PUBLIC REASON

Journal of Political and Moral Philosophy

Volume 3, Number 1, June 2011

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.

Public Reason is available online at **http://publicreason.ro**
ISSN 2065-7285
EISSN 2065-8958
Print version published by **Editura Universității din București** for *Public Reason*.
Comanda Nr. . Tipografia E.U.B.

PUBLIC REASON

Journal of Political and Moral Philosophy

Vol. 3, No. 1, June 2011

ARTICLES

BOOK REVIEWS

Public Reason 3 (1): 3-21

The Mutual Dependence of Institutions and Citizens' Dispositions in Liberal Democracies

Jeremy Neill
Houston Baptist University

Abstract. In recent centuries liberal democracies have amassed a remarkable record of performative excellence. But contemporary statistical evidence points to an ongoing and startling decline in the civic involvement of their citizens. In this article I use the problem of civic nonparticipation to unpack the interaction of institutions and the dispositions of citizens in liberal democracies. Currently the approach that political philosophers take to the civic nonparticipation problem is dominated by an institutional priority view. I develop a mutual reinforcement alternative to the institutional priority view. By mutual reinforcement I mean the idea that the successful operation of institutions and the dispositions of citizens in liberal democracies tend naturally to reinforce each other's orientation toward liberal-democratic ends. I conclude the article by noting that mutual reinforcement provides us with a more productive way of thinking about issues of civic participation than institutional priority.

Key words: institutions, dispositions, liberal, citizens, democracy.

"The basic structure is the primary subject of justice
because its effects are so profound and present from the start." (Rawls 1999, 7)

In recent centuries liberal democracies have amassed a remarkable record of performative excellence. They are on many accounts the most stable and popular forms of modern governance. But contemporary statistical evidence points to an ongoing and startling decline in the civic involvement of their citizens. Today the citizens of liberal democracies are participating in fewer civic associations, engaging in fewer public discourses about their fundamental political and moral values, and are ordering their daily lives with less and less frequency around the conviction that their civic participation contributes to the construction of a productive social order (Putnam 2000, 31-47, 277-84).[1] Citizens do not generally enjoy public involvement, do not believe that public involvement is valuable to their personal formation as human beings, and do not think that they personally can make a political difference in countries of so many millions.[2]

Political philosophers are generally agreed that the increasing rate of civic nonparticipation is a problematic trend (Putnam 2000, 402-14). The reasons that they have given for why it is problematic are numerous and varied. One reason is that civic nonparticipa-

1] Putnam's analysis concentrates on America; see also Elshtain 1995, 1-36; and Dagger 1997, 132-53. Similar trends have appeared outside of America, so it is appropriate to apply my comments in this article to liberal democracies in general. See Putnam 2002.

2] Among the practical political activities from which citizens are increasingly abstaining are voting in general elections, personally holding or running for political office, and participating in the operation of political parties. See Galston 2007, 631-34; 2004, 263-66; and 2001, 217-34.

tion causes citizens to lose track of their interests and rights – non-involved citizens understand significantly less about the importance of public events for their own views and well-being than do their more civically active counterparts. Another reason is that civic nonparticipation causes the political views of citizens to grow chaotic and uninformed. Non-involved citizens are less capable than their more active counterparts of productively absorbing political theories and developments into their overall worldviews; they also tend to be more distrustful and estranged than is warranted by the political situation. Still another downside of civic nonparticipation is that it deprives the public sphere of insightful viewpoints and diminishes the overall functional effectiveness of liberal democracies (Galston 2007, 636-38).

The various practical explanations that political philosophers have offered for the rise of civic nonparticipation have included such diverse things as generational changes, increasing occupational responsibilities, urban and suburban sprawl, isolated lifestyle patterns, and the development of television, mass entertainment, and leisured activities (Putnam 2000, 277-84). Certainly explanations which concentrate on practical social factors are productive initial approaches to the issue. But behind and beneath such explanations are deeper categories which shape our views of civic participation.

In this article I will use the problem of civic nonparticipation to unpack the interaction of institutions and the dispositions of citizens in liberal democracies. Currently the approach that political philosophers take to the civic nonparticipation problem is dominated by an institutional priority view. On this view institutional establishment ought to precede the development of the dispositions of citizens in the construction and analysis of liberal democracies.[3] As a first step in my argument, in sections one and two, I will expose the limitations of the institutional priority view by applying John Rawls' formulation of it to the civic nonparticipation problem.

As a second step in my argument I will develop in section three a mutual reinforcement alternative to the institutional priority view. By mutual reinforcement I mean the idea that the successful operation of institutions and the dispositions of citizens in liberal democracies tend naturally to reinforce each other's orientation toward liberal-democratic ends. I will conclude the article by applying the mutual reinforcement view to the civic nonparticipation problem and by noting that mutual reinforcement provides us with a more productive way of thinking about issues of civic participation than institutional priority. My overall thesis is that prudence entails an approach to liberal democracies that views institutions and dispositions as interacting in a way that is reciprocally determining and that meets the concerns of civic defectors. Throughout the article I will work at two different levels of inquiry – at an abstracted and idealized level, and at a practical and

3] Putnam implicitly views civic non-involvement from such a perspective when he lauds institutional initiatives like community libraries, coastal shipyards, and evangelical megachurches. (2003, 34-54, 55-74, 119-41). Others more explicitly view the cooperation of citizens as being influenced by government institutions, and as dependent on institutional engineering (see Rothstein and Stolle 2003, 191-210; Husseune 2003, 211-30; Hall 1999, 417-64; and Tarrow 1996, 389-97).

realistic level. I will argue that such a two-tiered investigation is justifiable and that the ideal-level commitments of philosophers in fact play a significant role in shaping their approach to real-world issues.

I. THE RAWLSIAN INSTITUTIONAL PRIORITY VIEW

John Rawls is the most prominent contemporary proponent of the institutional priority view. Rawls' arguments for the priority of institutions (or as he calls it, the 'basic structure') are made in the context of his discussion of the fundamental principles of justice: "For us the primary subject of justice is the basic structure of society, or more exactly, the way in which the major social institutions distribute fundamental rights and duties and determine the division of advantages from social cooperation." (2000, 7)[4] In justice as fairness the basic structure is depicted as a framework for the determination and preservation of the fundamental principles of justice.

As is evident from a listing of the social institutions that Rawls includes in the basic structure, the normative stipulations of the basic structure do not pertain to the everyday choices of citizens (1999, 8).[5] Rather, the basic structure encompasses only general institutions like "the legal protection of freedom of thought and liberty of conscience, competitive markets, private property in the means of production, and the monogamous family […]" (6). Rawls' intention in prioritizing the basic structure is to concentrate first on explicating the part of society that he identifies as the most fundamental, influential, and official.

Rawls views himself as prioritizing the basic structure in a sequentially primary sense: "on this view, a theory must develop principles for the relevant subjects step by step in some appropriate sequence." (1993, 258)[6] An established basic structure which sets as fixed a particular understanding of the classical questions of social justice is and ought to be the initial sequential consideration in the construction of the well-ordered society. On my interpretation Rawls intends at least three things by his sequential prioritization of the basic structure: first, to assert that the construction of the formal and fixed core of the well-ordered society ought chronologically to precede the construction of the rest of society; second, to establish the basic structure as a creative and conceptual source for the

4] For some of Rawls' other articulations of the priority of the basic structure cf. 1977, 159-65; 1978a, 48-52; 1978b, 47-71; 2005, 257-88; and 2001, 52-57.

5] A lengthy discussion of the scope of the basic structure has descended from G.A. Cohen's critique of the sustainability of justice as fairness in the absence of a just social ethos. See especially Cohen, 1997, 11, 17-19; and Murphy 1999: 251-91.

6] Elsewhere in the same work Rawls characterizes what he is doing as "starting with the basic structure and then developing other principles sequentially." (259-60); still elsewhere in the same work Rawls characterizes the idea underlying his sequential prioritization when he says that "It may be possible to find an appropriate sequence of kinds of subjects and to suppose that the parties to a social contract are to proceed through this sequence with the understanding that the principles of each later agreement are to be subordinate to those of all earlier agreements […]" (262).

development of the informal conventions and mores of the society; and third, to suggest that the construction of a just basic structure is a necessary condition for the health of the society. I will in what follows attempt to complicate Rawls' prioritization of the basic structure, in light of this three-part interpretation of his intentions.

II. RESPONSE TO THE RAWLSIAN INSTITUTIONAL PRIORITY VIEW

Rawls' prioritization of the basic structure has provided contemporary political philosophers with an immensely productive way of thinking about the infrastructures of liberal democracies. As such it might be thought that Rawls has equipped his theory with tools which are sufficient to resolve the civic nonparticipation problem. But in reality justice as fairness is not particularly well-prepared to approach high-percentage civic defection. In this section I will sketch a background for this claim by investigating the difficulties in Rawls' prioritization of the basic structure.

An inspection of the considered judgments of persons, which Rawls makes one of the key conceptual sources for the decision-making conditions of the original position, is sufficient to demonstrate that Rawls is not prioritizing the basic structure in the sequentially primary sense in which his comments on the subject seem to suggest. Rawls represents the original position as a reflective equilibrium between, on the one hand, our widely-held considered judgments about the appropriate way to construct contractual circumstances, and, on the other hand, the principled grounds upon which social contracts are characteristically built.[7] So at least half of the theoretical support for the basic structure comes from our reflective opinions about the kinds of limitations that are appropriate to place on decision-making situations.[8] For Rawls a reflective equilibrium is reached when the principled and historical understanding of the contract coincides with persons' agreement about the appropriate way to go about social decision-making.

How widespread is the general agreement which Rawls references when he constructs the original position? Close inspection, I think, demonstrates that the agreement is not as broad as might be thought from the descriptive claims that Rawls makes about it. In fact, it turns out not to be a general agreement at all, but rather an agreement among persons whose dispositions have already been shaped by a particular social tradition. Consider as evidence for this claim the understandings of freedom and equality that inform the construction of Rawls' original position. Rawls depicts the parties in the original position as being free to select principles of justice in the absence of externally-given

7] "It [viz. the original position] represents the attempt to accommodate within one scheme both reasonable philosophical conditions on principles as well as our considered judgments of justice." (Rawls 1999, 18-19; bracketed text is mine)

8] "I assume, for one thing, that there is a broad measure of agreement that principles of justice should be chosen under certain conditions. To justify a particular description of the initial situation one shows that it incorporates these commonly shared presumptions." (Rawls 1999, 16)

ends.[9] He depicts them also as possessing equality because they are moral persons "having a conception of their good and capable of a sense of justice" (1999, 17).[10]

Such views of freedom and equality arise from the heart of the liberal-democratic social tradition. That tradition, broadly construed, defines freedom as the ability to think or act in the absence of externally-given hindrances and ends (rather than, say, the ability to perform complex functions after having undergone a course of cognitive or moral development, or the ability to select a course of ethical action in accordance with the will of God).[11] It defines equality as the like status that all persons possess as persons, rational agents, or moral beings (rather than, say, a like status held only by persons who possess a certain natural excellence, or who are nobly born, or who manifest other intrinsic characteristics that set them apart from others).[12] So there are good reasons to think that the views of freedom and equality that inform Rawls' construction of the original position are primarily or even exclusively associated with the liberal-democratic tradition. Rawls himself confirms such an observation when he describes his project as offering "a conception of justice which generalizes and carries to a higher level of abstraction" (1999, 10) the staple commitments of key liberal-democratic thinkers..

It is not in itself controversial to note the embeddedness of the original position in the liberal-democratic tradition – many other Rawls observers have said as much (and Rawls himself obviously recognizes it). Yet in spite of the overall familiarity of the claim, there is an entailment that flows out of the claim that has perhaps *not* been sufficiently emphasized by Rawls interpreters, at least not in the context of the civic participation of citizens. This entailment is as follows: once Rawls' assertions about his sequential prioritization of the basic structure are considered in light of the embeddedness claim, it becomes evident that Rawls is not in reality prioritizing the basic structure in the sense in which we characteristically conceive of something as being 'sequentially' primary (i.e. the chronological, creative, and conditional senses that I have tried to capture with my three interpretations of sequential priority). In reality what Rawls is sequentially prioritizing are the norms of liberal-democratic contractual establishment and the dispositions of liberal-democratic persons. Contractual and dispositional considerations are not the first aspects of justice as fairness to which Rawls devotes his reflective attention, but they *are* the chronologically first considerations to inform Rawls' construction of justice as fairness. They are a

9] "I assume that the parties view themselves as *free* persons who have fundamental aims and interests […] *free* persons conceive of themselves as beings who can revise and alter their final end […] they […] have final ends that they are in principle *free* to pursue or to reject […]" (Rawls 1999, 131-32; italics are mine).

10] Elsewhere Rawls says that "It seems reasonable to suppose that the parties in the original position are equal […]. [By this is meant] equality between human beings as moral persons, as creatures having a conception of their good and capable of a sense of justice. The basis of equality is taken to be similarity in these two respects." (Rawls 1999, 17; bracketed text is mine)

11] See Mill 1989 [1859], 71; see also Berlin 2002, 166-218.

12] See for a general account of this kind of equality Dworkin 1981, 185-246.

pre-reflective body of knowledge for Rawls' claim that an abstracted hypothetical situation constrained by a veil of ignorance is an appropriate location for persons to conduct fundamental social negotiations.

There are two further things that I mean besides chronological, creative, and conditional precedence when I say that the dispositions of liberal-democratic persons are ultimately sequentially prioritized in the construction of justice as fairness. The first thing is that the dispositions of liberal-democratic persons are a pre-existing *social validation* of the original position.[13] The opinions of such persons are a pre-existing court of informed reflection which views the proceedings of a social contract as a valid means of collective decision-making. Prior to Rawls' development of justice as fairness, persons in the liberal-democratic tradition had already come to believe in the validity of a contractual arrangement among free and equal rational agents as an appropriate way to determine fundamental social principles (even if the contractual tradition had temporarily fallen dormant in the twentieth century). Rawls acknowledges this and designs the original position to tap into it as a pre-existing idea.

The second thing that I mean is that the dispositions and opinions of liberal-democratic persons are a pre-existing *imaginative source* for the conceptual material that Rawls includes in justice as fairness. The idea of using a contractual device to make social decisions has descended to us from countless attempts by past liberal-democratic philosophers to develop similar decision-making locations. Perhaps the philosophers of the past were not as successful at devising contractual theories which captured all of the goods and ideals that are captured by the original position, but certainly their efforts laid the creative groundwork for its view of the operation of contractual circumstances. Thus the opinions and dispositions of the participants in the liberal-democratic tradition function as an implicit but sequentially-prioritized imaginative basis out of which Rawls develops the original position.

