other similar cases is an indispensable ingredient to understanding. Sometimes, just *looking at* one case and *reflecting* on it might not help us to reach a justified moral judgement. I believe *looking away* is important because one of the most significant and familiar forms of moral thinking involves thinking about cases *like* the one at hand but with the crucial difference that the roles are reversed – e.g. we have to think of ourselves as being on the receiving end of what we are considering doing to someone else. Let me explain my view with the following example.

*The Robbery Case:* Consider the case of a man chatting to his friend that one day when he was in rush to his meeting, he saw two boys hitting an old woman to steal her money. Having seen this scene, he decided to stop them but he also wanted to be punctual at his job. When he tells the story, his friend responds that the man should

have helped the women even if he was in rush. The man tells his friend that he was trying to keep his promise to be on time. His friend however gives the man some more details by citing examples that what he would do if he sees his mother or sister getting robbed. His friend also says some more examples to illustrate that helping people in danger is more important than punctuality.

In fact, the man's friend tries to convince him that he would not do the same in the case of his mother's robbery. His friend also tries to convince him by referring to other *similar* cases that helping people in danger is more important than being on time. His friend does not appeal to common properties and intrinsic features of stealing to improve his argument. So, instead of just looking at this case, the man has to *look away* at similar and dissimilar cases and ask other people to arrive at the judgement.<sup>9</sup>

According to Ross, different *pro tanto* duties are combined together and contribute to the moral evaluation of cases in different ways. But, on the basis of the element that I highlighted in the Rossian framework, we can give an account to resolve the problem of explanation: *looking away* at similar combinations of *pro tanto* duties in different cases and getting opinions from other people can help.

We now have an account of how we behave reasonably in different ethical contexts that is grounded in the way in which we are engaged in *looking away* at *similar* cases. In doing so, our behaviour in different contexts makes sense. One can understand how different principles of *pro tanto* duties are combined and how we can arrive at tenable moral judgements by looking at other *similar* cases, by asking other people and by *reflecting* on the different comparable reasons in more detail.

In the *robbery case*, the more the man *looks away* at *similar* cases, the more he can see whether he should have helped the woman. Similarly, in the *manager case*, he will arrive at the justified moral judgement to the extent that he is engaged in seeing *similar* cases. There is no theoretical account available that can be used to crank out the justified moral judgement. Rather, he has to *look away* at *similar* cases to see what has to be done in the case. For example, if the manager sees other similar cases, it is very likely that he comes to the conclusion that he should abandon his promise under the new condition.

However, there is one possible objection here. One might ask: does not "looking away", "reflecting on other cases" or "getting opinions from others" just generate a regress? If our object is the judgements of others or merely other similar factual situations, the same problem arises. As a response, although there are some complex situations, ones where our reaching a final judgement is difficult and we might need to ask others' opinions, Rossians would believe that this does not necessarily create a regress. Almost all pluralists and at least some monists (e.g. Kantian Contractualists and Rule-Utilitarians) might do the

---

9] Although "looking away" might not be the best terminology, it does not make it sound as though one is ignoring the facts rather than thinking about them. This terminology came to my mind in the discussion with Soroush Dabbagh. For more details, see Dabbagh (2006, Dissertation). Using Hare's terminology, one can also say that in going from *prima facie* duties to *actual* duties, we use our capacity of apprehension, but also a "social consensus" (2000, 88), and this is similar to my solution of looking away.

same in reaching the final judgement. So, this is not a special problem – if it is at all – for intuitionists. They all need to think about the difficult case and consult with other people. In the cases where the judgements of others or similar situations are at stake, the only plausible answer which Rossians can advise is to practise more and more. All we have, in these cases, is to practise reflection, comparing cases and getting opinions from others. But looking away at some other cases (such as cases in which you are in my shoes and I in yours) and getting other people’s opinions might be enough to stop an infinite regress as they give us plausible reasons to reach judgement. Although there is no such thing as a valid principle for resolving all regresses in different moral cases, it does not follow from this that nothing can be said to explain how to deal with such regresses.

### III. PRACTICE BRINGS YOU MASTERY

Having seen that the problem of the explanation of *how* different principles of *pro tanto* duty are combined can be removed, the same approach can be assumed with regard to the way in which we become competent with the principles of *pro tanto* duties. We can see that promise-keeping is right with practice and through seeing the similarities and dissimilarities, e.g. through seeing that that is a promise, that this is a promise, that that is not a promise, and so on.<sup>10</sup> The more we are engaged in seeing moral cases, the more we see what the principles of *pro tanto* duties are. For instance, if we are wondering whether or not the new situation with which we are dealing can be regarded as an example of gratitude, we have to *look away* at other cases of gratitude and compare the case at hand with them.

According to Ross, we arrive at moral principles about *pro tanto* duties by seeing what features of actions count either morally for or morally against actions in different instances. Perhaps, when Ross is talking about *sufficient mental maturity* and its vital role in grasping a *pro tanto* duty, he has something similar in mind. Consider the following quotes:

[...] when I reflect on my own attitude towards particular acts, I seem to find that it is not by deduction but by direct insight that I see them to be right, or wrong. I never seem to be in the position of not seeing directly the rightness of a particular act of kindness, for instance [...]. (1939, 171)

Seeing a number of examples, for Ross, helps us to see the *rightness* or *wrongness* of promise-keeping (or seeing the fact which states the *reasonableness* of promise-keeping). One could arrive at the general principle that promise-keeping is a *pro tanto* duty by seeing different cases. The more we see different promise-keeping cases, the more we see what promise-keeping is. In effect, we come to know, in Ross’s ideas, the *rightness* or *wrongness* of moral duties or principles by knowing particular instances. For instance, we come to know that promise-breaking is *pro tanto* wrong by becoming acquainted with some particular cases of promise-breaking. (Ross 1930, 2002, 32-33)<sup>11</sup>

10] For more on practice theories, see Thompson (2008). For an alternative view, see Owens (2011).

11] See Hooker (2002) for an alternative view.

The whole point is that the procedure of grasping a moral principle is open-ended. It does not follow from this that there is no such thing as principle. What does follow is that the more we are engaged in seeing different cases, the more we see what the principle is. In other words, the more we are engaged in seeing different *reasons*, the more we acquire mastery of how the *reasons* can contribute in different moral cases.<sup>12</sup>

#### IV. CONCLUSIONS

The Rossian ethical framework, critics have argued, has a problem concerning explanation. According to the problem, the way in which different principles of *pro tanto* duties are combined together and make all-things-considered duties, both metaphysically and epistemologically, is vague and unclear. In order to tackle the problem, I introduced the reason-giving account of *pro tanto* duties and suggested that by bringing in the notion of “looking away” and “seeing other similarities”, which already exist in the Rossian framework, we can give an account of how we arrive at a justified moral judgement in a concrete situation. The Rossian notion of principles of *pro tanto* duties can be better understood in the light of such illumination.

#### *Acknowledgment*

I would like to thank Philip Stratton-Lake and Brad Hooker, and the editor and anonymous reviewers, for their critical comments.

*hdabbagh@dohainstitute.edu.qa*

#### REFERENCES

- Audi, Robert. 2004. *The Good in the Right: A Theory of Intuition and Intrinsic Value*. Princeton University Press.
- Dabbagh, Soroush. 2006. *Moral Reasons: Particularism, Patterns and Practice*. Warwick University Dissertation.
- Dabbagh, Hossein. 2017. Sinnott-Armstrong Meets Modest Epistemological Intuitionism. *Philosophical Forum* 48 (2):175-99.
- . 2018. Intuiting Intuition: The Seeming Account of Moral Intuition. *Croatian Journal of Philosophy* 18 (1):117-32.
- Dancy, Jonathan. 1981. On Moral Properties. *Mind*, 90: 367-83.
- . 2004. Enticing Reasons. In J. Wallace, P. Pettit, S. Scheffler, M. Smith (eds.), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*. Oxford: Oxford University Press.
- . 1991. An Ethics of Prima Facie Duties. In P. Singer (ed.), *A Companion to Ethics*. Oxford: Blackwell: 219-29.
- Ewing, A. C. 1959. *Second Thoughts in Moral Philosophy*. London: Routledge & Kegan Paul.