If the opinions of members of the liberal-democratic tradition are ultimately sequentially prioritized in Rawls' construction of the basic structure, and if the basic structure is *not* in reality sequentially prioritized, what is Rawls doing when he says that it is and makes other, similar claims like "the basic structure is the first subject of justice" (1993, 257) and "it is perfectly legitimate at first to restrict inquiry to the basic structure?" (1977, 159) The answer, I think, is that the sense of prioritization that Rawls intends is more appropriately identified as a *methodological* rather than a sequential prioritization. The basic structure is prioritized as a technical means to the development of the well-ordered society: it is the first aspect of the social construction process upon which Rawls trains his intensive and reflective creative attention; the first aspect for which he finds it necessary to invoke complex and creative philosophical categories; and the first aspect in the explication of which

13] I am talking about *social* validity here and not about validity in the normative sense in which Sandel talks about it when he argues that the justification of contracts is a complicated interweaving of procedure and principle. See Sandel 1982, 119.

he engages in a dialogue with philosophical interlocutors. The difference between Rawls' sequential prioritization of the liberal-democratic tradition and his methodological prioritization of the basic structure is approximately the difference between an assumption and a procedure: an assumption is prioritized in the sense that it is the knowledge or justificatory condition without which a theory is incapable of being formulated, whereas a procedure is prioritized in the sense that it is the technical means by which the practical elements of a theory unfold. Both assumptions and procedures are important players in the development of a theory, but an assumption serves as a chronological, creative, and conditional basis of a theory in a way that a procedure does not.

There are two reasons why it is important to note that Rawls' prioritization of the basic structure is in reality a methodological and not a sequential prioritization. The first is that the dispositions of real-world persons play a far more prominent role in the construction of the original position than might initially be thought. Through the original position the dispositions of real-world persons influence Rawls' determination of the principles of justice and his subsequent explication of the well-ordered society. This identification of the remote causal role of the dispositions of real-world persons demonstrates to us that there is space in the initial stages of justice as fairness for the development of a more nuanced and reciprocal view of the interaction of institutions and dispositions in liberal democracies.

The second thing that is entailed by my argument concerns the application of justice as fairness to issues of civic participation. The argument establishes a framework from the perspective of which it is possible to show two things: first, that justice as fairness in its current form is not productively applicable to the civic nonparticipation problem; and second, that the assumptions that underlie justice as fairness point us toward a view of the interaction of institutions and dispositions that ultimately *is* productively applicable to the civic nonparticipation problem. I will unpack the details of this somewhat technical argument in the next section.

III. JUSTICE AS FAIRNESS AND THE CIVIC NONPARTICIPATION PROBLEM

The application of justice as fairness to the civic nonparticipation problem is an enterprise that becomes perceptible to a Rawlsian when it is located within the Rawlsian's familiar conceptual categories. The most pertinent of these categories is Rawls' distinction between ideal and real political theory. Ideal political theory as Rawls defines it is predicated on the assumption that citizens are in full compliance with the principles of cooperation: "Everyone is presumed to act justly and to do his part in upholding just institutions." (1999, 8) Real political theory, by contrast, is predicated on the assumption that citizens are only partially compliant with such principles. Rawls identifies justice as fairness as an ideal rather than real political theory, and says that "[T]he reason for beginning with ideal theory is that it provides, I believe, the only basis for the systematic grasp

of these more pressing problems [of real theory]." (1999, 8; bracketed text is mine)[14] He is, of course, fully aware that there are no real-world political regimes whose citizens are universally compliant with the principles of cooperation in the way that is assumed by justice as fairness.

The civic nonparticipation problem concerns the partial compliance of citizens and occurs at the level of real rather than ideal political theory. As such, any discussion of how a Rawlsian would approach it is speculative. Nevertheless, there are reasons for thinking that such an enterprise would be profitable and that a Rawlsian would be comfortable engaging in it. The appropriateness of applying justice as fairness to the civic nonparticipation problem is evident from a consideration of the original position as a device of representation. Rawls intends for the original position to model real-world social conditions and to yield results that influence the further development of such conditions.[15]

The kinds of knowledge that are available to participants in the original position are generalized rather than particular or concrete. The conditions that inform the decision-making conditions of the original position are described by Rawls as the circumstances of justice: "As far as possible, then, the only particular facts which the parties know is that their society is subject to the circumstances of justice and whatever this implies." (1999, 119) Central to the circumstances of justice is the idea that real-world social decisions are made in conditions of interpersonal conflict and moderate scarcity.[16] The circumstances also model the idea that the motives of participants are self-interested: "The intention is to model men's conduct and motives in cases where questions of justice arise." (112) The fact that Rawls intends for the original position to be a device of representation for real-world social conditions does not in any way diminish its hypothetical status. The circumstances of justice are distinguishable from the more particularized forms of knowledge that are available to persons who are situated in such real-world social conditions. Among the kinds of knowledge that the circumstances do not model are the comprehensive world-views of citizens, the natural distribution of talents and advantages, hereditary legacies, and the availability of natural resources (118). Yet the fact that such particularized kinds of knowledge are not intended to inform the original position does not belie the status of the original position as a model that is drawn from real-world conditions and that is applicable to the civic nonparticipation problem (104).

On a stronger view, the application of the justice as fairness to the civic nonparticipation problem is an obligatory enterprise and not merely an appropriate one. As Rawlsian observer Liam Murphy has noted, an ideal-level political theory like justice as fairness ought to have acceptable implications both for the level at which it is conceived and also

14] For Rawls on ideal theory, see in the same work pp. 8-9, 245-46.

15] "It is natural to ask why, if this agreement is never actually entered into, we should take any interest in these principles, moral or otherwise. The answer is that the conditions embodied in the description of the original position are ones that we do in fact accept." (1999, 19)

16] For Rawls, "the circumstances of justice obtain whenever persons put forward conflicting claims to the division of social advantages under conditions of moderate scarcity." (1999,110)

for the non-ideal level as well (at the very least, it ought not to have demonstrably unacceptable implications for the non-ideal level) (1999, 279). None of us will ever in practice experience the political conditions of ideal theory. Our own actions and the actions of our neighbors will always to some extent resist the prevailing norms of our society. So any ideal-level political theory that we intend to accept as providing us with real-world normative guidance ought to be capable of illuminating political issues which assume only partial compliance (278-79). On such a view it is obligatory to apply an ideal-level theory like justice as fairness to the civic nonparticipation problem because it is paramount to expect from ideal-level theories a certain amount of routing assistance in the determination of real-world courses of action.

The most likely conceptual tool with which a Rawlsian would tackle the civic nonparticipation problem is the civic education plan that is developed in the third part of *Theory of Justice*.[17] The purpose of this plan is to develop in persons a reflective attachment to the principles of justice (Rawls 1999, 404). The plan elevates persons from a natural state of moral immaturity (a state in which they are unaware of the reasons why they are compelled to abide by moral principles) to a state of moral maturity (a state in which they are capable of reflectively endorsing the principles of justice for reasons that are wholly their own). Its core idea is that it is natural for a just basic structure to develop corresponding dispositions in persons: "we acquire a desire to act justly when we have lived under and benefited from just institutions." (Rawls 1999, 399) There are three stages in the plan through which persons must ascend on their way to moral maturity: At the first stage, parents and other social authorities instill in children the morality of authority. Children at this stage obey without comprehension the ethical injunctions of their parents and social authorities. At the second stage, social institutions instill in citizens the morality of association. Persons at this stage abide by the norms of their social associations and derive their identities as moral persons from these associations. At the plan's third stage, persons acquire the morality of principle – they come to understand the purposes of the principles of justice and endorse the principles for their own sake. The end-result of the civic education plan is a state of congruence between the dispositions of persons and the aims of the well-ordered society. The likely way in which a Rawlsian would apply the plan to the civic nonparticipation problem would be to support the development of a curriculum that is informed by the plan and to argue for the effectiveness of such a curriculum in promoting the participation of citizens in the public sphere.

Rawls' description of moral maturation in the well-ordered society is an elegant and inspirational ideal which is accurate in many ways to the social conditions of contemporary liberal democracies. Yet the ability of the Rawlsian education plan to be productively applied to the civic nonparticipation problem is diminished by the fact that the plan is

17] The plan is likely to be used because it is Rawls' attempt to connect the right with the good, institutions with dispositions, and the basic structure with the general expansion of the well-ordered society. See Rawls 1999, 405-19; Note that everything that I say in this argument is consistent with what Rawls says in *Political Liberalism*.

an instrument of a mono-directional view of the procession of transformative influence in liberal democracies.[18] According to this view the current of transformative influence in liberal democracies can and ought to flow in one direction only – away from the institutional infrastructure. Rawls takes this view when he depicts the basic structure as influencing the development of the general society in its regulation of the informal ethical and economic interaction of citizens, its influence on the social relations of families and friends, its shaping effect upon public discussions, and most of all in its alteration of the dispositions of citizens through the instrument of Rawls' civic education plan.

On the mono-directional view that informs Rawls' theory of justice the reciprocal influence of citizens upon the basic structure ought to be minimal and non-transformative. The only significant reciprocal impact that Rawls allows citizens to exert upon the basic structure is stabilization.[19] The four-stage practical sequence that Rawls employs to implement the principles of justice in society does provide a rudimentary sketch of the influence of citizens upon constitutional construction, practical legislative initiatives, and judicial rulings (1999, 171-76). But Rawls certainly does not envision the efforts of citizens as bringing about transformative changes in the theoretical framework that informs the basic structure or in the nature of the practical sequence by which the principles of justice are socially implemented. Nor, for that matter, does he really want the efforts of citizens to impact such fundamental elements of social construction, since citizens are situated in an information-rich social location that compromises their ability to impartially select principles of justice. Thus while the Rawlsian attempt to lay out the institutional infrastructure beforehand has certain advantages, the ability to involve citizens in the determination of their fundamental political values is not one of them.

The difficulty of applying Rawls' educational plan (or, for that matter, any other aspect of justice as fairness that is influenced by the mono-directional view) to the civic nonparticipation problem is that the plan does not offer its participants the opportunity to contribute to the further development of their fundamental political values. There is no creative reward for persons who have ascended through the stages of the plan and have placed themselves in a position to help in the ongoing evolution of the norms of the liberal-democratic tradition. Consider again the standard forms of belief and reasoning that are employed by contemporary civic defectors.[20] Central to all of these beliefs is the idea

18] The mono-directional view, as I interpret it, permeates all of the aspects of justice as fairness that develop out of the principles of justice. So other aspects of justice as fairness would be similarly prevented from being productively applied to the civic nonparticipation problem.

19] "Since a well-ordered society endures over time, its conception of justice is presumably stable: that is, when institutions are just those taking part in these arrangements acquire the corresponding sense of justice and desire to do their part in maintaining them." (Rawls 1999. 398) For Rawls the well-ordered society is permanently stable once its citizens have absorbed "…a sense of justice or a concern for those who would be disadvantaged by their defection […]." (435); Rawls' discussion of stability occurs in the same work on pp. 434-41.

20] To repeat, the beliefs that civic defectors acknowledge as fueling their abandonment of the public sphere are (1) that their individual political efforts are incapable of making a difference, (2) that their

that public participation ought to be a momentous and fulfilling undertaking. The opportunity to participate in the determination of one's fundamental political values would be momentous and fulfilling indeed.

But the Rawlsian view is unable to promise civic defectors such an opportunity, since its mono-directional understanding of the flow of transformative influence in liberal democracies is committed to the idea that fundamental political values are and ought to be beyond the reach of citizens. The Rawlsian view depicts the participation of citizens in the public sphere as being a matter of determining the appropriate practical implementation of already-selected principles of justice. At most what it is capable of offering civic defectors is the opportunity to determine the practical components of a preconceived ideal. Citizens who are denied the chance to determine their fundamental political values are denied a powerful reason for participating in public life. The overall meaningfulness of their participation in the public sphere is diminished, as is the significance of their immediate practical political activities (i.e. their legislative initiatives, judicial participation, public deliberation, etc.). So the inability of the Rawlsian view to offer citizens the opportunity to determine their fundamental political values provides them with more substantive grounds for defecting from the public sphere than would otherwise be the case.

The upshot is that justice as fairness in its current form diminishes the meaningfulness of participation in the public sphere. Its ability to be productively applied to the civic nonparticipation problem is undermined by its allegiance to a mono-directional view of the flow of conceptual influence in liberal democracies. The view is incapable of providing civic defectors with the fundamental goods they are seeking to achieve through public participation.

But the inability of justice as fairness in its current form to provide civic defectors with convincing reasons for participating in the liberal-democratic public sphere does not mean that it is in principle unable to serve as a template for a more appropriate depiction of the interaction of institutions and dispositions in liberal democracies. On further inspection justice as fairness actually contains within it conceptual antecedents for a more productive response to the reasoning patterns of contemporary civic defectors.

IV. MUTUAL REINFORCEMENT AS AN ALTERNATIVE TO INSTITUTIONAL PRIORITY

My claim that the opinions and dispositions of persons are in fact sequentially prioritized in Rawls' construction of the basic structure does not mean that I think that dispositions ought in principle to be prioritized over institutions in the construction and operation of liberal democracies. A closer inspection of the assumptions that inform the construction of justice as fairness demonstrates that Rawls' sequential prioritization of dispositions is really just a single link in a complex chain of reciprocal influence that in-

---

participation in the public sphere is no longer a meaningful enterprise, and (3) that the public sphere would function just as effectively without them.

stitutions and dispositions exert upon each other in the unfolding of liberal-democratic sentiments over time. The historical reality is that the development of the dispositions of the liberal-democratic persons to whose opinions Rawls appeals has been in large part a product of the influence of an institutional infrastructure. For centuries before Rawls made this appeal, liberal democracies had grown in size and strength and had developed powerful social institutions (market economies, political forums, the legal protection of freedom of thought and liberty of conscience, marriage and informal relational ties, etc.) to reinforce their citizens' commitment to liberal-democratic ends. By the time that Rawls developed justice as fairness in the late twentieth century, the citizens of liberal democracies had been convinced for centuries that free and equal contractual situations among rational agents were appropriate ethical decision-making locations. So although the dispositions of liberal-democratic citizens are locally sequentially prioritized in the development of justice as fairness, the dispositions of such persons are in reality the product of centuries of liberal-democratic institutional influence.

The evident historical impact of liberal-democratic institutions upon the dispositions of the persons whose opinions inform the construction of justice as fairness demonstrates that the issue of ultimate directional influence between institutions and dispositions in liberal democracies is far more muddled than my investigation of justice as fairness has so far suggested. In this section I will propose that a productive way to understand this complicated descent of reciprocal influence is as a relationship of mutual reinforcement. By this I mean that successfully-operating institutions and dispositions in liberal democracies tend naturally to reinforce each other's orientation toward liberal-democratic ends. On the mutual reinforcement view neither institutions nor dispositions ought to be systematically sequentially prioritized in the construction and operation of liberal democracies. Institutions and dispositions ought rather to be seen as intertwined in a complex relationship of reciprocal determination in which both exert an influence upon the other. The question of ultimate sequential priority causes both institutions and dispositions to recede into the primordial distance, as it were, with each one lining up chronologically behind the other.

The mutual reinforcement view is arguably a more appropriate depiction than the institutional priority view of the interaction of institutions and dispositions in liberal democracies: it arguably describes in more appropriate detail the actual interaction of institutions and dispositions, and it likewise prescribes in equally more appropriate detail a normative ideal of the way in which institutions and dispositions ought to interact with each other. The ability of the mutual reinforcement view to provide such a compelling descriptive and prescriptive account indicates that it is more able than the institutional priority view of providing real-world guidance about the operation of liberal democracies. Consider as an example of the descriptive efficacy of the mutual reinforcement view the ethically upright effects of a contemporary market economy upon the dispositions of persons. It is natural for a successfully-operating market economy to develop in its participants the virtues that promote its own flourishing and the flourishing of the overall value

system of liberal democracies. Among these virtues are *autonomy* in the planning and pursuit of market activities, *cooperativeness* with other participants in the achievement of common economic goals, and *decisiveness* in quickly and efficiently selecting among available market options. Persons whose dispositions have been positively impacted by market economies tend naturally to exert a reciprocal influence upon the market mechanisms that influenced them. Through *autonomy* participants regulate and filter their behaviors in accordance with socially-acceptable regulative norms. Through *cooperativeness* participants promote collective deliberation about social goals and are empowered to engage in uncomfortable but necessary specialized tasks. Through *decisiveness* participants prevent operational breakdowns and improve the efficiency of their transactions with producers and consumers. The reciprocal impact of virtuous participants upon the market economy is powerful enough to transform the very rules by which the market is regulated, and to influence in a profound and meaningful way their fellow participants' views of market operations. In the absence of external interference, the reciprocal impact of virtuous participants upon the market economy is sufficient to orient it more directly over time toward its fundamental ideals.

This account of the reciprocal influence of markets and the dispositions of persons is on the mutual reinforcement view a representative instance of the general tendency of persons who have been positively influenced by liberal-democratic institutions to act in ways that reinforce those institutions' fundamental ideals. The account exemplifies the ability of the mutual reinforcement view to provide us with a plausible explanation of the interaction of institutions and dispositions in real-world liberal democracies. If in real life the dispositions of virtuous citizens do exert a transforming reciprocal impact upon liberal-democratic institutions, and I think that it would be relatively easy to construct a series of similar empirical reflections to demonstrate that they do, then the mutual reinforcement view is better positioned than the institutional priority view to describe this reciprocal impact.

Rawls' emphasis upon the sequential unfolding of liberal-democratic moral principles indicates that justice as fairness, in spite of its ideal-level formulation, is intended to provide guidance on the development of real-world societies. But the Rawlsian view is poorly positioned to model the idea of a transforming reciprocal influence, since it depicts the institutional infrastructure as exerting a far more significant influence upon the dispositions of persons than the dispositions of persons exert in return upon the institutional infrastructure. Thus justice as fairness is not as descriptively equipped as the mutual reinforcement view to capture the idea of a transforming reciprocal impact that alters the ideals of liberal-democratic institutions over time.

On the mutual reinforcement view the reciprocal impact of virtuous citizens upon liberal-democratic institutions increases the efficiency, operational capacity, and structural harmony of those institutions. It also strengthens the orientation of those institutions toward fundamental liberal-democratic ideals. The reciprocal impact completes the reinforcement cycle and again places institutions in a position to reinforce the com-

mitment of their participants toward liberal-democratic ends. Repeated iterations of the mutual reinforcement relationship strengthen the respective orientations of institutions and dispositions toward liberal-democratic ends and promote in liberal democracies the greater internal structural cohesion that is characteristic of societies whose institutional and dispositional loci are both oriented toward the same goals.

In the absence of external interference, repeated iterations of the mutual reinforcement relationship would bring about permanently stable and healthy liberal democracies. But in real-life political situations there are almost always mitigating factors which diminish the functional success of the mutual reinforcement relationship. In the past century one such circumstance was the prevalence of large-scale social improvement movements (alcohol prohibition, franchise expansion, environmental protection, etc.). Such initiatives were frequently admirable for their correction of cultural injustices, but they also caused destabilizing disagreements among citizens about core liberal-democratic ideals. Other circumstances in liberal democracies which diminish the accumulating effects of the mutual reinforcement relationship are a large geographical size, a large population, and a diversity of comprehensive belief systems.[21] Thus the self-reinforcing stability of the mutual reinforcement relationship is regularly upset by circumstantial contingencies which diminish the reciprocal positive influence of institutions and dispositions.

In addition to being a more accurate modeling than the institutional priority view of the actual interaction of institutions and dispositions in liberal democracies, the mutual reinforcement view arguably prescribes a better account than the institutional priority view of the way in which institutions and dispositions *ought* to interact in liberal democracies. Its prescriptive account unfolds naturally out of its explanation of why some liberal democracies are healthier than others. All liberal democracies, including unhealthy ones, manifest the mutual reinforcement relationship to a minimal extent. But liberal democracies which do so in a robust manner are more functionally successful than their counterparts. A minimally-instantiated mutual reinforcement relationship promotes the stabilization of the political processes of a liberal democracy. A robustly-instantiated relationship exerts a far more substantial impact upon liberal-democratic political processes – it preserves civil freedoms, regulates the distribution of material resources, and develops a public sphere which is conducive to discussions of foundational political principles. Liberal democracies do not generally work well unless and until their institutions and dispositions exert a significant positive influence upon each other.

It is natural for liberal democracies to improve morally in response to the constructive interaction of their institutions and the dispositions of their citizens. The moral improvement is a move from a minimal to a robust instantiation of the mutual reinforcement relationship, from a lesser to a greater political self-consciousness, and from an inferior to

21] All three of these conditions are capable of diminishing the reciprocal influence of institutions and dispositions upon each other – a large geographical or demographical size causes logistical problems; ideological diversity causes consensus difficulties.

a superior public flourishing. Consider as an example of such improvement the expanded realization of liberal-democratic principles that has occurred over the last three centuries. The improvement has taken place along two axes – the first of which is the spread of liberal-democratic principles to a greater number of peoples, and the second of which is a refinement of liberal-democratic principles among those peoples who already practice them. Upward development along the first axis has meant that many peoples are now enjoying a greater degree of individual and collective flourishing than they did under the illiberal regimes of the past. Upward development along the second axis has meant that such peoples are now endorsing a better and more refined version of liberal-democratic principles than they did three centuries ago.[22] Among the principles that are now more clearly understood as being central to liberal democracies are an acknowledgment of a universal right to religious liberty (where previously religious liberty had been restricted to certain groups), an extension of civil rights to minorities and other underprivileged groups (where previously only a limited number of citizens had enjoyed such rights), and an increasing recognition of a right to dignified working conditions and basic health benefits (where previously there had been no such recognition). Together the two axes of moral improvement have made liberal democracies among the most popular and stable political regimes in recent human history.[23]

The axes of moral improvement have been propelled upward in large part by the constructive interaction of institutions and dispositions. Over time the operation of the mutual reinforcement relationship has exerted an increasingly positive impact upon the political values of liberal democracies – one that preserves civil freedoms, stabilizes the distribution of material resources, and develops a public sphere that is conducive to foundational public discussions. On the basis of this natural positive impact it is reasonable to make the prescriptive claim that the mutual reinforcement relationship functions best when it is allowed to develop naturally. I mean by this that liberal democracies undergo their most rapid and permanent moral improvement when there is minimal external interference in the natural operation of their mutual reinforcement relationship. The moral ascent that results from the constructive interaction of institutions and dispositions is usually in itself sufficient to orient liberal democracies more directly and deeply toward their core ideals over time.

The key prescriptive move of the mutual reinforcement view is the claim that liberal democracies experience optimal moral development when the mutual reinforcement relationship is allowed to operate in a natural way. As a prescription the claim is markedly different from the Rawlsian claim that liberal democracies ought to unfold sequentially

22] Liberal democracies are capable of becoming oriented toward more refined versions of liberal-democratic principles over time, as T.H. Marshall's account of the historical development of citizenship has shown (1997, 291-319).

23] With a few notable exceptions (Kerensky's White Russia, the Weimar Republic, etc.), all of which were impacted by substantial external factors, modern liberal democracies have not typically failed for internal reasons.

out of predetermined principles. On the Rawlsian view liberal democracies ought to be characterized by a mono-directional sequential unfolding in which moral principles are determined in a privileged decision-making location, the institutional infrastructure is organized to instantiate these moral principles, and the dispositions of citizens are educated to conform to the deliverances of the institutional infrastructure. The mutual reinforcement view is unlike the Rawlsian view in that it depicts liberal democracies as developing optimally through a natural (i.e. non-artificial) increase in the instantiation of liberal-democratic principles in their institutions and the dispositions of their citizens. As a prescriptive ideal the mutual reinforcement view is meant to be procedural and is not a derivation of substantive moral content. So it is continuous with the core Rawlsian value commitments for those who wish to see it as such. But as a prescriptive ideal it is certainly not intended to be continuous with the Rawlsian procedural pronouncement in favor of a mono-directional unfolding of liberal-democratic principles.

V. MUTUAL REINFORCEMENT: THE CIVIC NONPARTICIPATION PROBLEM REVISITED

In this article the civic nonparticipation problem has functioned as a catalyst for an investigation of the interaction of institutions and the dispositions of citizens in liberal democracies. The problem is in many ways a breakdown in this relationship, and as such it is appropriate for political philosophers to approach the problem from a particular understanding of how the relationship ought to operate. The dominant contemporary view is that the establishment of institutions ought to be prioritized over dispositional development in the construction of liberal democracies. But the limitations of the dominant view are evident when it is applied to the civic nonparticipation problem.

A more productive approach to the civic nonparticipation problem is to rethink the categories in terms of which we view the interaction of institutions and dispositions in liberal democracies. I have sketched the mutual reinforcement view as a more or less accurate account of this interaction. The grounding of the mutual reinforcement view in real-world political circumstances renders it more capable of interpreting a breakdown in the development of liberal-democratic sentiments than the institutional priority view.

The natural tendency of liberal democracies is to improve morally in response to the positive interaction of their institutions and the dispositions of their citizens. The civic nonparticipation problem is a breakdown in this natural tendency – one that has been caused by an accumulation of conditions over many decades, and one that must be countered by a restoration program. On the mutual reinforcement view the interference of the restoration program ought not to proceed from a perspective that is artificially external to the natural development of liberal-democratic values. By this I mean an interference that is descended from an abstracted and privileged normative location (i.e. like Rawls' original position). Such interference is designed to prevent liberal-democratic citizens from influencing the formation of their fundamental political values rather than to promote the reciprocal positive interaction of institutions and dispositions. It tends to diminish the

ability of liberal democracies to respond productively to real-world contingencies and to progress upwards along a path of moral improvement. Thus on the mutual reinforcement view it is reasonable at the outset to build into the restoration program the constraint that the requisite interference ought not to be inordinately disassociated from the natural development of liberal-democratic values.

The emphasis of the mutual reinforcement view on the mutability of liberal-democratic values highlights the interactive connection between institutions and the dispositions of citizens in liberal democracies. The instantiation of a certain set of values in liberal-democratic institutions causes the dispositions of participants in those institutions over time to be oriented toward such values (and vice-versa, in accordance with the principle of reciprocal influence). The natural reciprocity of institutions and dispositions makes it reasonable to conclude that breakdowns in the functioning of institutions are likely to be accompanied by breakdowns in the dispositions of the participants in such institutions. On the mutual reinforcement view the practical social factors which have been highlighted by contemporary political philosophers as the causes of the civic nonparticipation problem (i.e. television, work concerns, private entertainments, etc.) are the external and visible manifestations of a problem that also exists internally within the dispositions of citizens.

The failure of civic defectors to instantiate liberal-democratic values in their dispositions is on the mutual reinforcement view as much a part of the civic nonparticipation problem as is the external and visible breakdown in institutions. Prime examples of this failure are the cognitive defense mechanisms that civic defectors employ to justify their abandonment of the public sphere: citizens' cynicism and lack of trust in the operation of contemporary political processes, their belief that their personal involvement in the public sphere is not capable of making a significant political difference, and their failure to identify their personal interests with the overall interests of their society (Galston 2007, 625-26). Such cognitive defenses are indicators that the malformed dispositions of civic defectors no longer provide them with the resources to recognize the value of public involvement. The approach of the mutual reinforcement view to the civic nonparticipation problem places a much greater emphasis than the institutional priority view upon the dispositions of citizens. The equal value that it attributes to institutions and dispositions suggests that attempts to restore the moral improvement of liberal democracies ought not to concentrate exclusively on institutional infrastructure in the way that would be typical of a restoration program that is influenced by the Rawlsian institutional priority view. Rather, it suggests that political philosophers would more easily accomplish their restorative aims by targeting the malformed dispositions of civic defectors and by providing civic defectors with reasons they are capable of accepting.

The balance of this article is not lengthy enough to survey the many techniques by which political philosophers are capable of shaping the dispositions of liberal-democratic citizens. One technique that does deserve mentioning, however, is the particularly promising one of providing citizens with a goal toward which to concentrate their political efforts. Consider as an example of such a goal the idea that citizens have the right and the

ability to influence the development of their fundamental political values. If this goal were promoted in liberal democracies it would be sufficiently closely tied to the natural progress of liberal-democratic values to be an acceptable form of restorative interference on the mutual reinforcement view. It would, of course, be incompatible with the now-dominant Rawlsian assumption that the fundamental political values of liberal democracies ought to be beyond the ability of citizens to criticize or reject. But as a goal it would meet the objections of civic defectors and would promote a renewed interest among citizens in the central values of the liberal-democratic tradition – one that would be likely to lead to the improvement of such values over time.