[12] Scanlon and Stratton-Lake believe that we can define moral concepts such as *good* and *bad* in terms of reasons. In the same way, I think, the more we are engaged in seeing different cases, the more we find reasons. In effect, reasoning is a kind of practice. See Scanlon (1998, 96 & 96ff) and Stratton-lake (2002, xx and 2004).

- Hare, R. M. 2000. *Sorting Out Ethics*. Oxford University Press.
- Hooker, Brad. 2002. Intuitions and Moral Theorizing. In P. Stratton-Lake (ed.), *Ethical Intuitionism: Re-evaluations*. Oxford: Oxford University Press: 161-83.
- Hooker, Brad. 1996. Ross-Style Pluralism vs. Rule-Consequentialism. *Mind*, 105: 531-52.
- Hooker, Brad., Bakhurst, David., Little, Margaret. 2013. *Thinking about Reasons: Themes from the Philosophy of Jonathan Dancy*. Oxford University Press.
- Owens, David. 2012a. The Value of Duty. *Proceedings of the Aristotelian Society Supplementary*, Vol. 86 (1): 199-215.
- . 2012b. Promises. *The International Encyclopedia of Ethics*. H. LaFollette (ed.). Blackwell, pp. 4143-4155.
- . 2011. The Problem with Promising. In Hanoch Sheinman (ed.), *Understanding Promises: Philosophical Essays*. Oxford: Oxford University Press: 58-79.
- Ross, W. D. 1927. The Basis of Objective Judgments in Ethics. *International Journal of Ethics*, 37: 113-27.
- . 1939. *The Foundations of Ethics*. Oxford: Clarendon Press.
- . 1930 [2002]. *The Right and the Good*. P. Stratton-Lake, (ed.), Oxford: Clarendon Press.
- Scanlon, Tim. 1998. *What We Owe To Each Other*. Cambridge: Harvard University Press.
- Searle, John. 1987. Prima facie Obligations. In J. Raz (ed.), *Practical Reasoning*. Oxford University Press: 81-90.
- Stratton, Philip. 2004. Introduction. In P. Stratton-Lake (ed.), *On What We Owe to Each Other*. Blackwell, Oxford: 1-17.
- . 2011a. Recalcitrant pluralism. *Ratio*, 24 (4): 364-83.
- . 2011b. Eliminativism about derivative prima facie duties. In T. Hurka (ed.), *Underivative Duty*. Oxford University Press: 146-65.
- . 2013. Ross, W. D. In H. LaFollette (ed.), *The International Encyclopedia of Ethics*. Wiley-Blackwell.
- . 2002. Introduction. In P. Stratton-Lake (ed.), W. D. Ross's *The Right and the Good*. Oxford University Press: ix-xlix.
- Thompson, Michael. 2008. *Life and Action: Elementary Structures of Practice and Practical Thought*. Harvard University Press.
- Urmson, J. O. 1975. A Defence of Intuitionism. *Proceedings of the Aristotelian Society*, 75: 111-19.



# Feminism, Agency and Objectivity

Adelin Costin Dumitru

National University of Political Studies and Public Administration

**Abstract:** In this article I defend the capability approach by focusing on its built-in gender-sensitivity and on its concern with comprehensive outcomes and informationally-rich evaluation of well-being, two elements of Sen's work that are too rarely put together. I then try to show what the capability approach would have to gain by focusing on trans-positional objectivity (as Elizabeth Anderson does) and by leaving behind the narrow confines of states in favor of a more cosmopolitan stance. These preliminary discussions are followed by two more precise applications. At first, I show how a gender-sensitive capability approach that respects the criteria of trans-positional objectivity and cosmopolitanism can enhance the agency of women inhabiting third-world societies. I turn next to show how mainstream feminism can insulate itself against criticisms such as bell hooks' by switching to trans-positional objectivity in public reasoning.

**Key words:** agency, capabilities, gender-sensitive theory, informational parsimony, open impartiality, positional objectivity.

The capability approach was proposed as a way to account for the ubiquitous element of human diversity (Robeyns 2000, 6-7). Sen's background in social choice led to the development of an approach to justice that sacrifices simplicity for the sake of providing more adequate answers to the problems human beings encounter. In order to accomplish that objective, the capability approach is concerned not merely with the outcomes of applying distributive principles, but with what Sen calls *comprehensive outcomes*, which assess consequences as well as "social process, including the exercise of duties and responsibilities" (Sen 2010, 22). Furthermore, the capability approach has a built-in gender sensitivity that lacks from many other accounts of justice.

The capability approach thus has various elements that make it attractive for feminist theorizing.<sup>1</sup> By focusing on real freedom and on comprehensive outcomes, it can identify and provide adequate answers to problems that have a negative impact on the well-being of women, such as adaptive preferences (Teschl and Comim 2005). The main claim of the paper is that feminist studies would have much to gain by focusing on another, less discussed aspect of Sen's work, the concept of trans-positional objectivity (Sen 1982; 1992; 1993; 2010, 155-74). Furthermore, unlike current extensions of Sen's work on positional objectivity to feminist deliberative democracy (Anderson 2003), I argue that we must renounce at the arbitrary confinement of the scope of justice to states and instead advocate a global scope. Albeit Sen endorses open impartiality and public

---

1] I have in mind here Western feminist theories, committed to the idea that women are oppressed in various ways in patriarchal societies and that this should cease to be the case. I do not refer particularly to one or another feminist theory.

reasoning at a global scale, certain feminist applications of the capability approach have the aforementioned shortcoming.<sup>2</sup>

In order to accomplish this objective I proceed as follows. In the first section I introduce the various lines of Sen's arguments, focusing on his criticism of informationally-parsimonious theories of justice and on his defense of capabilities as the proper metric of distributive justice. In the second section I present Sen's case for trans-positional objectivity and his criticism of Rawls' "closed impartiality" (Sen 2010, 128). In the third section I argue that focusing on a narrow scope of justice, the state, flaws Anderson's application of Sen's work on positional objectivity. The fourth section shows what mainstream feminism would have to gain by incorporating the aforementioned considerations. At first, I show how a gender-sensitive capability approach that respects the criteria of trans-positional objectivity and cosmopolitanism can enhance the agency of women inhabiting third-world societies. I turn next to show how mainstream feminism can insulate itself against criticisms such as bell hooks' by switching to trans-positional objectivity in public reasoning (hooks 1999). Section V concludes.

#### I. ESCAPING INFORMATIONAL PARSIMONY – A JUSTIFICATION FOR THE CAPABILITIES APPROACH

One of Amartya Sen's far-reaching contributions has been the emphasis on informational constraints in social choice theory, political philosophy or ethics. Although the focus of this paper is on the latter two, Sen's work on social choice theory plays an important role in understanding some of his criticisms regarding utilitarianism. In his main contribution to the field of social choice, Sen showed how the anonymity and neutrality conditions left out important information regarding relations between individuals or the intensity of agents' preferences<sup>3</sup> (Sen 1970, 198-99). In the field of justice, Sen similarly showed how moral principles exclude certain categories of information and that usability of information is highly dependent on what principles are chosen (1985, 169-184; 1993, 73-88). For instance, utilitarianism proposes a flawed evaluation of well-being, since it comprises three distinct elements that combined reduce its informational scope, *i.e.* consequentialism, welfarism and sum-ranking (Sen 1985, 175). Sen argues that utilitarianism is informationally parsimonious and that it lacks the means to judge adequately the goodness of states of affairs. The main problem is with the welfarist component, which deliberately ignores all information that is not based on utility when evaluating end states, which reverberates on its muteness pertaining to evaluating actions: "had welfarism not been additionally imposed, consequentialism could have coexisted

---

2] I refer here mainly to Anderson, as she is the most prominent theoretician who has linked feminist problems with Sen's work on positional objectivity.

3] For an account of the lessons that can be learned by feminists from social choice and vice versa see Peter (2003).

with taking note of such things as the values and disvalues of actions through the valuation of states, which include these actions” (Sen 1985, 182)<sup>4</sup>.