The particular promise of the claim that citizens are allowed and equipped to influence their fundamental political values is its ability to meet the concerns of civic defectors and inspire them to return to the public sphere. The opportunity to influence the fundamental political values of one's society is a great and momentous good. As a good it is perhaps resonant enough to motivate citizens to reject the lure of private entertainments and the other forms of civic defection. At the very least, citizens who are given the opportunity to influence their fundamental political values are denied the excuse that the structure of the political process prevents them from personally making a difference. If such an opportunity were made available to liberal-democratic citizens, it would provide them with a compelling account of the operation of the public sphere and of their role in its ongoing development.

jneill@hbu.edu

REFERENCES

Berlin, Isaiah. 2002. Two concepts of liberty. In *Four Essays on Liberty*, 166-218. Oxford: Oxford University Press.

Cohen, G.A. 1997. Where the Action Is: On the Site of Distributive Justice. *Philosophy and Public Affairs* 26: 3-30.

Dagger, Richard. 1997. *Civic Virtues: Rights, Citizenship, and Republican Liberalism*. New York, NY: Oxford University Press.

Dworkin, Ronald. 1981. What is Equality? Part 1: Equality of Welfare. *Philosophy and Public Affairs* 10: 185-246.

Galston, William. 2001. Political Knowledge, Political Engagement, and Civic Education. *Annual Review of Political Science* 4: 17-24.

———. 2004. Civic Education and Political Participation. *Political Science and Politics* 37: 263-66.

———. 2007. Civic Knowledge, Civic Education, and Civic Engagement: A Summary of Recent Research. *International Journal of Public Administration* 30: 623-42.

Goldman, Alvin, and Jaegwon Kim, eds. 1978. *Values and Morals: Essays in Honor of William Frankena, Charles Stevenson, and Richard B. Brandt*. Dordrecht, Holland & Boston: Reidel.

Hall, Peter. 1999. Social Capital in Britain. *British Journal of Political Science* 29: 417-64.

Husseune, Michel. 2003. Institutions and their Impact on Social Capital and Civic Culture: the Case of Italy. In *Generating Social Capital: Civil Society and Institutions in Comparative Perspective*, eds. Marc Hooghe and Dietlind Stolle, 211-30. Macmillan.

Sandel, Michael. 1982. *Liberalism and the Limits of Justice*. Cambridge: Cambridge University Press.

Jean, Elshtain. 1995. *Democracy on Trial*. New York, NY: Basic Books.

Marshall, T.H. 1997. Citizenship and Social Class, In *Contemporary political philosophy: an anthology*, ed. Robert Goodin and Philip Pettit, 291-319. Oxford, UK: Blackwell Publishers.

Mill, J.S. 1989 [1859]. *On Liberty*. New York, NY: Cambridge University Press.

Murphy, Liam. 1999. Institutions and the Demands of Justice. *Philosophy and Public Affairs* 27: 251-91.

Putnam, Robert. 2000. *Bowling Alone*. New York, NY: Simon and Schuster.

———. 2003. *Better Together*. New York, NY: Simon & Schuster.

Rawls, John. 1977. The Basic Structure as Subject. *American Philosophical Quarterly* 14: 159-65.

———. 1978a. [Addendum:] The Basic Structure as Subject. In *Justice and Economic Distribution*, eds. John Arthur and William H. Shaw. Englewood Cliffs, New Jersey: Prentice-Hall.

———. 1978b. The Basic Structure as Subject. In *Values and Morals: Essays in Honor of William Frankena, Charles Stevenson, and Richard B. Brandt*, eds. Alvin Goldman and Jaegwon Kim, 47-71. Dordrecht, Holland & Boston: Reidel.

———. 1993. *Political Liberalism*. New York, NY: Columbia University Press.

———. 1999. *Theory of Justice*. 2nd ed. Cambridge, MA: Belknap Press.

———. 2001. *Justice as Fairness*. Ed. Erin Kelly. Cambridge, MA: Belknap Press.

Robert, Putnam, ed. 2002. *Democracies in Flux: the Evolution of Social Capital in Contemporary Society*. New York, NY: Oxford University Press.

Rothstein, Bo, and Dietlind Stolle. 2003. The Importance of Passive Membership for Social Capital Formation. In *Generating Social Capital: Civil Society and Institutions in Comparative Perspective*, eds. Marc Hooghe and Dietlind Stolle, 191-210. Macmillan.

Sandel, Michael. 1982. *Liberalism and the Limits of Justice*. Cambridge: Cambridge University Press.

Tarrow, Sidney. 1996. Making Social Science Work Across Space and time: a Critical Reflection on Robert Putnam's *Making Democracy Work*. *American Political Science Review* 90: 389-97.

Public Reason 3 (1): 22-38

Legalizing Selective Conscientious Objection

George Clifford

Abstract. Currently, most nations make no legal provision for selective conscientious objection by potential conscripts or by military personnel who morally object to fighting a particular war. This article argues that nations for multiple moral reasons should grant both potential conscripts and persons already in the military the legal right to apply for and to receive recognition as selective conscientious objectors. Focusing on a nation's collective moral responsibility to permit selective conscientious objection complements arguments centered on individual rights and duties because acknowledging an individual's moral right to selective conscientious objection does not require a nation to recognize that right in law. Nations that authorize selective conscientious objection, from a Kantian deontological perspective, respect the moral autonomy of warriors and potential conscripts. From the perspective of Rawls' concept of moral equality, nations with an option for selective conscientious objection preserve moral equality between warriors, between warriors and civilians, and between conscientious and selective conscientious objectors. Selective conscientious objection, from the perspective of Mill's utilitarianism, benefits the common good by potentially ending an unprofitable war more quickly and avoiding the significant diminution of utility that disrespecting dissent causes. Even if one judges that none of the three arguments sufficiently justifies a nation establishing a provision for selective conscientious objection, the arguments are cumulatively persuasive. Finally, the article refutes several practical objections that prior opponents to selective conscientious objection have interposed against nations implementing provisions for selective conscientious objection.

Key words: selective conscientious objection, conscientious objection, conscription.

Currently, most nations make no legal provision for selective conscientious objection by potential draftees in the event of conscription or by military personnel who morally object to fighting a particular war. This article argues that nations for multiple moral reasons should grant both potential conscripts and persons already in the military (warriors) the legal right to apply for and to receive recognition as selective conscientious objectors. Focusing on a nation's collective moral responsibility to permit selective conscientious objection complements arguments centered on individuals, e.g., selective conscientious objection is a right inherent in a warrior's integrity (Robinson 2009) or freedom of conscience (Hammer 2002, 169), is a warrior's moral duty (Wolfendale 2009), or is not a moral option for warriors (French 2009, and Lucas 2009). Acknowledging an individual's moral right or duty to selective conscientious objection does not require a nation to recognize that right in law (Gans 2002, 44).

After defining key terms, the article's first three sections each present a new or substantially revised moral argument (i.e., a categorical imperative, moral equality, and utilitarian argument) for legalizing selective conscientious objection for potential conscripts and warriors. Even if one judges that none of the three arguments sufficiently justifies a nation establishing a provision for selective conscientious objection, the arguments are cumulatively persuasive. Finally, the fourth section reviews and rebuts practical arguments interposed against selective conscientious objection in prior debates. Emphasizing the practical alongside the philosophical is important because proponents of selective

conscientious objection often have failed to move even those policymakers who seem sympathetic to selective conscientious objection to act, for the latter have contended that creating a selective conscientious objection option entails an administrative nightmare (Wertheimer 2007, 64; Rohr 1971, 103-80).

Although the United States provides the primary context for analysis and source of examples, the article remains persistently international in perspective and relevance. Focusing on a specific nation simplifies the analysis by ignoring largely tangential differences between nations, e.g., some nations permit enlisted personnel to request a discharge before completing obligated service (Great Britain) and others do not (the U.S.). The U.S. provides a helpful context for analysis because neither the U.S. armed forces (U.S. Department of Defense 2007) nor its conscription law (U.S. Military Selective Service Act 1967) currently allows selective conscientious objection.[1] Using Israel, where the issue has recently received much attention, as the primary context risks introducing more heat than light into the debate because of the emotional issues involved in Israeli occupation of Palestinian territories and intrusions into Lebanon, a problem characteristic of selective conscientious objection debates (Capizzi 1994, 340). Additionally, several authors have relied on a false dichotomy between defending Israeli settlements outside Israel's borders or putting those settlers at risk of perishing (Israel Supreme Court 2002; Kasher 2002) when in fact those settlers could withdraw into Israel's internationally recognized borders. Indeed, the Israeli debate on selective conscientious objection has largely ignored the underlying issues (Enoch 2002, 253).

A conscientious objector (CO) opposes all war; a selective conscientious objector (SCO) opposes particular wars (Moskos and Chambers 1993a, 5). Moral reasons why a person may decide to become a selective conscientious objector include the person believing that:

> The aims of a war are unjust, e.g., defending an unjust regime or odious political policy such as apartheid;
> The war is fought in an immoral manner, e.g., consistently targeting civilians;
> The war's evil effects will probably outweigh any good achieved. (Harries-Jenkins 1983, 73; Childress 1982, 201)

An option for selective conscientious objection constitutes an important indicator of a society's respect for freedom: "The ultimate test of a free society is the extent to which

1] Although the 1967 Selective Service Act addresses only conscientious objection based on religious belief, the U.S. Supreme Court has extended that option to those who hold sincere and meaningful objections, including atheists (Chambers 1993, 42) and upheld the constitutionality of the law (Gillette v. United States, 401 US 4370). A small Christian denomination, the Jehovah Witnesses, constitutes the only possible exception to U.S. policies barring selective conscientious objection. Witnesses teach their adherents that they should only fight in theocratic wars, such as the Battle of Armageddon that they believe the Christian Bible foresees (Revelation 16:12-16). In Sicurella v. U.S., the U.S. Supreme Court upheld the right of individuals to claim CO status based being a Jehovah Witness, concluding that theocratic wars like the Battle of Armageddon are legally irrelevant (Childress 1982, 199).

individuals are able to carve out their own destiny on the basis of reflective choice. In shaping one's destiny, few options are more fundamental than the choice between killing and not killing." (Kaufman 1968, 262)

I. THE CATEGORICAL IMPERATIVE ARGUMENT

Although Immanuel Kant formulated multiple versions of the categorical imperative, he staunchly maintained that only one categorical imperative exists: "There is therefore but one categorical imperative, namely this: Act only on that maxim whereby thou canst at the same time will that it should become a universal law" (Kant 2001a, 178). His second formulation of the categorical imperative is particularly relevant to debates about selective conscientious objection: "So act as to treat humanity, whether in thine own person or in that of any other, in every case as an end withal, never as means only" (Kant 2001a, 186). According to Kant, the universality of human moral agency – an individual expressing his/her reason through autonomous action – justifies the categorical imperative to treat people always as ends and never as the means to an end (Kant 2001a, 159-60, 170, 185). Consequently, Kant regards moral autonomy as morality's supreme principle (Kant 2001a, 197). Loss of agency violates the categorical imperative and reduces a person to a means to an end (Kant 2001a, 187).

A nation with military conscription, whose potential conscripts lack a legal right to refuse to serve in what the individual believes to be an unjust war, deprives its civilian populace of their individual moral agency, thereby dehumanizing them (Tramel 2008). From a Kantian perspective, this constitutes heteronomy rather than autonomy, people subject to multiple laws rather than a single moral imperative (Kant 2001a, 190). Rational beings co-located within the space-time matrix reflecting on the categorical imperative should reach the same conclusion about what everyone must do (the universality of the imperative). However, a government that usurps this autonomy by giving those who are not COs only one option (i.e., go to war) deprives potential conscripts of the choice integral to moral autonomy, thereby dehumanizing them. This is not an issue in nations with an all-volunteer military, as citizens who object to a particular war may simply refuse to join the military during that conflict. However, in nations with conscription, the law should establish provision for both conscientious objection (arguing the case for this lies beyond the purview of this article) and selective conscientious objection. The existence of a selective conscientious objection option preserves an individual conscript's moral agency.

More broadly, a government that requires civilians who morally object to particular wars to serve in its military during such a war powerfully communicates that the government regards its citizenry en masse as the means by which to achieve the government's goals (Sturm 1983, 273). This collective dehumanization is in addition to dehumanizing individual objectors (Kant 2001a, 190). Collective dehumanizing, in time, can permeate a nation, as happened in Nazi Germany, the former Soviet Union, and elsewhere. The

Vietnam War evidenced a similar, though significantly less severe, dehumanizing in the U.S. Selective conscientious objectors fled to Canada, engaged in acts of civil disobedience, served in spite of moral objections, suffered criminal penalties for refusing to serve, and became lightning rods for those opposed to other, unrelated forces of social change, e.g., the civil rights movement. Cumulatively, these individual actions diminished the collective humanity of U.S. society: internally polarizing those for and against the war, increasing alienation toward and distrust of government, and spreading a perception that government was no longer by and for the people. Many of those effects linger today. Admittedly, a selective conscientious objection option would not have prevented the entirety of those problems. Yet the lack of an option for selective conscientious objection certainly magnified the Vietnam War's dehumanization of U.S. society, violating Kant's categorical imperative by reducing people – U.S. citizens in this case – to a means to an end.

Arguing that warriors should have an option for conscientious objection to a particular war poses more difficulties than arguing that civilians should have an option for selective conscientious objection. A nation rightly requires warriors to relinquish much of their personal autonomy while serving in the military (Friedman 2006, 81). Military effectiveness necessitates that warriors obey orders promulgated by their hierarchical chain of command. Subordinates perform duties when, where, and as assigned. Subordinates must follow strategies and tactics dictated by higher authority. A nation and its warriors accept these infringements of individual autonomy as necessary preconditions for maintaining an effective fighting force (McMahan 2004, 704). Attempts to respect individual autonomy by introducing an element of democracy (e.g., some U.S. militia units electing their officers at the beginning of the Civil War) proved less than satisfactory (incompetent leaders led to unnecessary fatalities and mission failure). Modern warfare's rapid pace and greater lethality make effective command and control even more critical as prerequisites to victory. In other words, individuals necessarily become a means to the end of national defense while serving in the military (Friedman 2006, 88).

However, many nations, international institutions, and ethicists concurrently expect warriors to maintain a degree of personal autonomy. Warriors in the U.S. have a legal as well as moral responsibility to refuse to obey an illegal or egregiously immoral order (Cook 2004, 63). Anyone who serves very long in the military is likely to receive an order the individual perceives as immoral. Most immoral orders require obedience because they are neither illegal nor egregiously immoral, e.g., ordering someone to work uncompensated overtime on an unimportant, non-urgent project. An egregiously immoral order causes significant unnecessary personal harm, infringes on basic human rights, or results in the destruction of valuable property. A warrior relies upon his or her reason to recognize illegal or egregiously immoral orders. Then the warrior must exercise autonomy, refusing to obey the illegal or egregiously immoral order.

The pragmatic realization that only warriors can prevent certain atrocities (e.g., the My Lai massacre), rather than any intrinsic respect for warriors, probably explains the eti-

ology of these expectations. One clear message of the post-World War II Nuremberg trials is that warriors cannot justify illegal and immoral acts with the defense that they were only obeying orders (Wakin 2000, 149-50; Orend 2006, 171-73). Clearly, life for warriors without this responsibility would be simpler, i.e., a warrior's moral duty would be to obey all orders.

In classic Just War terms, many nations, international institutions, and legal conventions expect warriors to exercise moral agency with respect to the jus in bello non-combatant discrimination and proportionality criteria. International and many national laws incorporate essential elements of the jus in bello standards. The massacre of civilians at My Lai by U.S. Army personnel was criminal because most (all?) of those killed were obviously non-combatants. Iraq's employment of chemical weapons in the 1980s Iran-Iraq war violated a widely accepted international definition of proportionality. The U.S. Army tried Lieutenant Calley and Captain Medina for the events at My Lai; Iraq's new government conducted a trial for former Iraqi government officials who authorized using chemical weapons against Iran. Alleged perpetrators of war crimes in Asia, Africa, and Europe have also faced (or are facing) war crime trials. Holding warriors accused of war crimes accountable and establishing justice for the victims of those crimes remains far from perfect. Nevertheless, a growing international consensus about the need for and importance of the jus in bello standards is emerging, a standard that presumes individuals do not completely surrender their moral autonomy upon entering military service. The larger community – national and international – becomes less safe, less just, and less moral when warriors become only means to ends. Persons who become selective conscientious objectors because they believe a war is being waged immorally (the second reason for becoming an SCO) reinforce the expectation that warriors will refuse to obey illegal and egregiously immoral orders.

Nations and the community of nations should also expect warriors to exercise moral autonomy with respect to Just War Theory's jus ad bellum criteria (Lynd 2011). This argument updates Just War Theory's classical formulation and requires two important caveats. Just War Theory developed prior to contemporary representative democracies and the information age: "Just war theory is unique in contemporary practical ethics in two respects: it is widely and uncritically accepted and differs very little in content from what Western religious thinkers have believed from the Middle Ages to the present" (McMahan 2004, 731). In non-democratic societies, a follower's role and responsibility are more circumscribed and differentiated than that of rulers. Followers may express doubt about the morality of a war, but the decision to go to war belongs exclusively to the ruler and just war theorists have reasonably absolved followers of any moral guilt for fighting in an unjust war (Walters 1973, 201-11; McMahan 2004, 705). Rulers also often had access to fuller and more accurate information than did followers, theoretically enabling rulers to make better decisions (McMahan 2004, 703-4).

Kantian arguments against granting warriors the option for selective conscientious objection, like the one Fiala makes (2008), incorrectly presume those conditions remain

true. Representative democracy – government of and by the people through elected representatives –shares some measure of responsibility for government decisions between those who hold elected office and the electorate. Warriors retain the privilege and responsibility of voting. The advent of the information age provides the electorate more timely and fuller access to knowledge of world events (Wolfendale 2009, 134). Indeed, during both Gulf Wars U.S. government officials relied on public sources such as CNN and the Internet to obtain more timely and accurate information than the officials could obtain from classified government sources. Warriors may have first-hand knowledge or direct access to pertinent information that others cannot access or receive only second hand. In other words, neither of the rationales for treating warriors as a means to an end with respect to jus ad bellum criteria remains valid in a modern representative democracy. This coheres well with Kant's concept of "the will of every rational being as a universally legislative will" (Kant 2001a, 188).

The important caveats for granting warriors a right to selective conscientious objection are that warriors must object that a war is unjust as soon as possible and must remain in situ until the military accepts an individual's request for discharge. The soldier cannot wait until ordered to make a frontal assault upon a heavily fortified enemy emplacement and the sailor cannot wait until five minutes before a ship deploys. The timing of those requests suggests a lack of courage or other ulterior motive rather than the exercise of reason and moral agency. Instead, a person in the military who objects to a particular war has a moral responsibility to act upon the objection immediately after reaching that assessment. This reasonably mirrors the U.S. military's expectation that a warrior who becomes a conscientious objector to all war will immediately act upon his or her new belief. The requirement to remain in situ balances honoring the individual's duty to serve with preserving the individual's status as an end not merely a means to an end. During the U.S. Civil War, some U.S. warriors who aligned themselves with the Confederacy remained in situ until their resignation from the military was accepted, honoring their fiduciary responsibility while preserving their moral autonomy. Other personnel simply walked away from their post or turned over Union military assets to Confederate forces (Avins 1962, 431-61). These breaches of fiduciary responsibility clearly violated Kant's first formulation of the categorical imperative because reliably and consistently fulfilling contracts is a necessary element of human interaction. Unlike the usually immediate consequences associated with an immoral order that violates jus in bello standards, the warrior who objects to a particular war is unlikely to receive an order requiring decisive, personal participation in that war before the military can process his or her request for discharge as a selective conscientious objector. Establishing an option for selective conscientious objection within the military will not significantly degrade fighting effectiveness nor undercut the efficacy of the chain of command's authority but will increase respect for the moral agency of warriors and civilians.

Kantian arguments against selective conscientious objection may also misinterpret Kant's position on the primacy of the categorical imperative over not only individual

choice but also over national law and policies. Kant recognized that a warrior's duty to obey orders potentially conflicts with the duty the categorical imperative imposes (Kant 2001b, 462). He offers an example of this, a war with unjust aims, the first of the three reasons listed in this article's introduction why a person might become a selective conscientious objector (Kant 2001b, 438). In other words, Kant implicitly maintains that warriors should morally object to a war waged primarily for the unjust aim of implementing regime change. Genocidal regimes in the last century certainly call the validity of Kant's example into question. Nevertheless, the very existence of this example emphasizes that in Kant's thinking a warrior's duty defined by the universal categorical imperative overrides the warrior's duty to follow orders.

In sum, denying military personnel and potential conscripts the right of selective conscientious objection presupposes that human nature is incapable of determining the right course of action. In doing so, a nation unnecessarily dehumanizes its citizenry and military by needlessly depriving them of moral autonomy to the detriment of the individual and the global community (Kant 2001b, 467; Rietkerk 2008).

II. MORAL EQUALITY ARGUMENTS

Moral equality in this article denotes equal respect for the basic human liberty of acting upon one's own sense of right and wrong. This is an admittedly incomplete definition, derived from John Rawls' broader definition (2001, 28, 113). This narrow definition focuses on the aspect of moral equality pertinent to the issue of selective conscientious objection. According to Rawls, only the protection of other basic liberties justifies the injustice that abrogating moral equality causes (2001, 23, 111), e.g., conscription to defend a nation's security (1999, 333-34).

Rawls defends a warrior's responsibility to disobey illegal or egregiously immoral orders, i.e., selective conscientious objection by military personnel with respect to the jus in bello criteria (the second reason a person may become a selective conscientious objector) (1999, 333). He also explicitly supports the right of potential conscripts to become selective conscientious objectors exempted from conscription because they object to the aims of a particular war (the first reason a person may become a selective conscientious objector) (1999, 333-35; 2001, 47). In his discussion of selective conscientious objection, Rawls concludes: "What is needed, then, is not a general pacifism but a discriminating conscientious refusal to engage in war in certain circumstances. … Given the often predatory aims of state power, and the tendency of men to defer to their government's decision to wage war, a general willingness to resist the state's claims is all the more necessary." (1999, 335)

A nation granting potential conscripts and active duty personnel an option for selective conscientious objection respects their moral equality without posing a conflict to the basic liberties of other people. This argument addresses three moral equality issues with respect to selective conscientious objection:

> Authorizing selective conscientious objection for military personnel will eliminate one unjustifiable cause of moral inequality among some nation's warriors;
>
> Authorizing selective conscientious objection for military personnel will eliminate one unjustifiable cause of moral inequality between civilians and warriors in nations with all-volunteer armed forces;
>
> Authorizing selective conscientious objection establishes moral equality between a nation's universal and selective conscientious objectors.

First, not granting all warriors an explicit option for selective conscientious objection creates an unjustifiable moral inequality between officers and between officers and enlisted personnel. Officers in many militaries may submit a resignation at any time. In general, militaries prefer officers who want to serve. An unmotivated or negatively motivated leader, rarely, if ever, leads as well as a motivated leader. In battle, this may cost a nation the lives of warriors, perhaps even a defeat. However, no requirement exists for the military to accept an officer's resignation. Indeed, militaries routinely deny an officer's request to resign when the officer has an obligated service commitment connected to education, training, promotion, transfer, etc. Although a military may have the discretion to waive that obligation, militaries are more prone to accept a resignation from one of the many officers who serve without a current obligation than from an officer with obligated service. Enlisted personnel, as in the U.S., may have no option to resign or to request discharge for any reason, including selective conscientious objection to a particular war. The difference between officer and enlisted dates back to an era when officers and enlisted personnel came from different social classes, often with different levels of education, a distinction preserved in U.S. by enlisted personnel signing a contract to serve for a specified period. Since the advent of universal education and the extension of the voting franchise to all citizens, arguments that officers are better prepared to make moral judgments than enlisted personnel have lost their cogency and force (Rawls 2001, 87). By establishing a selective conscientious objection option for all warriors, regardless of rank, nations avoid or eliminate an anachronistic and unjustifiable moral inequality between officers and enlisted.

Second, nations like the U.S. with all-volunteer armed services respect warriors who have conscientious objections to a particular war less than the nation respects civilians who hold the same selective conscientious objection (Tramel 2008). The civilian can act upon his or her moral conclusion by refusing to join the military during the conflict. Although the U.S. may accept protest resignations from some officers who object to a particular war, no option for discharge as a selective conscientious objector exists for enlisted personnel and for officers whose resignation the government refuses. Expecting even the most prescient individual to anticipate every objectionable war during his or her service before entering the military is unreasonable. Thus, the absence of an option for discharge as a selective conscientious objector creates an inherent moral inequality between civil-

ians free to follow the dictates of their conscience and warriors duty bound to obey the orders of a government whose policies they may abhor. Rawls rightly argues that a nation should treat all of its citizens equally unless doing so jeopardizes basic liberties, which is clearly not the case with respect to selective conscientious objection (2001, 87, 132).

Third, in nations that permit conscientious objection because an individual objects to all war, not allowing an option for selective conscientious objection creates a morally unjustifiable inequality between people who subscribe to different belief systems (Rawls 2001, 20). This distinction is especially egregious during conscription. For example, the U.S. generally approved twentieth century requests for exemption from conscription by Quakers who adhered to their Church's teaching of absolute pacifism. Meanwhile, the Roman Catholic, Lutheran, or Anglican who sincerely adhered to his Church's Just War teachings, who concluded that a particular war, such as the Vietnam War, was immoral, and who applied for exemption had that request denied because the U.S. did not exempt selective conscientious objectors. Such individuals were then eligible for conscription. In short, U.S. government policy favored Quaker dissent over dissent from Roman Catholics, Lutherans, Anglicans, and others who dissented based on Just War Theory. This morally privileged Quakers and others who objected to all war. This policy, intentionally or unintentionally, ironically embodied a lack of respect and moral equality for the selective conscientious objector vis-à-vis the universal conscientious objector even though the SCO may develop his position from universal principles that the law recognizes, e.g., moral principles for waging war (Childress 1982, 205-6; Capizzi 1996, 345).

III. UTILITARIAN ARGUMENTS

According to John Stuart Mill, utilitarianism argues that the morally correct action is that which best promotes the common good (2002b, 244). Mill thus seeks to avoid the allegations of hedonism sometimes leveled at utilitarians, insisting that an individual should act not only for personal benefit but also for the public good when he or she has the power to do so (252). He defines utility "in the largest sense, grounded on the permanent interests of man as a progressive being" (2002a, 13). Actions that promote utility (the common and individual good) are right; those that diminish utility are wrong (Mill 2002b, 239). Selective conscientious objection promotes the communal good both by potentially allowing dissent against a war in which the costs appear to outweigh gains (the third reason for becoming a selective conscientious objector) to end that war more quickly and by avoiding the significant diminution of utility that results from disrespecting dissent (Wolfendale 2009, 135; Foster 2009).

Quantifying the effect of selective conscientious objection on terminating a hypothetical war is impossible. The effect may be so slight as to minimize the weight assigned to this argument since both citizens and soldiers generally have reasons for supporting a war, at least in its initial stages (McMahan 2004, 705-6). Citizens generally want to trust an elected government to act in the nation's best interests; military personnel usually want

to believe that any war they fight is just. The Korean War, for example, was unpopular in the U.S. when waged. Although the war's lack of popularity resulted in a ten-fold increase in the number of COs, growing from .15% of conscripts in both world wars to 1.5%, that remained a still relatively insignificant number (Chambers 1993, 39). On the other hand, not all of the evidence about conscientious objection's influence is negative. During the Reagan presidency, U.S. law required eighteen-year-old males to register for a possible resumption of the draft. At least half a million failed to register. There were no prosecutions. This passive protest added momentum to ending mandatory registration.[2] The growing numbers of COs in both the Vietnam War and the first Israeli invasion of Lebanon (1982-1985) may have expedited the conclusion of both wars (Peri 1993, 150-56). The potential for a war's lack of popular support expressed through selective conscientious objection helping to pressure the government to end the war rapidly is real, if unquantifiable and at times small (Goldberg 2006, 70; Knapp 2009; Enoch 2002, 244).

The number of conscientious objectors tends to increase with increasing social opposition to a particular war (Chambers 1993, 39). Rampant militarism in many nations, including the U.S, heightens the importance of creating a selective conscientious objection option (Bacevich 2005, 2). The conditions under which a person may quality as a selective conscientious objector will be one factor in determining the importance of selective conscientious objection for ending a war sooner. In Britain, for example, persons can seek selective conscientious objector status when "materially affected by a specified stimulus." These include belief that the armed forces no longer defend their interests (e.g., in Northern Ireland where some were convinced that the Army did not protect) and belief that the military is a divisive element within society (e.g., socialists who criticized Britain's participation in WWI) (Harries-Jenkins 1983, 74-75). Similarly, but probably less likely, an option for selective conscientious objection might cause a nation's leaders to consider more carefully any future war or military deployment (Ruesga 1995, 68), perhaps even undermining their ability to wage war (Friedman 2006, 91; Foster 2009, 390).