What motivates Sen’s (and Nussbaum’s) capability approach is human diversity, which is purportedly not satisfactorily taken into account by other approaches. The human diversity aspect is strongly related to the evaluation of states of affairs. Capabilities have been proposed as replacements to other metrics of justice – what Sen calls evaluative space (1992, 20) and Cohen “currency of justice” (1989) – because those are severely limited by informational constraints. For instance, Rawls’ primary goods approach<sup>5</sup> is inadequate because it would give nothing to a cripple who has access to primary goods.<sup>6</sup> Rawls’ contractarian theory of justice is based on the premise that rational and reasonable parties placed behind a veil of ignorance will choose the principles of justice as fairness<sup>7</sup> (which stand in a lexicographical order) – the principle of an equal right to sets of equal basic liberties and the second principle, comprising the fair equality of opportunity and the difference principle (2001, 302). Rawls justifies the device of the veil of ignorance by holding that any personal contingencies have to be removed in order to reach a truly fair agreement (2001, 15)<sup>8</sup>. What he proposes is a political conception of justice for a democratic society, conceiving citizens as free and equal (Rawls 1999, 31) and taking part in the affairs of a “mutually advantageous cooperative venture” (1971, 112).

---

4] Elsewhere, Sen criticizes utilitarianism, for its leading to counterintuitive implications. For instance, utilitarians would distribute more to a pleasure wizard than to a cripple, because the pleasure wizard’s marginal utility is higher than the cripple’s: “the cripple would then be doubly worse off [...] both since he gets less utility from the same level of income, and since he will also get less income” (1979, 203).

5] Rawls considers that the index of well-being should comprise primary goods, which are “all-purpose goods” that individuals with diverse conceptions of the good would want in the original position, because “other things equal, they prefer a wider to a narrower liberty and opportunity, and a greater rather than a smaller share of wealth and income (1971, 396).

6] Another variant of resourcism, Dworkin’s equality of resources, would uphold distributions to those who lack physical endowments – what he calls “personal resources”: “Someone who is born with a serious handicap faces his life with what we concede to be fewer resources [...]. This circumstance justifies compensation” (through his favored hypothetical insurance market) (Dworkin 2002, 81). However, Dworkin’s view faces other problems. For some of the most relevant criticisms, see Burley (2004).

7] Rawls writes that “[...] the principles of justice are the result of a fair agreement or bargain” (1971, 12). The idea of agreement has been interpreted differently. One operationalization of the agreement in the original position does not require ontologically distinct parties. Alexander (1974) argued that Rawls’ is not a contract theory, because it could be replaced by an individual choice theory or a sympathetic observer theory at no loss for its substantial claims. Jean Hampton (1980, 337) holds that because behind the veil of ignorance parties are indistinguishable it is better to interpret the deliberation as being the result of the reasoning of a single party.

8] Earlier (Rawls 1971, 12), he mentioned that parties do not know “their place in society, their class position, social status, intelligence, fortune in the distribution of natural assets and abilities, their conceptions of the good or their special psychological propensities”. In his later work he includes comprehensive doctrines since “not allowing the parties to know people’s comprehensive doctrines is one way in which the veil of ignorance is thick as opposed to thin” (Rawls 1999, 31).

The definition advanced by Rawls of society as a “mutually advantageous cooperative venture” has attracted the criticism of feminists such as Young (2006) for ignoring those who are not able-bodied and thus are not contributing parts of the society. One related criticism is that Rawls fails to include in his index care and dependency related primary goods, such as “the understanding that we will be cared for if we become dependent, the support we require if we have to take on the work of caring for a dependent, the assurance that if we become dependent, someone will take on the job of caring for those who are dependent upon us” (Kittay 1999, 102)<sup>9</sup>. Similarly, Nussbaum holds that caregivers are disregarded by Rawls’ theory of justice, a deficiency that her capability approach tries to solve. These considerations are integrated in her own version of the capability approach, which acknowledges that securing capabilities entails a process of designing the material and institutional framework “so that it provides the requisite affirmative support for all the relevant capabilities” (Nussbaum 2003, 51-55). Capabilities are combinations of functionings, which in turn are doings and beings that a person can achieve if she desires (Sen 1992, 40). While capabilities represent the opportunity/freedom aspect, functionings concern the outcome/achievement aspect (Robeyns 2003, 63). In order to differentiate between these two aspects, the capability approach needed to escape the narrow confines of informationally-parsimonious accounts of justice such as utilitarianism. Thus, it could be argued that informational richness and the case for assessing comprehensive outcomes are the main sources of the capabilities approach.

The capability approach is not a full theory of justice and by itself it proposes just an evaluative space – the metric of capabilities. According to Robeyns, the capability approach operates at three levels, being capable of taking various forms according to one’s needs. It can be a framework of thought, a critique of other approaches to well-being assessment and a formula for making interpersonal comparisons of well-being (Robeyns 2000, 3). The fact that the capability approach is “underspecified” can be double-edged: the sensitivity of the approach to gender is dependent on what additional theories are brought to complement it. Robeyns develops a gender-sensitive capability approach, which shows what feminism has to gain from resorting to this conceptual apparatus. She emphasizes 3 elements of the capability approach that make it a good candidate for being complemented by feminist theories, *i.e.* i) its evaluative space, ii) the attention given to human diversity and iii) its employment of ethical individualism. The second of these is of paramount importance for feminism because it can be seen as a way to internalize Okin’s (1989) criticism of theories of justice for their failure to acknowledge the way in which the different experiences of females and women affect their subsequent development.

i) One of the reasons why Sen endorses the capability metric and not a functioning metric is because of what he calls an “informationally inclusive advantage” of the former, since focusing on capabilities allows one to account for the opportunities and choices of the agent, who is not perceived as a simple recipient, but whose interests and engagement

---

9] For a defense of Rawls, see Wong (2010, 127-146).

are taken into account in a proper way (2010, 236). This is not without criticism – Cohen argues that behind the capability approach is a certain “athleticism”, which consists of “overestimating the place of freedom and activity in well-being” (1990, 377)<sup>10</sup>. Sen argues that in assessing well-being freedom, one should look at the capability levels, but concedes that “limits of practicality may force the analysis to be confined to examining the achieved functioning bundle only” (1992, 53). This is nonetheless not desirable from a normative standpoint, and capabilities should be the evaluative space. Of course, going further, well-being is not the only aspect that people should be enabled to reach – agency plays an important role too, as I will show later on. For feminists, the capability metric is more appropriate than a resourcist or a welfarist one because it focuses on people’s doings and beings and thus has something to say on the inequalities within the family. According to Nussbaum (2003, 39) “inequalities in resources and opportunities, educational deprivations, the failure of work to be recognized as work, insults to bodily integrity” are all hindrances that diminish the capability sets of women and thus should be solved. The fact that the capability approach is so sensitive to structural injustices affecting the conversion of bundles of goods into functionings is held to be a powerful argument in favor of this evaluative space (Anderson 2010a, 87).

ii) Robeyns (2000, 6) holds that the capability approach takes account of human diversity in two distinct ways: by focusing on capabilities as the evaluative space (assessing thus the opportunities held by people and not imposing outcome desirability) and by looking at individual and social factors of converting commodities into functionings.<sup>11</sup> Among the social factors affecting the conversion of commodities into functionings is gender discrimination. Furthermore, by looking not only at market processes, but also at what happens in the private households, the capability approach internalizes another criticism raised by feminists against other theories of justice (Robeyns 2008, 89-90). Capabilitarians hold that personal heterogeneities cannot be accounted for by other evaluative spaces, which cannot deal with feminine experiences such as breast-feeding. However, there have been arguments that gender-sensitive versions of resourcism could be developed, which would take into account the needs of the infants. This is closely related to a criticism addressed to the capability approach that it proposes a vertical conception of the human nature – “that it falsely suggests that women’s terrible and disproportionate suffering in most of this world is due to their being insufficiently compensated for their inferior natural endowments” (Pogge 2002, 181-83). One capabilitarian response has been that this criticism is based on the idea that capability theorists try to compensate people for bad luck in the natural allocation of internal resources (Anderson 2010a,

---

10] Arneson (2006, 37) also expresses concern towards the focus on what he calls real option freedom – “in many contexts, adding extra valuable options and thus increasing a person’s real option freedom would come at a cost or lesser achieved well-being, and in these contexts we should favor achieving more functioning. Beyond some point, expansion of real option freedom is wasteful of resources that are better spent targeted at boosting functionings”.

11] Commodities here represent the resources distributed, i.e. income or rights.

95). While this is a powerful argument, it is based on Anderson's democratic egalitarian proposal, which constrains, as I will show, the scope of justice in an illegitimate way.

iii) Robeyns argues that the capability approach is based only on ethical individualism, which postulates that the fundamental unit of human concern is the individual. This does not imply ontological individualism, on which it has nothing to say. It can be complemented by various theories regarding ontology. This is important for those feminists who emphasize the necessity of accounting in a theory of justice for the relations between people or between people and nature (Robeyns 2000, 19). Nussbaum's (2000, 79) list of central human capabilities includes the capability of being able to have a connection with the nature, which should make this view attractive to eco-feminism, for instance.

Having established how the capability approach is responsive to gender, I turn now to the main concern of this paper, the problem of positional objectivity and its impact on feminist agendas. In the next section I will briefly present Sen's arguments for open impartiality, his account of positional objectivity and his endorsement of transpositional objectivity. In order to understand what open impartiality is and why Sen criticizes Rawls' contractarianism, it is necessary to explain the relation between justice as fairness and global justice.

## II. SKEWED IMPARTIALITY, POSITIONAL OBJECTIVITY AND THE BIASES BEHIND PUBLIC REASON

Although Rawls (1971) leaves an element of ambiguity regarding the scope of justice<sup>12</sup>, in later works it becomes clear that justice as fairness is meant to be a political conception of justice for well-ordered peoples. Rawls holds that principles of global justice are to be established following a second original position, where the parties are now representatives of peoples. The veil of ignorance device is once again implemented and this time it excludes from knowledge facts such as the size of territory or of the population, the relative strength of the peoples whom the parties represent, the natural resource endowments, the level of economic development (Rawls 1999, 31-32). The peoples and their representatives are moved by reasonable interests, which is congruent with a fair equality and a due respect for all peoples. The idea of public reason is generalized, since the parties do not invoke principles related to their comprehensive doctrines of truth or right, but in terms that can represent a common ground to all the different peoples taking part in the deliberation. More controversial, Rawls includes decent hierarchical societies along the liberal ones. These are not reasonable, but decent (this is a weaker criterion than reasonability) (Rawls 1999, 83-84). Decent hierarchical societies have a process of public consultation and minorities are encouraged to have a flourishing life, but are not allowed

---

12] The scope of justice refers to "the range of persons who have claims upon and responsibilities to each other arising from considerations of justice" (Abizadeh 2007, 323).

to hold offices (in contradistinction with the fair equality of opportunity part of Rawls' second principle of justice as fairness) (Rawls 1999, 77-78). This aspect is important for the subject discussed here because Rawls employs a "narrow doctrine of human rights" (Macleod 2006, 134-49). In one of his examples, a decent hierarchical society can be one in which women are relegated to a status of second-order citizens, if important reforms have been undertaken in regard to their rights.

Furthermore, Rawls does not uphold a global difference principle, but a weaker duty of "assistance to burdened society" (1999, 37). The fact that Rawls includes decent hierarchical societies and scales down the principles of justice and the requirements of justice in order to appeal to these has been criticized as an inadequate concession. Kok Chor Tan (2004, 75) argues that "the normative individualism fundamental to Rawls' domestic theory is replaced by a communitarianism of a sort that takes societies or peoples to be the basic subject of justice". Tolerating and accepting as justified (although not as fair) decent hierarchical societies leads to ignoring dissenting individuals from those nonliberal societies (Tan 2006, 75).

Now we can begin to understand Sen's arguments against Rawls' *closed impartiality*. According to Sen (2010, 182), limiting the original position to peoples inhabiting a society "extracts a heavy price in the absence of any procedural guarantee that local values will be subjected to an open scrutiny". Sen (2010, 125) advocates Adam Smith's impartial spectator device, which allows one to examine values, convictions and principles "at a distance". Rawls, unlike Smith, limits the deliberation to a "given focal group" which might reflect pre-contractarian biases (Sen 2010, 133). In this light, there are several limitations of the original position device, such as "exclusionary neglect" (people who do not belong in that society are excluded, although they might be affected by the outcomes of the deliberation) or "procedural parochialism" (other worldviews are excluded from considerations, and even the veil of ignorance cannot prevent the decisions from "being swayed only by local group prejudices") (Sen 2010, 139, 150). These are some reasons why closed impartiality is morally flawed and why Sen endorses open impartiality. It is important to note here that Sen's endorsement of open impartiality is partially (and implicitly) based on his work on informational constraints.

Sen (1985, 184) engages in a discussion on objectivity insisting on cases that are position-relative but authorship-invariant. For instance, irrespective of the identity of a poor-sighted professor, she would not know that the student in the last row is cheating. If that student had been in the professor's position, having the same characteristics (poor sightedness), she would hold the same belief. This is what he later calls objective positionality (1992), a concept with implications for decision theory, gender discrimination, morbidity rates assessment and ethics. It is to be distinguished from subjectivity, since it does not depend on psychological processes peculiar to a person or another – anyone being in the same circumstances and being endowed with the same quantity of information

would hold the same view. Later, Sen insists on the relevance of this sense of objectivity, proceeding with a criticism of Nagel's "view from nowhere" – type of objectivity.<sup>13</sup>

Against Nagel, he argues that this view is not "from nowhere, but from a delineated somewhere" (Sen 1993, 127). This concept is important because it shows how beliefs can be *objective*, irrespective of their truth value.<sup>14</sup> What values one has, what position in society she occupies, what knowledge she holds, all these elements influence her objective assessment of the world – Sen calls these features *positional parameters*. In order to avoid the inevitable biases associated with this positional relativity, Sen recommends trans-positional assessments, the aggregation of different positional observations into a whole. Trans-positional assessments transcend the biases of positional parameters by making more information accessible to others and by allowing the public agenda to be influenced by opinions who hitherto have been ignored in an *objective* way (Sen 1992, 6). Pursuing trans-positional objectivity has important implications for feminism. Sen shows that living in a society where women are considered second-order citizens and accordingly paid less than men might contribute to one's holding an objective positional belief that women value less than men. It is an objective assessment because anyone living in that society, having access to the same sources of information, would hold the same opinion. However, the need for trans-positionality is relevant because this is the only way in which dissenting views could be taken into account. Trans-positionality allows criticism to be *internal* and not a form of cultural imperialism (Sen 1992, 6). Women who had access to more sources of information than their male and female counterparts would then be allowed to expose others to these new sources and maybe to alleviate through public deliberation the condition of women. Sen argues that a similar procedure is proposed by Adam Smith, and that the impartial spectator is introduced as a way to obtain such a trans-positional view (Sen 1992, 5-6).

The connections between positional objectivity and the capability approach are easy to understand, especially when one sees through this filter the criticisms raised against desire satisfaction or happiness-based assessments of the quality of life. Sen gives the example of self-reported morbidity rates among Indian women. According to Sen, this reflects not only their lack of education, but also of accepting this state of affairs. Sen (1992, 14) admits that:

by constraining the positional parameters very thoroughly, it would be possible to attribute positional objectivity to the Indian rural women's lack of sense of relative deprivation in health or well-being...on the other hand, this positional objectivity, achieving through extensive constraining, would not readily translate into trans-positional objectivity of women's relative deprivation.

---

13] In which Nagel argues that "the standpoint of political theory is necessarily detached and objective" (1986, 188).

14] "Truth is quite a different issue from the objectivity of the reasoning leading to a particular belief, given the access to information that the person has" (Sen 1993, 127).

Would it be paternalistic or imperialistic to criticize this state of affairs and promote women participation in public debates? Martha Nussbaum (2000) offers many examples of how educated women set out to reform their communities, examples which contradict the thesis that this is how the Indian culture is and any attempt to change the state of affairs would be intrusive. Positional objectivity can and should be transcended in order to leave behind such cultural biases, nonetheless. I end this section by quoting at length Sen's own view on the necessity of trans-positional objectivity, which provides a direct link to the next section as well:

"[...] in denying the objectivity of the belief in women's inferiority one can of course invoke the need for a trans-positional assessment involving international perspectives, drawing on observations and beliefs from vantage points prevailing in other societies where women have more opportunity to show their ability. But the more immediate issue is the non-necessity of taking an establishment view of feminine inferiority even for those living in such a society. Contrary views can be taken consistently with living in such a society, and the critique of that view can be internal" (1993, 139).