Respecting the right of individuals to express dissent through selective conscientious objection increases a nation's utility. Dissent by a selective conscientious objector is the personal and sometimes political expression of moral ideas: "the objector is an 'officer of society' giving witness to the priority of peace as a political virtue and symbolizing war, even in its most justified forms, is a morally ambiguous endeavor" (Sturm 1983, 276). Representative democracies invest ultimate responsibility for political decisions in the electorate; if sufficient numbers of military personnel or future conscripts became SCOs to diminish the nation's ability to wage a particular war, then perhaps the government lacked justification for the war or for fighting it with conscripts (Rawls 1999, 335; Malament 1972, 382). Rigorous controls, as exist now in the U.S. for officially recognizing

2] Most of those who failed to register were not conscientious objectors opposed to all war. They simply wanted to end the draft, were too lazy to comply with the law, or were ignorant of the requirement to register. (Ginerich 1983, 135-43)

COs, can ensure that those recognized as SCOs are in fact motivated by moral objections instead of other concerns (McMahan 2004, 707). Granting individuals the opportunity to express dissent through selective conscientious objection is consistent with the idea that "respect for personal integrity is a central principle of any just legal and political order" (Sturm 1983, 276; also, Robinson 2008 and Cohen 1968, 276; more broadly, Mill 2002b, 297). From Mill's utilitarian perspective, a just political order is a system that treats all equally; doing otherwise diminishes utility (Mill 2002b, 298-300). An option for selective conscientious objection is the only way a nation can treat both a war's proponents and opponents equally, allowing each to act in accordance with his or her views.

Conversely, the lack of a selective conscientious objection option diminishes a nation's utility. Conscripting an individual who objects to a particular war violates personal integrity, diminishing the conscript's sense of worth. Because of the military's self-image as an honorable profession, an even greater diminution of utility occurs when a warrior who objects to fighting a particular war must nonetheless fight that war (Wolfendale 2009; Robinson 2009, 46). In either case, compelling an individual to kill in spite of the person's moral objections constitutes an especially "inexcusable form of human cruelty" (Ruesga 1995, 71), thereby diminishing the collective utility of that nation.

Selective conscientious objectors who refuse to serve by committing acts of military, criminal, or civil disobedience diminish a nation's utility by eroding its democratic institutions and rule of law while jeopardizing, if the individual is in the military, her or his unit's safety, morale, and ability to accomplish its mission (Kasher 2002, 174-78). Legalizing an option for selective conscientious objection obviously reduces these problems substantially.

Furthermore, the violation of personal integrity through coerced military service and the resultant diminished sense of self-worth often contribute to an individual's moral desensitization. Admittedly, the lack of an option for selective conscientious objection will rarely result in complete moral desensitization. However, for individuals whose personal development locates them on the cusp between general patterns of acting morally or immorally, requiring the person to compartmentalize or otherwise override moral objections to a particular war may very well constitute the tipping point that pushes the balance of their actions from moral to immoral. This same individual subsequently confronted with an illegal, egregious immoral order, as at My Lai, may then obey that order. Before dismissing this line of reasoning as too hypothetical or unlikely, recall the youthfulness and immaturity of many military personnel and potential conscripts. Conversely, the existence of an option for selective conscientious objection promotes moral discourse within the military, encouraging moral development within the military and the nation. Although only a minority of personnel may inquire why a particular war is moral or immoral, and the criteria for distinguishing between the two, that person may subsequently engage others in that discourse thereby contributing to their moral development. During the Vietnam War, Paul Ramsey, concluded that this improvement from an option for se-

lective conscientious objection, in total, might be substantial enough to improve the level of political discourse in the U.S. (1968, 35).

The major adverse consequence for a nation that has a selective conscientious objection option appears to be that the dissent selective conscientious objection generates may badly fray the nation's social fabric. Given the historically small number of COs in all nations, existence of an option for selective conscientious objection will rarely produce a sufficient magnitude of SCOs to tear the social fabric seriously – unless the war is widely perceived as immoral in which case the dissent may prove beneficial. Israeli reservists who refused to mobilize during the first Lebanon war and first Intifada (1987-1994) did not tatter their society in a nation in which Army service is the "entrance ticket to Israeli society" (Linn 2002, 60).

Other potential disadvantages seem slight or remote:

> In most cases, the level of noncooperation that might reasonably be expected in an unjust war would be unlikely to incapacitate or imperil the viability of just democratic or military institutions. Indeed, it seems that those who refuse to fight in an unjust war might in the long term actually benefit their country's institutions by setting a precedent that would help to deter those in positions of authority within the institutions from initiating further unjust wars. It is also possible that those who refuse to participate in an unjust war could prompt the institutions to shield themselves from the instability that such challenges can cause by adapting themselves to anticipate and accommodate instances of conscientious refusal to fight. The enhanced institutional flexibility would almost certainly be healthy and would presumably involve more generous provisions for conscientious refusal to fight. It also seems unlikely that allowing or even encouraging conscientious refusal would seriously impair a country's ability to fight just wars. … Nor is it likely that more liberal provisions for conscientious refusal would prompt malingering in the guise of moral scruple. (McMahan 2004, 705-6)

The utility to a nation of selective conscientious objection, based on the limited evidence that exists and given the difficulties in hypothesizing about the impact of future events (Ruesga 1995, 68), seems to outweigh any disadvantage.

IV. PRAGMATIC ARGUMENTS AGAINST

Four additional arguments against allowing selective conscientious objection warrant brief consideration: the practicality of administering selective conscientious objection, the effect of selective conscientious objection on fighting effectiveness, the alleged impossibility of differentiating political and moral issues, and the inappropriate entanglement of government with religion. These arguments do not represent, prima facie, ethical objections to selective conscientious objection. However, opponents relied upon these arguments to block 1967 efforts to create an option of selective conscientious objection in the U.S. conscription system (Edwards 1972, 120). Arguing for the practicality of establishing a legal option for selective conscientious objection necessitates rebutting these four arguments, exposing them for the "straw men" that they are (Rohr 1971, 181).

Administering an option for selective conscientious objection, within a conscription system or the military, poses no greater challenge than administering an option for conscientious objection within the same system (Malament 1972, 380; Paz-Fuchs and Sfard 2002, 138).[3] The ability of some nations to provide an option for selective conscientious objection suggests that all nations should be able to follow suit. The British, notably, even at their moment of greatest peril during World War II offered potential conscripts an option to seek exemption as an SCO (Malament 1972, 384-85; Brock and Young 1999, 45-47, 156-64, 166). With only sixty-seven thousand World War II applicants for CO/SCO status, the existence of an option for selective conscientious objection did not create an unmanageable burden for Britain nor did it jeopardize national security (Malament 1972, 383-84). Furthermore, the option of selective conscientious objection did not adversely the morale or effectiveness of British warriors (Childress 1986, 119; Moskos and Chambers 1993b, 204). Given that between 2002 and 2010, the U.S. military had roughly 600 applicants for CO status (Vitello 2011), establishing an option for SCO status seems unlikely to create an unmanageable administrative burden or to diminish combat effectiveness significantly.

Concomitantly, alleging the impossibility of differentiating between political dissent and moral objection in the case of selective conscientious objection is a red herring that deflects attention from an effort to suppress dissent. Those evaluating the claims of an applicant for SCO status must focus on the real issues, i.e. the depth and cogency of the applicant's moral objection to a particular war. An applicant's political dissent may indicate, in fact, the degree of sincerity with which the applicant holds the underlying moral objections (Childress 1982, 202). Obviously, applicants for either CO or SCO status may cite moral objections when their actual motive is cowardice, avarice, emotional entanglements, etc. Germany, as a measure of sincerity, requires conscripts granted CO status to serve a longer period in unpaid alternative civilian service than paid military conscripts serve (Kuhlman and Lippert 1993, 101-3). Additionally, the social cost that COs and SCOs usually pay (Chambers 1993, 23-26, 35-36, 43-46; Kuhlman and Lippert 1993, 103-4; Peri 1993, 154-55) helps to deter the insincere from seeking that status, whether in or out of the military. Attempting to stifle political dissent by opposing the existence of an option for legal selective conscientious objection violates basic principles of representative democracy (Murray 1968, 22-23).

Moral objection to a particular war may originate from philosophical analysis or from religious belief. Either offers a paradigm clearly distinguishable from purely political dissent. Harvard philosopher Michael Walzer's *Just and Unjust Wars* (1977) is an enduring classic that offers the primary non-religious basis by which an individual can assess the morality of particular wars (Foster 2009, 392). Religious objections to a particular

3] Ramsey and Israel's Supreme Court disagree without providing explaining their rationales or showing problems experienced by nations that currently permit selective conscientious objection (Ramsey 1968, 31, 34; Israel Supreme Court 2002).

war take diverse forms. Christians often rely upon a Christian version Just War Theory (Ramsey 1983, 94, 124-37; Edwards 1972, 118). In addition to the Roman Catholic Church, Anglicans, and Lutherans, diverse groups including the American Baptist Churches, the United Presbyterian Church, the U.S. National Council of Churches, and the World Council of Churches all support selective conscientious objection that relies upon Just War Theory to assess the morality of a particular war (Finn 1968, vii, x). Some Muslims believe that Muslims should not wage war against other Muslims (Kelsay 2007, 101; Koran 9:5). A Muslim might therefore legitimately seek SCO status because a particular war will involve fighting against other Muslims. Judaism teaches that only defensive wars are moral (Goldberg 2006, 35). A Jew might therefore object to fighting a war that he or she believes is not required for national defense (Peri 1993, 150-56). Free exercise of religion, to be meaningful, must allow individuals not only freedom of belief but also reasonable freedom to act upon those beliefs. Contending that an option for selective conscientious objection excessively entangles government with religion is not only false but also has the effect, hopefully unintended, of unnecessarily limiting religious freedom while, as already noted, creating moral inequality between objectors to all war and selective conscientious objectors (Capizzi 1996, 339).

This article has argued that nations have a moral obligation to establish a legal option for selective conscientious objection for both military personnel and civilians eligible for conscription. Persons may rightly object to a particular war because they believe the war's aims immoral, the war's conduct immoral, or the war's costs to exceed its benefits. Selective conscientious objection affords such objectors a meaningful opportunity to voice those objections and, with government concurrence, to refrain from direct participation in a morally repugnant war. Nations that create an option for selective conscientious objection, from a Kantian deontological perspective, honor and preserve the moral autonomy of warriors and potential conscripts. From the perspective of Rawls' concept of moral equality, nations that permit selective conscientious objection importantly maintain moral equality between warriors, between warriors and civilians, and between a CO and a SCO. Finally, selective conscientious objection, from the perspective of Mill's utilitarianism, benefits the common good by potentially ending an unprofitable war more quickly and avoiding the significant diminution of utility that disrespecting dissent causes. The several practical objections prior opponents to selective conscientious objection have raised all lack validity and are thus do not pose obstacles to nations establishing an option for selective conscientious objection. Moral arguments from a national perspective conclude, as do moral arguments from an individual's perspective, that selective conscientious objection is a moral imperative, underscoring the importance of nations providing warriors and conscripts with an option for selective conscientious objection.

georgeclifford@bellsouth.net

REFERENCES

Avins, Alfred. 1962. Right of military Officers to Resign: An Historical Footnote. *George Washington Law Review* 3: 431-61.

Bacevich, Andrew J. 2005. *The New American Militarism*. New York: Oxford University Press.

Brock Peter and Nigel Young. 1999. *Pacifism in the Twentieth Century* Syracuse. New York: Syracuse University Press.

Capizzi, Joseph E. 1996. Selective conscientious objection in the United States. *Journal of Church and State* 38 (2): 339-63.

Chambers, John Whiteclay. 1993. Conscientious Objectors and the American State from Colonial Times to the Present. In *The New Conscientious Objection*, edited by Charles C. Moskos, and John Whiteclay Chambers II, 23-46. New York: Oxford University Press.

Childress, James F. 1982. *Moral Responsibility in Conflicts.* Baton Rouge, LA: Louisiana State University Press.

———. 1986. Conscientious Objection. In *The Westminster Dictionary of Christian Ethics*, edited by James Childress, and John Macquarrie, 118-20. Philadelphia: Westminster.

Cohen, Carl. 1968. Conscientious Objection. *Ethics* 78 (4): 269-79.

Cook, Martin L. 2004. *The Moral Warrior: Ethics and Service in the U.S. Military.* Albany: State University Press of New York.

Edwards, George R. 1972. *Jesus and the Politics of Violence.* New York: Harper and Row.

Enoch, David. 2002. Some Arguments against Conscientious Objection and Civil Disobedience Refuted. *Israel Law Review* 36: 227-53.

Fiala, Andrew. 2008. Professional Obligation, Conscientious Objection, and the Military: A Kantian Approach. Paper presented at the 2008 International Society for Military Ethics. http://www.usafa.edu/isme/ISME08/Fiala08.html (accessed March 21, 2011).

Finn, James. 1968. Introduction. In *A Conflict of Loyalties: The Case for Selective Conscientious Objection*, edited by James Finn, vii-xiv. New York: Pegasus.

Foster, Gregory. 2009. Selective Conscientious Objection. *Society* 46: 390-93.

French, Shannon. 2009. Sergeant Davis's Stern Charge: The Obligation of Offices to Preserve the Humanity of Their Troops. *Journal of Military Ethics* 8 (2): 116-26.

Friedman, Randy. 2006. The Challenge of Selective Conscientious Objection in Israel. *Theoria* 109 (April): 79-99.

Gans, Chaim. 2002. Right and Left: Ideological Disobedience in Israel, *Israel Law Review*, 36: 19-71.

Gingerich, Andre. 1983. A Draft Resister. In *Peacemakers: Christian Voices From the New Abolitionist Movement*, edited by Jim Wallis, 135-43. San Francisco: Harper and Row.

Goldberg, Edwin C. 2006. *Swords and Plowshares: Jewish Views of War and Peace.* New York: URJ Press.

Hammer, Leonard. 2002. Selective Conscientious Objection and International Human Rights. *Israel Law Review* 36: 145-69.

Harries-Jenkins, Gwyn. 1993. Britain: From Individual Conscience to Social Movement. In *The New Conscientious Objection*, edited by Charles C. Moskos, and John Whiteclay Chambers II, 67-79. New York: Oxford University Press.

Kant, Immanuel. 2001a. Fundamental Principles of the Metaphysics of Morals. Translated by Thomas K. Abbott. In *Basic Writings of Kant*, edited by Allen W. Wood, 143-222. New York: Modern Library.