### III. WHOSE DEMOCRATIC EQUALITY? NARROW SCOPE, NARROWER CAPABILITY SETS

In this section I try to accomplish three things: 1. to show how Sen's concept of positional objectivity has been applied to deliberative democracy by Elizabeth Anderson; 2. to refute Anderson's capability-based democratic equality on cosmopolitan grounds and 3. to argue for a gender-sensitive and cosmopolitan relational egalitarianism. This section is meant to defend an approach to justice inspired by Sen's conceptualization of informational constraints, positional objectivity and capabilities, with further inputs from global justice theories and democratic equality philosophers. I will briefly present the central line of argumentation in the next section, so that the uninterested reader can skip this section.

Anderson interprets Sen's positional objectivity as a conceptual apparatus that could reduce the tension between woman perspectives and moral objectivity (2003, 239). She argues that we could place evaluative perspectives on a continuum ranging from local to global. Those who employ the global perspective hold that local perspectives are biased, sectarian and inadequate. This is the view adopted, *inter alia*, by Nagel or Rawls. The other end of the continuum is occupied, among others, by feminists, and it advocates perspective pluralism. Each different perspective, being forged in different circumstances, is unique and brings something new to the debate. For good reason, Anderson quotes among those holding this position Sandra Harding (1993, 49-82). Harding proposes a standpoint feminist epistemology that bears resemblance to Sen's justification for open impartiality and trans-positional objectivity. According to Harding, the experiences of those at the bottom of society, disadvantaged on racial, ethnical or gender criteria, should be brought on research agenda. The fact that their experiences, opinions and values have been ignored means that what happens in the public sphere lacks comprehensiveness.

As it accounts for the fact that “one’s social situation enables and sets limits on what one can know”, her proposed standpoint feminist epistemology is based on the premise that basing research on women’s lives will generate “less partial and distorted accounts not only of women’s lives but also of men’s lives and of the whole social order” (Harding 1993, 55-6). Similarly, to Anderson, she contests the view that the only alternatives to “the view from nowhere” are ethnocentrism and relativism (1993, 58). It is interesting to see how a version of epistemology feminism has been developed in parallel to Sen’s work on positional objectivity leading basically to the same conclusions.

Anderson (2003, 242-43) shows, however, that obtaining a trans-positional account can be done on distinct grounds, some with more appealing implications than others. She offers four main such reasons: ascriptive identification (on the basis of gender, ethnicity, race), sympathy, practical identification and respect. Ascriptive identities refer to relations established by birth (kinship, ethnicity, race, caste) or through religious or cultural affiliation and are hopelessly parochial and sectarian, creating conflict between different viewpoints and unable to generate a fully universal evaluative position. For Anderson, attempting to obtain an inclusive trans-positionality is meant to solve some informational constraints of democratic rules. Sen has argued for the inclusion of “procedural considerations in consequential analysis” (1995, 13) – although they may seem efficient, some outcomes cannot be judged irrespective of the procedure that led to their appearance. Democratic aggregative procedures are intrinsically flawed because they tend to ignore minorities’ needs from consideration. Sen gives the example of famines, which would be easily preventable, had people been aware of their occurrence and recurrence in many parts of the world. However, the fact that less than 5% of the globe’s population is stricken by poverty and has to suffer the hardships of famine means that only by public discussion could this problem be solved (Sen 1995, 17). As Anderson puts it, “only a vivid awareness of the feasibility of alternatives inspires dissatisfaction with normal states of chronic deprivation” (2003, 248).<sup>15</sup> Democracy in its aggregative form is weakly equipped to respond to people’s needs. This is why Sen has endorsed a deliberative conception of democracy, with broader informational bases (Anderson 2003, 248-49). Anderson takes Sen’s instrumental arguments for democracy and extends them in order to make these compatible with her democratic egalitarian framework. Anderson holds that for Sen democracy is a universal value in that it reflects, promotes and is the result of practical reason. Anderson provides in this regard the following example: if a society

---

15] She links this with the problem of adaptive preferences. Subjective preferences are introduced by Sen as a criticism of subjective metrics such as welfarism, which are informationally parsimonious and do not take into account the fact that a person may have rationally developed preferences for her deplorable state just because she sees no other feasible alternatives. For an interdisciplinary approach to adaptive preferences, which combines the capability approach with research from behavioral economics, psychology and cognitive dissonance, see Teschl and Comin (2005). See also Peter (2003) for an overview of how Sen has attempted to accommodate the problem of adaptive preferences within social choice (and thus to enlarge the informational scope of the latter domain).

afflicted by adaptive preferences problems such as women's underestimation of their own problems would become more responsive to public demands for gender justice, this would have a transformative effect on individual women's desires: "once women no longer perceive women's lesser access to (healthcare) as normal, they may no longer adapt their desires to this condition" (Anderson 2003, 251). However, raising women's problems on the public agenda should not be based on identity politics strategies. Anderson emphasizes that one of Sen's concerns is that "resigning to separate, mistrustful identities can be a self-fulfilling prophecy, entrenching suboptimal parochial self-understandings" (2003, 255).

The incorporation of the concept of positional objectivity in deliberative democracy implies a cosmopolitan conception. Sympathy and respect are applicable on a global scale. The endpoint of deliberative processes that bring together different positional perspectives is a synthesis that can legitimize the decisions. The global transpositional outcome is meant to "pay due regard to the interests and perspectives of all". Anderson considers that the political significance of positionality is given by its being an epistemological mechanism, not "a matter of parochial solidarity" (Anderson 2003, 255-58). All in all, it seems that Anderson takes further Sen's project by emphasizing the role of public reasoning in improving disadvantaged individuals' capability prospects. However, one problem that has to be solved is what is the scope of justice. If public reasoning takes place within the boundaries of a state, then it fails to account for the perspective of those from distant lands. This narrower scope of justice is suggested by Anderson's relational egalitarianism.

In several papers (1999, 2007, 2010b), Anderson has endorsed what she calls relational egalitarianism or democratic equality, which represents a reaction to the luck egalitarianism of Arneson (1989) or Cohen (1989). According to Anderson, modern egalitarian writings have lost their grip with the realworld problems. The politically oppressed, race, gender, class inequalities, victims of nationalist genocide, slavery and ethnic subordination, all have been more or less ignored by recent egalitarian thought. In their attempt to eliminate the effects of brute luck on distributions, luck egalitarians have forgotten that the purpose of equality is to end oppression and to ensure that people can effectively stand as equals in the community (Anderson 1999, 288). For Anderson, luck egalitarians have put the wrong questions, to which they have later given the wrong answers. Luck egalitarians have searched for states of affairs whose distributions of goods are morally desirable. Relational egalitarians are asking what justice demands of each agent, institutional or individual. Justice, for relational egalitarianism, is constrained by several elements: principles of justice must be feasible so that agent can realistically pursue them (the "ought implies can" maxim), they must be publicly articulable, stable and satisfying a Pareto improvement condition, in that no action can be forbidden if it can advance some people's interests to no one else's loss (Anderson 2010b, 16-18). Luck egalitarianism fails to respect the basic principles of any egalitarian theory: that people be treated with equal respect and equal concern. It would leave faulty drivers to die on the side of the road (*the abandonment of negligent victims objection*). It would leave people who

live in risk-prone areas to incur all costs if they fail to insure, where such an opportunity exists (*problem of geographical discrimination*). It considers that the desire to procreate is an expensive taste, and it has no concern for dependent caretakers (*problem of vulnerability of dependent caretakers*). The only way it can take care of such problems is by becoming as paternalistic as it gets, telling citizens that they lack the ability to conduct their own lives as they should (Anderson 1999, 296-301).

In relational egalitarianism, equality is a notion characterizing a type of social relations between people, instead of being a distribution of non-relational goods. Equality entails not a distributive pattern but reflects the idea that all people are equally moral agents. Everyone should have the power to develop and exercise moral responsibility, to cooperate with others according to some principles of justice, to shape and fulfill a conception of their good (Anderson 1999, 312). To this end, goods have to be distributed according to principles and processes that express respect for all. Being capable of functioning as an equal citizen has to be understood in a very broad sense, from having the ability to exercise specific political rights, to participate in the activities of civil society, in economy. It is linked to being capable of functioning as a human being, which in turn involves effective access to means of subsistence, such as food, shelter, clothing, medical care, further backed by the capability to function as a human *agent* in the proper sense of the term. This involves “knowledge of one’s circumstances and options, the ability to deliberate about means and ends, psychological conditions of autonomy, freedom of thought, of movement, access to education, freedom of occupational choice, the right to receive fair value for one’s labor”. Democratic equality’s purpose is not to ensure effective access to equal levels of functioning, but to guarantee access to levels of functioning sufficient to stand as an equal in society. Anderson considered that democratic equality ought to guarantee effective access to a package of capabilities, and that people have to function as equals over the course of their entire lives (1999, 317-18).

It is unclear if Anderson’s democratic equality is meant to apply just to Western, developed states, or globally. At one point, she mentions that “the point is to identify the demands of justice that flow from citizenship in a democratic state [...] Citizens have a claim to a capability set sufficient to enable them to function as equals in society” (2010a, 83). It is obvious, however, that in order for trans-positional objectivity to be achieved, it is necessary to surpass the confines of the modern nation-state, towards global democratic deliberative processes. Achieving an equal standing in a modern democratic society could still mean that the interests of those far from that society would be ignored. If people were to be concerned over securing a sufficient level of capabilities throughout a democratic state, would they really be concerned to include distant others’ points of view? If people were indeed “entitled to access to a level of functioning in virtue of their citizenship” (2010a, 95), this would be a narrow conception of justice. Such a narrow scope would restrict the capability sets of the distant others. What is needed is to include all viewpoints in a deliberative process that is global in scope. Furthermore, it seems peculiar to resort to what can be called principle asymmetry – upholding an egalitarian principle of justice at

the national level and a sufficientarian conception at the global level. Anderson mentions that, on the one hand, all citizens should function as equals, but at the global level, they should function as *humans* (2010a, 88).

Thus, democratic equality should not be perceived as a distributive view confined to a state, but in cosmopolitan terms. Simply put, cosmopolitanism is based on the idea that each human being has equal moral worth, and that we have certain responsibilities towards all human beings *qua* human beings (Brock and Brighouse 2005, 1-10)<sup>16</sup>. This is in the spirit of Sen's open impartiality, who actively endorses public reasoning at the global level (2010, 151). Ensuring that democratic equality perceives justice's scope as global would be beneficial for those who nowadays are excluded from deliberation processes and whose viewpoints are (objectively) neglected. Fortunately, there is nothing in democratic equality that precludes it from dropping this principle asymmetry and from embracing a global scope.

#### IV. APPLICATIONS

In the previous section I presented Elizabeth Anderson's extension of Sen's concepts to feminist deliberative democracy. This is a view according to which positional objectivity – and its associated misdeeds such as women's neglect of their own problems pertaining to well-being and health – should be transcended towards trans-positional objectivity. Trans-positional objectivity is epistemologically richer because it comprises the values, viewpoints and interests of numerous distinct groups, which are all biased by their own positional parameters. I argued that although Anderson seems to favor a global deliberation, her democratic equality approach can act as a counterweight. Therefore, I argued for the extension of the democratic equality approach to justice at the global level. I have not provided a particular account, though this should be, based on what has been mentioned before a gender-sensitive capability approach. In what follows I will briefly present two distinct applications of the Sen-Anderson work on positional objectivity.

##### *Women, agency and objectivity. Once upon a time in a patriarchy*

The concept of agency is one of the fundamental differences between Sen's and Nussbaum's capability approach (Nussbaum 2000, 13). While Sen distinguishes between agency and well-being, Nussbaum considers that the capability/functioning distinction can accommodate both concepts. Although in some cases this might be true, problematic here can be those cases where agency promotion comes at odds with well-being promotion.<sup>17</sup> These fringe cases are important in the context of the capability

<sup>16</sup> See Dumitru (2017, 234) for some distinctions pertaining to cosmopolitanism and global justice.

<sup>17</sup> It seems as if Sen employs the fox' view, while Nussbaum is more of a hedgehog: "the hedgehog [...] has faith that all true values form an interlocking network, that each of our convictions about what is good or right plays some role in supporting each of our other convictions in each of these domains of value",

approach because Sen himself has emphasized that one way in which a moral theory can be criticized is by the *case-implication* techniques, which sets out to show that in particular circumstances what a normative theory advocates would have counterintuitive results (1979, 197). In this subsection I will show how women's agency can be enhanced. The benefits will accrue especially for women from disadvantaged societies, societies which in turn might benefit from this enriched objectivity.

Sen distinguishes between the well-being and the agency aspects of a person. Each of these can be conceived either in terms of results or of opportunities. Thus, we have four logical possibilities: well-being freedom, well-being achievement, agency freedom, agency achievement. Agency refers to "what the person is free to do and achieve in pursuit of whatever goals or values she regards as important [...] A person's agency aspect cannot be understood without taking note of his or her aims, objectives, allegiances and the person's conception of the good" (Sen 1985, 204). The freedom aspect is emphasized more in the concept of agency, because it is directly related to people being responsible for their condition. Agency and well-being draw attention to different aspects of a person's life – whereas well-being conveys information regarding a person's advantage, the agency aspect provides an assessment of a person's ability to do certain things in accordance with her conception of the good (Sen 1985, 204-206).

As mentioned before, the two aspects can come in direct conflict. Imagine a rich society in which women scored well in terms of well-being freedom and achievement. Take the particular case of Anne, who is a middle-class citizen whose conception of the good gravitates around helping the poor. In order to show that Nussbaum's value-unity endorsement is flawed, let us further suppose that Anne has reached the threshold in regard to all 10 capabilities on Nussbaum's list (2000). For instance, she developed a social network that aims at helping the poor by raising funds (the "Affiliation" capability). She is well-regarded in society, she really cares about the poor (Emotions). She has been able to form her own conception of the good, and she holds a degree in Political Philosophy (Practical Reason). She has Control over her environment, being able to "participate effectively in political choices that govern one's life" (Nussbaum 2000, 80). There is no necessity for trades off between capabilities, because she scores so high in all of them. However, among those subjects that enter the public debate in which she is involved and on which she can vote there is no *redistribution* proposal. Let us suppose that this happens in a parallel universe, where they had no Marx, no Rawls, no Nozick, no Sen, no Nussbaum, no Anderson. There is an implicit bias towards a voluntarist ideology, and free market processes are supposed to solve all poverty-related problems. The poor, who could argue for redistribution, lack the capability Control over one's environment, because they have to struggle with their condition and are not able to engage in political deliberation. Thus, the fact that free market processes are fair is positionally objective, since, had the poor been in the exact same position as Anne or other rich citizens, they would have had the

---

while the fox acknowledges value-conflict (Dworkin 2011, 120; Knight 2015).

same *Weltanschauung*. Nonetheless, Anne feels that something is not right, and she would deeply want to do something more for the poor. Had the poor been helped to engage in public reasoning and to bring their own positional parameters on the agenda, Anne would realize that there is in fact something more that she could do in order to help the poor. Her *agency freedom* would go up, and perhaps, if she convinced enough of her friends to vote for redistributive programs, her *agency achievement* would go up too. Nevertheless, her well-being freedom and her well-being achievement would go down a little. This example shows that there is conflict between agency and well-being. Furthermore, it shows how the concept of positional objectivity can be used in connection with the capability approach.

Imagine now what could happen in today's patriarchal societies if women's voices had been heard and if they had had a word to say in political decisions. Not only would this help those particular women improve their life chances and their well-being freedom and achievement, but many others would have a boost in their agency freedom. Of course, this process might result in losses of well-being freedom, but as long as everybody's capabilities remain above a certain threshold – that I will not specify here – this is not an argument against this reform. In this light, policies to adopt gender quotas can be justified on new grounds. Furthermore, they would not reflect a subsumption/abstraction strategy, but

a pragmatic-epistemic strategy to dealing with positional differences [...] concern for the representation of disadvantaged groups arises not from a desire to reinforce parochial group identities as ends in themselves, but from a desire to construct a more global perspective that can pay due regard to everyone's interests and perspectives (Anderson 2003, 253-55).