———. 2001b. To Eternal Peace. Translated by Carl J. Friedrich. In *Basic Writings of Kant*, edited by Allen W. Wood, 433-76. New York: Modern Library.

Kasher, Asa. 2002. Refusals: Neglected Aspects. *Israel Law Review* 36: 171-80.

Kaufman, Arnold S. 1968. The Selective Service System: Actualities and Alternatives. In *A Conflict of Loyalties: The Case for Selective Conscientious Objection*, edited by James Finn, 240-62. New York: Pegasus.

Kelsay, John. 2007. *Arguing the Just War in Islam.* Cambridge, MA: Harvard University Press.

Knapp, Kathryn. 2009. Oaths, Dissent, and Selective Participation in the United States Military and the Israeli Defence Forces. Paper presented at the 2009 Meeting of the International Society on Military Ethics. http://www.usafa.edu/isme/ISME09/Knapp09.html (accessed March 26, 2009).

Kuhlman, Jürgen, and Ekkehard Lippert. 1993. The Federal Republic of Germany: Conscientious Objection as Social Welfare. In *The New Conscientious Objection*, edited by Charles C. Moskos, and John Whiteclay Chambers II, 98-105. New York: Oxford University Press.

Linn, Ruth. 1996. When the Individual Soldier Says 'No' to War: A Look at Selective Refusal during the Intifada. *Journal of Peace Research* 33 (4): 421-31.

———. 2002. Soldiers with Conscience Never Die – They are Just Ignored by their Society. Moral Disobedience in the Israel Defense Forces. *Journal of Military Ethics* 1 (2): 57-76.

Lucas, George. 2009. Advice and Dissent: 'The Uniform Perspective.' *Journal of Military Ethics* 8 (2): 141-61.

Lynd, Staughton. 2011. Someday They'll Have a War and Nobody Will Come. *Peace and Change* 36 (2): 156-71.

Malament, David B. 1972. Selective Conscientious Objection and Gillette Decision. *Philosophy and Public Affairs* 1 (4): 363-86.

McMahan, Jeff. 2004. The Ethics of Killing in War. *Ethics* 114 (July): 693-733.

Mill, John Stuart. 2002a. On Liberty. *The Basic Writings of John Stuart Mill*, 3-119. New York: Modern Library.

———. 2002b. Utilitarianism. *The Basic Writings of John Stuart Mill*, 233-301. New York: Modern Library.

Moskos, Charles C. and John Whiteclay Chambers II. 1993a. The Secularization of Conscience. In *The New Conscientious Objection*, edited by Charles C. Moskos, and John Whiteclay Chambers II, 3-20. New York: Oxford University Press.

———. 1993b. Conclusion: The Secularization of Conscience Reconsidered. In *The New Conscientious Objection*, edited by Charles C. Moskos, and John Whiteclay Chambers II, 196-208. New York: Oxford University Press.

Murray, John Courtney. 1968. War and Conscience." In *A Conflict of Loyalties: The Case for Selective Conscientious Objection*, edited by James Finn, 19-30. New York: Pegasus.

Orend, Brian. 2006. *The Morality of War.* Orchard Park, NY: Broadview Press.

Paz-Fuchs, Amir and Michael Sfard. 2002. The Fallacies of Objections to Selective Conscientious Objection. *Israel Law Review* (36): 111-43.

Peri, Yoram. 1993. Conscientious Objection in a Democracy Under Siege. In *The New Conscientious Objection, edited by* Charles C. Moskos and John Whiteclay Chambers II, 146-57. New York: Oxford University Press.

Ramsey, Paul. 1968. Selective Conscientious Objection: Warrants and Reservations. In *A Conflict of Loyalties: The Case for Selective Conscientious Objection*, edited by James Finn, 31-77. New York: Pegasus.

———. 1983. *The Just War: Force and Political Responsibility* Lanham, MD: Rowman and Littlefield Publishers, Inc.

Rawls, John. 1999. *A Theory of Justice*, rev. ed. Cambridge, MA: Harvard University Press.

———. 2001. *Justice as Fairness*, edited by Erin Kelly. Cambridge, MA: Harvard University Press.

Rietkerk, Tim. 2008. The Duty of the Selective Conscientious Objector in a Values Based Army, *The Army Chaplaincy*. Spring-Summer: 70-76. http://www.ncmaf.org/resources/ArmyChaplaincyNewsletterJuly08.pdf (accessed March 21, 2011).

Robinson, Paul. 2009. Integrity and Selective Conscientious Objection. *Journal of Military Ethics* 8 (1): 34-47.

Rohr, John A. 1971. *Prophets without Honor.* Nashville: Abingdon.

Ruesga, G. Albert. 1995. Selective Conscientious Objection and the Right Not to Kill. *Social Theory and Practice* 21 (1): 61-81.

Sturm, Douglas. 1983. Constitutionalism and Conscientiousness: The Dignity of Objection to Military Service. *Journal of Law and Religion* 1 (2): 265-77.

Supreme Court of Israel. 2002. H.C. 7622/02, Zonshein v. Judge Advocate General. Reprinted in *Israel Law Review* 36: 1-17.

Tramel, Peter. 2008. Immoral Military Service and Selective Conscientious Objection. Draft paper presented at 2008 International Society on Military Ethics. http://www.usafa.edu/isme/ISME08/Tramel08.html (accessed August 13, 2008).

U.S. Department of Defense 2007. *Conscientious Objectors.* Directive 1300.6.

U.S. Military Selective Service Act 1967. Section 6j.

Vitello, Paul. 2011. Midshipman, Then Pacifist: Rare Victory to Leave the Navy. *New York Times*, February 22, 2011. http://www.nytimes.com/2011/02/23/nyregion/23objector.html (accessed March 10, 2011).

Wakin, Malham M. 2000. *Integrity First: Reflections of a Military Philosopher.* Lanham, MD: Lexington Books.

Walters, LeRoy. 1973. A Historical Perspective on Selective Conscientious Objection. *Journal of the American Academy of Religion* 41 (2): 201-11.

Walzer, Michael. 1977. *Just and Unjust Wars.* New York: Basic Books.

Wertheimer, Roger. 2007. Reconnoitering Combatant Moral Equality. *Journal of Military Ethics* 6 (1): 60-74.

Wolfendale, Jessica. 2009. Professional Integrity and Disobedience in the Military. *Journal of Military Ethics* 8 (2): 127-40.

Public Reason 3 (1): 39-49

The Extension and Limits of the Duty to Rescue

Per Bauhn
Linnaeus University

Abstract. While many people believe that we have a moral duty to rescue another human being from a mortal danger when we can do so at little cost to ourselves or to other persons, there is less agreement concerning the extension and limits of this duty. When do we have a duty to rescue? What exactly is meant by "little cost"? In this paper we will examine a consequentialist as well as a deontological way of approaching the duty to rescue. We will point to some significant problems for both versions, but also indicate a way in which at least the deontological position can be improved. Finally, we will try to indicate how the duty to rescue may be applied to the international scene.

Key words: duty to rescue, consequentialism, deontology, duty of necessity, relational duty.

I. THE CONSEQUENTIALIST VERSION OF THE DUTY TO RESCUE

Here is how Peter Singer famously discusses the duty to rescue, relating it to a consequentialist principle of preventing bad things from happening:

> "[I]f it is in our power to prevent something bad from happening, without thereby sacrificing anything of comparable moral importance, we ought, morally, to do it.... An application of this principle would be as follows: if I am walking past a shallow pond and see a child drowning in it, I ought to wade in and pull the child out. This will mean getting my clothes muddy, but this is insignificant, while the death of the child would presumably be a very bad thing [...]. [T]he principle takes, firstly, no account of proximity or distance. It makes no moral difference whether the person I can help is a neighbor's child ten yards from me or a Bengali whose name I shall never know, ten thousand miles away. Secondly, the principle makes no distinction between cases in which I am the only person who could possibly do anything and cases in which I am just one among millions in the same position." (Singer 1972, 231–32)

Hence, according to Singer, we have a duty to rescue whenever our intervention is sufficient to prevent something bad from happening and when it does not involve a sacrifice of "comparable moral importance." His account raises at least two objections, both of which emanate from the consequentialist idea that we have a duty to rescue whenever our intervention is sufficient to prevent something bad from happening. The first objection concerns consequentialism's inability to make an important distinction between two different ways in which we can have a duty to rescue. The second objection concerns the maximization inherent in consequentialism and how it may cause confusion as to whom the duty to rescue applies.

Sufficiency vs. Necessity and Different Kinds of Duties

The idea that we have a duty to rescue whenever it is *sufficient* to prevent something bad from happening, regardless of whether our intervention is also *necessary*, seems to ignore the fact that some agents may have more of a duty to intervene than others, at least in

a given situation. Of course, if you are the only person present, and you can easily rescue the child from drowning, then it is both necessary and sufficient that you intervene, and therefore it would also be obvious that you have a duty to rescue the child.

But what if the child's parents are present at the scene and do nothing, while being both aware of what is happening and capable of intervening? Or what if a trained lifeguard, employed to make the pond safe, is present without intervening, although he is capable of doing so? Or what if the person who pushed the child into the water in the first place is there, watching the child drown, without intervening, although she could easily do so?

In all these cases there exist certain relationships between the agents mentioned and the drowning child, and these relationships also create special duties on the part of the agents. We may call these *relational duties*. Parents have a special responsibility for their children that nobody else has. They are morally obliged to look after their children and see to it that they come to no harm. Likewise, a lifeguard employed to maintain the safety of a certain pond has a special responsibility to rescue a child about to drown there that nobody else has. And a person who has pushed a child into the water and thereby exposed it to a life-threatening danger has a duty that nobody else has to correct her wrongdoing by pulling the child out of the water.

If you, an unrelated passer-by, find yourself by the pond in the company of any or all of these agents, it is they rather than you who have a moral duty to rescue the child. It may be the case that it is sufficient that only one of you intervenes to rescue the child, but you are not all equally morally obliged to do so. The parents, the lifeguard, and the person who pushed the child into the water have, for different reasons, a duty to intervene that you do not have.

However, if all of these people who have a relational duty to rescue the child either are unable to fulfil that duty, or simply refuse to act in accordance with it, then it would be necessary for you to intervene to save the child from drowning. Then you would have what we may call a *duty of necessity*, emanating from the causal necessity of your intervention, rather than from any special relationship between you and the drowning child. Hence, we do not adhere to the view, criticized by Joel Feinberg, that "apart from special moral relationships, our moral claim against others is only to be let alone" (Feinberg 1987, 131). If your intervention is necessary to save the child's life, and if you can intervene at no cost to yourself, then the child has a right to your intervention.

But the primary duty to intervene rests with the people who stand in a special relationship to the child. If you have to intervene because they refuse to act in accordance with their relational duty to rescue, they are morally responsible not only for failing to help the child but also for leaving it to you to make up for their failure. It is a weakness of the consequentialist version of the duty to rescue that it fails to distinguish in this way between relational duties and duties of necessity, and between agents with different degrees of responsibility for fulfilling the duty to rescue.

Sufficiency vs. Necessity and the Maximization of Outcomes

The consequentialist version of the duty to rescue also suffers from a tension between, on the one hand, the intutions evoked by its premises and, on the other hand, the way it justifies its conclusions. We believe that a child, in a situation of the kind described by Singer, has a *right* to our assistance, and that we have a corresponding duty to rescue the child. Unless our intervention brings with it a serious threat to our own well-being, there can be no justification whatsoever for us not rescuing the child.

However, consequentialism, concerned with total outcomes rather than with individual rights, is open to the possibility that we are not only permitted, but actually morally required to leave the child to drown. Singer's justification of the duty to rescue takes place against the background of a more general argument, holding that we should pay attention to all the interests of all those affected by our planned course of action and "weigh up all these interests and adopt the course of action most likely to maximize the interests of those affected" (Singer 1979, 12). This means, however, that we may well face situations in which the duty to rescue the child in the pond may be outweighed by a duty to save other people somewhere else.

Consider, for instance, a case in which you are on your way to the post office to mail a sum of money which will save ten persons from dying from starvation in some far-away country – but only if the money is mailed today. You pass a pond where a child is about to drown. You realize that you can easily pull the child out of the water, but it will take some time to do so, and the post office is about to close. So if you stop to save the drowning child, you will not make it in time to the post office, and then you will not be able to save the ten starving persons. From a consequentialist point of view, the loss of ten lives would be a morally worse outcome than the loss of one life. Hence, your duty is to proceed to the post office and let the child drown.

Now, this result would seem counterintuitive to many of us. Whatever duty you may have to save the starving ten, it cannot set aside the more urgent duty to rescue the drowning child. In the words of Patricia Greenspan, "I do not have moral leeway [...] to pass by an accident victim whom no one else is available to help, on the grounds that I have given or plan to give enough aid elsewhere" (Greenspan 2010, 197).

Some might think that this is because "we have greater obligations to take care of *what is in the area near us*, whether this is threats that will cause harm at a distance, or persons who are or will be victims" (Kamm 2000, 671; emphasis in original). However, this is not a strong argument against the consequentialist position, since the consequentialist may simply counter by questioning the validity of our intuitions. She could claim that they just reflect our prejudice for proximity and against distance (in time as well as in space). We are moved to act by what we have in front of us here and now, and find it easy to ignore consequences more remote in time and space, regardless of the relative importance of the interests at stake.

A stronger objection against the consequentialist position would instead point to the fact that while it is necessary that you intervene to rescue the child from drowning, it is not necessary that you get to the post office in time in order for the starving ten to be saved. Of course, it is necessary that you arrive in time at the post office for the ten to be saved *by you*, but even if you do not get there in time, there is still the possibility that someone else can contribute the sum of money required for saving their lives. Hence, your money is *sufficient* to save the starving ten, but not *necessary*. The child in the pond, on the other hand, can only be saved by you. If you do not pull the child out of the water, no one else will do it. Hence, you have a duty of necessity regarding the child that you do not have regarding the starving ten. Once again, we find that the consequentialist idea that we have a duty to rescue whenever our intervention is sufficient to save someone's life tends to confuse our moral priorities.

What if your intervention would be as necessary for the rescue of the starving ten as it is for the child in the pond? That is, only you can save the child in the pond, and only you can save the starving ten, but you cannot do both. Here we would have to admit that you should give priority to the starving ten. Numbers do not decide by themselves what is the morally right action, but in a case where two groups of people have equally good claims on your support (your intervention is necessary to both of them, and none of them stands in a relationship to you that gives you special relational duties to its members) it seems reasonable that you should choose to help the more numerous group. But unlike consequentialism, quantity is here secondary to quality, in the sense that it is first when we have ascertained that there is no difference in terms of necessity or special relationships that we let numbers decide our duty.