Peter (2003, 24) draws attention to another usage of the concept of, in Sen's reconceptualization of social choice. Peter notes that the informational basis provided by agency considerations enriches social choice theory. As a consequence of assessing agency levels, attention is refocused "from problems of aggregation of unexamined individual preferences to participation and inclusion in democratic decision-making". According to Peter (2003, 27), "the challenge for social evaluation of policy alternatives is to register and take seriously the interpretations and evaluations of women as situated agents, thus identifying the means by which their participation in policy discourses can be enhanced and their effectiveness reinforced". These arguments should not be too surprising. Sen's concept of agency has been considered one of his main contributions for gender-equality research programs (Qizilbash 2005).

### *The race, the deliberation and the trans-positional objectivity*

No one bothered to discuss the way in which sexism operates both independently of and simultaneously with racism to oppress us. No other group in America has so had their identity socialized out of existence as have black women. We are rarely recognized as a group separate and distinct from black men, or as a present part of the larger group "women" in this culture. When black people are talked about, sexism militates against the acknowledgment of the interests of black women; when women

are talked about racism militates against a recognition of black women interests. When black people are talked about the focus tends to be on black men, and when women are talked about the focus tends to be on white women. No where is this more evident than in the vast body of feminist literature [...]. All too often in our society, it is assumed that one can know all there is to know about black people by merely hearing the life story and opinions of one black person [...]. From the onset of my involvement with the women's movement I was disturbed by the white women's liberationists' insistence that race and sex were two separate issues [...]. My life experience had shown me that the two issues were inseparable (hooks 1999, 7-12).

There could be employed several filters through which bell hooks' words could be interpreted. I resort now to one of them, which is Sen's (2010, 122) urge to validate ethical claims by subjecting these to public reasoning processes characterized by open impartiality. Sen (2010,123) argued convincingly that in order to achieve such open impartiality we have to invoke judgments from "others, outside our focal group". Similarly, bell hooks holds that the woman liberation movement has been characterized by a lack of concern for the experiences of the black women. This disregard does not come from a malfeasance of white women, but from their positional parameters which disallow certain kinds of information from entering the public debate. Sen (2010, 169) argues that the fact that our perception of the world is inevitably tied to our position in the world affects our understanding and our way of thinking about ethical subjects: "in the pursuit of justice, positional illusions can impose serious barriers that have to be overcome through broadening the informational basis of evaluations". Whereas Smith has proposed a device such as the impartial spectator, Sen advocates transpositionality. We cannot have a view from nowhere, but we can have a view from a plurality of somewhere. Sen uses a great deal of the space in *The Idea of Justice* endorsing the idea that "judgments about justice have to take on board the task of accommodating different kinds of reasons and evaluative concerns" (2010, 395). Although broadening the plurality of justice might lead to incomplete preference orderings, this is not necessarily a problem (2010, 398). The richness of human diversity should not be hidden but taken advantage of.

The implications of Sen's approach to justice can be shown by listing the reasons why he considers social choice to be relevant as a "framework for reasoning": "recognition of the inescapable plurality of competing principles, allowing and facilitating re-examination, permissibility of partial orderings, diverse interpretations and inputs" (Sen 2010, 106-109). These are at odds with approaches that exclude from consideration inputs from social groups whose preference orderings might be different than those of the majority. Furthermore, there are numerous biases related to one's positional parameters that can be overcome only by switching to a trans-positional view.

These are some of the conceptual tools that can be employed to analyze bell hooks' criticisms. Black women's views have been neglected, since the woman liberation movement had incorporated only certain parameters – those inspired by the lives

of white women. Albeit discrimination against white women is not to be ignored, it cannot be considered the only form of discrimination. The absence of black women voices from public debate meant however that their particular experiences had been neglected. By advocating the broadening of positional perspectives, the Sen-Anderson feminist deliberative democracy can internalize bell hooks' criticism. The concern for including in the public debate the voice of neglected categories is of crucial importance for Sen, and this is what provides democratic deliberation with legitimacy (Anderson 2003, 225). Had black women interests, values and opinions been incorporated earlier into the public agenda, trans-positional objectivity would have characterized the woman liberation movement. Instead, what happened was that "the efforts of white women activists to expand employment opportunities for women were focused exclusively on improving the lot of white women workers, who did not identify with black women workers" (hooks 1999, 132). Although bell hooks notes that there had also been a persistent rivalry between white women and black women for access to the job market, this does not mean that a conciliatory relationship would have solved anything. No matter how good their intentions, white women could not have known the positional specificities of black women experiences. In this subsection I have argued for interpreting this criticism of feminism by resorting to Sen's conceptual apparatus. This puts in a new light radical feminists' criticism of "white woman feminism" and shows that it is justified both ethically and epistemologically. Although the white women liberation movement had made objective assessments of the state of affairs, these evaluations had been hopelessly positionally objective, biased by their own positional parameters.

## V. CONCLUSIONS

I advocated in this paper a version of gender-sensitive capability approach which accounts for positional objectivity. I presented the capability approach, insisting on Sen's endorsement for open impartiality and trans-positional objectivity. I followed Sen and Anderson in showing what public reason has to gain from the inclusion of women problems on the agenda. I held that the capability approach can contribute to women's gaining equal standing in society. On the other hand, I have argued for a version of cosmopolitan democratic equality, which satisfies the criterion of open impartiality and goes beyond the confines of a narrow set of denizens of the globe. I have not advanced a specific proposal in this regard, which is a task that I leave for another paper. However, on a more constructive note, I have shown how we can use Sen's research on positional objectivity, agency and social choice in order to accommodate some radical feminist criticisms of mainstream women's movements, specifically regarding bell hooks' diatribe against the lack of concern for the black women's particular situation. I reconceptualized bell hook's arguments in Sen's terms, since her criticisms can be considered to refer to the informational parsimony of mainstream feminist theories.

The stake of the paper, however, is not limited to the applications mentioned above. The concept of positional objectivity can be employed in regard to other issues as well, bridging together feminist theories, political philosophy and social choice.<sup>18</sup>

adelin.dumitru1@gmail.com

#### REFERENCES

- Abizadeh, Arash. 2007. Cooperation, Pervasive Impact and Coercion: On the Scope (Not Site) of Distributive Justice. *Philosophy & Public Affairs* 35 (4): 318-58.
- Alexander, Sidney. 1974. Social Evaluation through Notional Choice. *The Quarterly Journal of Economics* 88 (4): 597-624.
- Anderson, Elizabeth. 1999. What is the Point of Equality? *Ethics* 109 (2): 287-337.
- . 2003. Sen, Ethics and Democracy. *Feminist Economics* 9 (2-3): 239-61.
- . 2007. Fair Opportunity in Education: a Democratic Equality Perspective. *Ethics* 117 (4): 595-622.
- . 2010a. Justifying the Capabilities Approach to Justice. In *Measuring Justice. Primary Goods and Capabilities*, edited by Harry Brighouse and Ingrid Robeyns, 81-101. New York: Cambridge University Press.
- . 2010b. The Fundamental Disagreement Between Luck Egalitarians and Relational Egalitarians. *Canadian Journal of Philosophy* 40 (1): 1-23.
- Arneson, Richard. 2006. Distributive Justice and Basic Capability Equality: Good Enough is Not Good Enough. In *Capabilities Equality. Basic Issues and Problems*, edited by Alexander Kaufman, 17-44. New York and London: Routledge.
- Brock, Gillian and Brighouse, Harry. 2005. Introduction. In *The Political Philosophy of Cosmopolitanism*, edited by Gillian Brock and Harry Brighouse, 1-10. New York: Cambridge University Press.
- Burley, Justine. 2004. *Dworkin and His Critics, with Replies by Dworkin*. Malden: Blackwell Publishing.
- Chor-Tan, Kok. 2004. *Justice Without Borders. Cosmopolitanism, Nationalism and Patriotism*. New York: Cambridge University Press.
- . 2006. The Problem of Decent Peoples. In *Rawls' Law of Peoples. A Realistic Utopia?* edited by Rex Martin and David Reidy, 76-95. Malden: Blackwell Publishing.
- Cohen, G.A. 1989. On the Currency of Egalitarian Justice. *Ethics* 99 (4): 906-944.
- . 1990. Equality of What? On Welfare, Goods and Capabilities. *Louvain Economic Review* 56 (3-4): 357-82.
- Dumitru, Adelin. 2017. On the Moral Irrelevance of a Global Basic Structure: Prospects for a Satisficing Sufficiency Theory of Global Justice. *Croatian Journal of Philosophy* XVII (50): 233-64.
- Dworkin, Ronald. 2002. *Sovereign Virtue. The Theory and Practice of Equality*. Cambridge: Harvard University Press.
- . 2011. *Justice for Hedgehogs*. Cambridge: Harvard University Press.
- Hampton, Jean. 1980. Contracts and Choices: Does Rawls Have a Social Contract Theory? *The Journal of Philosophy* 77 (6): 315-80.

---

18] I thank Mihaela Miroiu and Alexandru Volacu for their helpful comments that significantly improved upon the quality of the paper.

- Harding, Sandra. 1993. Rethinking Standpoint Epistemology: What is Strong Objectivity? In *Feminist Epistemologies*, edited by Linda Alcoff and Elizabeth Porter, 49-82. New York and London: Routledge.
- hooks, bell. 1999. *Ain't I a Woman Black Women and Feminism*. London: Pluto Press.
- Kittay, Eva Feder. 1999. *Love's Labor. Essays on Women, Equality and Dependency*. New York and London: Routledge.
- Knight, Carl. 2015. Justice for Foxes. *Law and Philosophy* 34 (6): 633-59.
- Macleod, Alistair. 2006. Rawls' Narrow Doctrine of Human Rights. In *Rawls' Law of Peoples. A Realistic Utopia?* edited by Rex Martin and David Reidy, 134-49. Malden: Blackwell Publishing.
- Nagel, Thomas. 1986. *The View from Nowhere*. New York: Oxford University Press.
- Nussbaum, Martha. 2000. *Women and Human Development. The Capabilities Approach*. New York: Cambridge University Press.
- . 2003. Capabilities as Fundamental Entitlements: Sen and Social Justice. *Feminist Economics* 9 (2-3): 33-59.
- Okin, Susan Moller. 1989. *Justice, Gender and the Family*. New York: Basic Books.
- Peter, Fabienne. 2003. Gender and the Foundations of Social Choice: the Role of Situated Agency. *Feminist Economics* 9 (2-3): 13-32.
- Pogge, Thomas. 2002. Can the Capability Approach Be Justified? *Philosophical Topics* 30 (2): 167- 228.
- Qizilbash, Mozaffar. 2005. Sen on Freedom and Gender Justice. *Feminist Economics* 11 (2): 151-166.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge: The Belknap Press of Harvard University Press.
- . 1999. *The Law of Peoples, with "The Idea of Public Reason Revisited"*. Cambridge: Harvard University Press.
- . 2001. *Justice as Fairness. A Restatement*. Cambridge: The Belknap Press of Harvard University Press.
- Robeyns, Ingrid. 2000. An Unworkable Idea or a Promising Alternative? Sen's Capability Approach Re-examined. Center for Economic Studies.
- . 2003. Sen's Capability Approach and Gender Inequality: Selecting Relevant Capabilities. *Feminist Economics* 9 (2-3): 33-59.
- . 2008. Sen's Capability Approach and Feminist Concerns. In *The Capability Approach. Concepts, Measures and Applications*, edited by Sabina Alkire, Mozaffar Qizilbash, and Flavio Comim, 82-105. New York: Cambridge University Press.
- Sen, Amartya. 1970. *Collective Choice and Social Welfare*. Amsterdam: Elsevier.
- . 1979. Equality of What? Stanford University: The Tanner Lecture on Human Values: 197-220.
- . 1982. Rights and Agency. *Philosophy & Public Affairs* 11(1): 3-39.
- . 1985. Well-Being, Agency and Freedom: the Dewey Lectures 1984. *The Journal of Philosophy* 82 (4): 169-221.
- . 1992a. *Inequality Reexamined*. New York: Oxford University Press.
- . 1992b. *Objectivity and Position*. University of Kansas: The Lindley Lecture.
- . 1993. Positional Objectivity. *Philosophy & Public Affairs* 22 (2): 126-46.
- . 1995. Rationality and Social Choice. *The American Economic Review* 85 (1): 1-24.
- . 2010. *The Idea of Justice*. London: Penguin Books.

- Teschl, Miriam, and Flavio Comim. 2005. Adaptive Preferences and Capabilities: Some Preliminary Conceptual Explorations. *Review of Social Economy* LXIII (2): 229-47.
- Wong, Sophia Isako. 2010. Duties of Justice to Citizens with Cognitive Disabilities. In *Cognitive Disability and Its Challenge to Moral Philosophy*, edited by Eva Feder Kittay and Licia Carlson, 127-46. West Sussex: Wiley-Blackwell.
- Young, Iris Marion. 2006. Taking the Basic Structure Seriously. *Perspectives on Politics* 4 (1): 91-97.

# PUBLIC REASON

## Journal of Political and Moral Philosophy

**Submission information:** All contributions whether articles or reviews are welcome and should be sent by e-mail as attachment at [publicreason@ub-filosofie.ro](mailto:publicreason@ub-filosofie.ro). Papers will be refereed on a blind basis by *Public Reason's* referee board. Acceptance notices will be sent as soon as possible. The editors may ask for revisions of accepted manuscripts. Papers should have an abstract and key words. Please limit your paper submission to 8000 words, including footnotes and references, and format it for blind review (the text should be free of personal and institutional information). Along with the article, but not in the article, please send contact details: current affiliation, address, telephone number and email address. Authors are responsible for reinserting self identifying citations and references when manuscripts are prepared for final submission. Book reviews should have no more than 4000 words.

**Manuscript preparation:** All manuscripts should be formatted in Rich Text Format file (\*.rtf) or Microsoft Word document (\*.doc) with 12 pt. font size and double-spaced. Please use the author-date system documentation style, as it is presented in *The Chicago Manual of Style*. For more information about our editing rules, please go to: <http://www.publicreason.ro/submission-info>

**Notes:** The contributions have to be original and not published before. *Public Reason* does not require exclusivity in submission of articles. However, if you have submitted your article to another journal, we kindly require that you let us know about it, and notify us without any delay about other journals' decisions about your article. The work has to be approved by all co-authors. Authors wishing to use text passages, figures, tables etc. that have already been published elsewhere are required to obtain permission from the copyright owner(s). Evidence that permissions have been granted are to be sent along with the manuscripts submitted for the blind review. The authors assume all responsibility for the content published in the journal.

**Book reviews:** Books for review should be sent to: CSRB, Department of Philosophy, University of Bucharest, 204 Splaiul Independentei, Sect. 6, 060024, Bucharest, Romania.

**Advertising Information:** If you want to advertise your books, department programs, conferences, events etc. please contact us at [publicreason@ub-filosofie.ro](mailto:publicreason@ub-filosofie.ro) in order to get general information or information on costs, specifications, and deadlines.

**Subscription and delivery:** *Public Reason* is an open access e-journal. For the print version of the journal please contact us at [publicreason@ub-filosofie.ro](mailto:publicreason@ub-filosofie.ro).

**Permission:** *Public Reason* holds the copyright of all published materials. Requests for permission to reprint material from *Public Reason* should be sent to [publicreason@ub-filosofie.ro](mailto:publicreason@ub-filosofie.ro).

**Online publication:** *Public Reason* is available online at <http://publicreason.ro>

The journal *Public Reason* is published by *The Center for the Study of Rationality and Beliefs*. CSRB is a unit of scientific research within the Faculty of Philosophy of the University of Bucharest (<http://www.ub-filosofie.ro>). CSRB is located in the Faculty of Philosophy, Splaiul Independenței 204, Sector 6, postcode 060024, Bucharest. For more information see <http://www.csrc.ro/EN/home>

ISSN 2065-7285

EISSN 2065-8958

© 2018 by *Public Reason*

## ARTICLES

Julian Katz

Kant's 'Bund': A Voluntary Reading

Leonidas Makris

Mill's Political Perception of Liberty: Idiosyncratic, Perfectionist but essentially Liberal

Alexandros Manolatos

Rawlsian Stability and the Hazards of Envy

Jens Peter Brune

Capabiltarianism without Paternalism?

Hossein Dabbagh

The Problem of Explanation and Reason-Giving Account  
of pro tanto Duties in the Rossian Ethical Framework

Adelin Costin Dumitru

Feminism, Agency and Objectivity