II. THE DEONTOLOGICAL VERSION OF THE DUTY TO RESCUE

Like Peter Singer, Alan Gewirth, in his discussion of the duty to rescue, takes his point of departure in a case of a drowning person:

> "[W]henever some person knows that unless he acts in certain ways other persons will suffer basic harms, and he is proximately able to act in these ways with no comparable cost to himself, it is his moral duty to act to prevent these harms.... Suppose Carr, who is an excellent swimmer, is lolling in the sun on a deserted beach. On the edge of the beach near him is his motorboat, to which is attached a long, stout rope. Suddenly he becomes aware that another person, whom I shall call Davis, is struggling in the water some yards away. Carr knows that the water is about thirty feet deep at that point. Davis shouts for help; he is obviously in immediate danger of drowning. Carr sees that he could easily save Davis by swimming out to him, or at least by throwing him the rope from his boat. But Carr simply doesn't want to bother even though he is aware that Davis probably will drown unless he rescues him. Davis drowns." (Gewirth 1978, 217–18)

Here the background is not the consequentialist one of being required to prevent bad things from happening whenever one can do so at little cost to oneself, that is, when one's intervention is *sufficient* to prevent bad outcomes. Instead, as the formulation "*un-*

less he acts in certain ways" indicates, we are required to intervene when it is *necessary* to prevent basic harms and when we can do so at "no comparable cost" to ourselves. The reference to "basic harms" also implies that we have no duty to intervene to prevent bad outcomes in general, but only to prevent certain harms that interfere with other people's right to basic well-being.

However, we still need to clarify the meaning of "comparable cost". The deontological version presupposes that the rescuer as well as the rescuee have *rights* to basic well-being. That is why we have a moral duty to intervene when it is necessary to prevent someone else from suffering basic harm. But this also raises questions of conflicts of rights. To what extent is the rescuing agent supposed to sacrifice aspects of her own well-being for the sake of maintaining the basic well-being of some other person? Gewirth's story about Carr and Davis does not provide us with any clue here, since Carr obviously does not risk any bodily harm at all by rescuing Davis.

The comparable cost condition

At one extreme, we have the possibility that every loss up to the level of loss that the rescuee is confronted with can be required of the rescuing agent. If the person in need of rescue is about to lose her life, then everything except loss of her own life can be required of the rescuing agent. This seems much too demanding, however. As Jonathan Quong has pointed out:

> "If a child is drowning and X can rescue the child at the cost of muddying their trousers, most will agree X is required to save the child. But suppose instead X can only save the child at the cost of becoming a paraplegic. Here I think many would agree X is no longer required by morality to save the child. Since the death of the child is worse than the cost of becoming a paraplegic, the only explanation is that agent-relative considerations have altered what morality permits." (Quong 2009, 517)

To be sure, professional lifeguards, close friends, relatives, and others who stand in a special professional, contractual, or emotional relationship to a drowning person may have a relational duty to risk even their lives when necessary to save that person. But in the absence of such special relationships, we assume that the rescuing agent, too, has a right to basic well-being that cannot be set aside for the sake of maintaining that same right of a drowning person. This is what the comparable cost condition is about.

Now, granting that becoming a paraplegic would violate the comparable cost condition, we have not said anything about what kinds of harm a rescuing agent *should* be morally required to accept for himself. Certain kinds of harm seem trivial compared to what a drowning person is about to lose, and would hence be consistent with the comparable cost condition. For instance, an opera singer may catch a common cold if she tries to rescue a drowning man, and as a consequence she will be unable to perform arias for some time. But does this imply that she is entitled to refuse to rescue the drowning man? That would be absurd. After all, the opera singer's loss is limited and temporary, while the drowning person is about to suffer a loss that is total and permanent.

Here we may have a point of departure for a more principled argument concerning the contents of the comparable cost condition. We will claim that total and permanent losses of capacities for action constitute attacks on any agent's basic well-being of a kind that is ruled out by the comparable cost condition, while limited and temporary losses of capacities for action may be acceptable according to this condition.

The Right to Basic Well-being

Basic well-being, then, will be taken to include not only life, but also those other physical and mental abilities that are required for agency in general (and not only for certain specific actions). Hence, no agent is morally obliged to risk her life, nor to expose herself to a total and permanent loss of such capacities as the ability to use one's limbs, the ability to see, hear, speak, and so on, the ability to stay concentrated and focused, the ability to perceive and interpret one's natural and social environment correctly (not suffering from delusions), and so on.

Accordingly, no agent can be morally required to make a total and permanent sacrifice of any aspect of her basic well-being for the sake of rescuing another agent's life. (That is, unless her duty to rescue is also a relational duty obliging her to do more than is required by the comparable cost condition.) However, a rescuing agent may be required to make a limited and temporary sacrifice of aspects of basic well-being, such as having her hearing impaired (without becoming completely deaf) for a few days, or suffering from a mild headache (not severe enough to make thinking and concentration impossible) for a day or two. This is how we should understand the comparable cost condition.

Of course, "temporary" is a vague term. Is a headache that lasts a year still to be called temporary? It is certainly not a permanent affliction, but it is not a short-term experience either. We might avoid this problem by simply saying that when it is unclear whether a loss of basic well-being will be limited and temporary, we should leave it to the rescuing agent to decide whether she should intervene or not. In these cases of uncertainty then, the duty to rescue will become an *imperfect duty*, in the sense indicated by Kant and Mill. According to Kant, this means that "the duty has in it a latitude for doing more or less, and no specific limits can be assigned to what should be done" (Kant 1996[1797], 156; Akademieausgabe 6:393). In Mill's terminology, those duties are imperfect "in which, though the act is obligatory, the particular occasions of performing it are left to our choice" (Mill 1987[1863], 66).

When we hold that an agent has no duty to expose herself to a total and permanent loss of any aspect of basic well-being for the sake of rescuing a person in mortal danger, this applies not only to the dangers of the rescue operation itself, but also to the expected effects of rescuing the person in question. Just as the agent has no duty to jump into the ocean to save a drowning person if she herself cannot swim, so she has no duty to rescue a drowning person who in the past has made credible threats that he will kill or mutilate

the agent if he ever gets the opportunity to do so. We are under no duty to risk our own basic well-being by providing opportunties for somebody else to do us unjustifiable harm.

Possible Exceptions: Relational Duties and Fairness

We should note, however, that the comparable cost condition can be set aside by an agent's relational duties. For instance, if the agent has chosen to become a bodyguard, she may well be morally obliged to risk even her life for the sake of protecting her client's basic well-being. Likewise, parents, lovers, and close friends are morally expected to risk their basic well-being when it is necessary to maintain that same basic well-being of their children, partners, and loved ones. But all these relationships should have a background in the agent's voluntary commitments and hence be at least indirectly consistent with her right to freedom. Only in this way can there be a morally justified duty for the rescuing agent to take risks beyond what is required by the comparable cost condition for the sake of saving another person's life.

Could an agent's duty to rescue be limited for other reasons than for being inconsistent with the agent's own right to basic well-being? One such reason is suggested by Liam Murphy in his discussion of fairness in relation to the duty to rescue. Given a situation in which there are many potential rescuers – you being one of them – and many potential rescuees, we may claim that each rescuer is responsible for a certain *share* of rescuees. But what will happen if the other rescuers refuse to intervene? Will you have to rescue more persons than is your fair share, or are you entitled to limit your rescue activities to include just that number of persons that is your fair share? According to Murphy, it seems at least intuitively plausible to hold that "[w]e should do our fair share, which can amount to a great sacrifice in certain circumstances; what we cannot be required to do is other people's shares as well as our own" (Murphy 1993, 278).

Here I would like to turn Murphy's conclusion on its head, however. While you cannot be morally obliged to make "a great sacrifice", since this seems to imply a loss of your basic well-being, you may well be morally obliged to do more than your "fair share" of a rescue operation. If other potential rescuers do not intervene to save some persons in need of being rescued, and you can save these rescuees at little cost to yourself, then it is your duty to rescue them, even if this means that you will have to save more persons than would have been the case had all potential rescuers done their fair share. It is the cost to yourself, not the fairness of your share, that might limit your duty to rescue.

III. THE DUTY TO RESCUE AND THE DISTANT STARVING

As we have already seen, Peter Singer claims that "[i]t makes no moral difference whether the person I can help is a neighbor's child ten yards from me or a Bengali whose name I shall never know, ten thousand miles away." Can the duty to rescue really be extended to global humanitarian aid in this way?

Do Individuals Have a Global Duty of Necessity to Rescue?

As we have already noted, we have a duty of necessity to rescue someone whose basic well-being is endangered, if our intervention is indeed necessary to prevent this from happening. If it is true that certain people will die from starvation in some distant third world country unless you contribute a certain sum of money, and if it is also true that this will not deprive you of any aspect of your basic well-being, then you have a duty to make that contribution.

However, it is very rare that we can establish a causal relationship between one potential donor here and some victims of starvation there, such that if this particular donor does not contribute her money, these people will die from starvation. It is not like the case of the drowning child, where a limited number of people are present, and you know that if none of the others intervene, then it is necessary that you intervene. In the case of the distant starving, it is not at all obvious who has a duty of necessity to help them. Why you? Why not your neighbour? Why not any other citizen of your country? Why not any other citizen of any other wealthy country?

To place the burden of contribution on your shoulders alone would be unfair, given that there are countless other individuals who are equally well off and who could provide the contribution required. On the other hand, for each and every one of these other individuals it will also be true that her specific contribution is not necessary, since it could be provided by some other member of this group of wealthy potential donors. Hence, we seem to lack what Violetta Igneski calls a *morally determinate* situation, connecting a particular rescuer with a particular rescuee (Igneski 2001, 606–7). So who has a duty to rescue the distant starving?

Relational Duties of States

As we have already noted, the duty to rescue comes not only in the form of duties of necessity, but also in the form of relational duties. Now, relational duties apply not only to individuals who stand in a special emotional or professional relationship to the rescuee, such as parents, friends, lifeguards, and so on. Relational duties also apply to institutions and, indirectly, to the persons who are in charge of or work for them. Especially important when it comes to dealing with human afflictions like starvation is the institution of the *state*. This is so, since starvation, unlike the case of a drowning person, cannot be categorized as a sudden occurrence of danger that threatens the basic well-being of some individual and that can be averted by the intervention of some other individual.

Instead, starvation is often the final outcome of a long process of deterioration of communal life in which citizens are deprived of political rights and hence also of the means to voice their grievances. It has been pointed out that "there has never been a famine in a functioning multiparty democracy" (Sen 1999, 178). And since the state, and more precisely the government of the state, has a relational duty to maintain and protect the political rights of its citizens – this is the responsibility that comes with political sover-

eignty – the government will also have a duty to support those of its citizens who starve as a consequence of governmental misrule.

Of course, to the extent that the governments of other countries have contributed to create adverse conditions of development for the starving nation by, for instance, maintaining very unfair conditions of trade, it could be argued that they, too, have a relational duty to support the starving. (This would be a compensatory kind of relational duty, similar to the one that figures in the argument that the person who pushed the child into the water also has a duty to pull the child out of the water.) Thomas Pogge has made an argument to this effect, claiming that world poverty is a consequence of "institutional arrangements [...] to which most of the world's affluent are making uncompensated contributions" (Pogge 2005, 721).

An International Duty to Rescue?

Could there be a duty for another state to intervene in support of victims of starvation in a poor country, that has nothing to do with past wrongdoing on the part of the intervening state? For instance, if it is necessary to remove the dictatorial government of the poor country in order to end starvation there, and a neighbouring country has the military means to do so – would that country thereby also have a duty to intervene? What if this military intervention can be expected to result in casualties among the intervening soldiers – is it still a duty for the neighbouring country to intervene?

Whether or not we accept that there is a duty to intervene militarily in another country for the sake of rescuing its starving people will depend on how we conceive of international relations in general. If we think of states as being members of a "community of nations", we might well accept that "[p]eoples have a duty to assist other peoples living under unfavorable conditions that prevent their having a just or decent political and social regime" (Rawls 1999, 37). This international duty of assistance could then include military interventions, given that such interventions are necessary and feasible.

However, to the extent that we accept such an international version of the duty to rescue, it will be more like a duty of necessity than a relational duty – that is, in the absence of alliances or treaties that commit one state to pay special attention to the communal well-being of another state. Just as individuals are supposed to intervene to each other's rescue when it is necessary and when they can do so without sacrificing their own basic well-being, so one state will have a duty to intervene for the sake of saving the people of another state – but only when it is necessary, and when the intervening state can do so without sacrificing important aspects of its own communal basic well-being.

Hence, the duty of a government to rescue another nation will be limited by the government's relational duty to its own citizens not to endanger their basic well-being, at least not without their consent or unless it is necessary to preserve their own political community. However, governments have a relational duty to its own citizens not only not to endanger their basic well-being, but also to *promote* their well-being. This can be expected

to set further limits to their duty to intervene militarily for the sake of rescuing citizens of another country.

Hence, we may conclude that the duty to rescue indeed can be applied at an international level. However, the extension of the international application of the duty to rescue will be limited by the negative and positive relational duties of governments to their citizens.

IV. SUMMARY OF CONCLUSIONS

The consequentialist version of the duty to rescue exhibited certain weaknesses in its inability to distinguish between relational duties and duties of necessity, and hence of different degrees of moral responsibility among potential rescuers. According to the consequentialist, you have a duty to prevent bad things from happening whenever you can do so, regardless of whether it is necessary that you do it or not. This focus on *sufficiency* rather than *necessity*, in combination with the maximizing aim of consequentialism, also imply that you might have to ignore some rescuees whom only you can rescue for the sake of rescuing a larger group of rescuees who could be rescued by other people as well.

The deontological version of the duty to rescue, on the other hand, accepts that you have a duty to rescue only when it is necessary that you do so. However, since the deontological version is based on the idea that both rescuer and rescuee have moral *rights*, it brings with it an obvious risk that these rights will conflict. The formula of "comparable cost" hence needs to be clarified, in order that we should be able to know how far the rescuing agent's duties extend and what sacrifices she has to accept for herself.

We outlined an idea of the right to basic well-being according to which no agent should have to risk neither her life nor a total and permanent loss of the physical and mental capacities generally needed for agency for the sake of rescuing another agent. (That is, unless the agent has voluntarily entered relationships which bring with them relational duties that go beyond the comparable cost condition.)

With this elaboration of the deontological version of the duty to rescue we will arrive at a relatively clear picture of what is involved in the duty to rescue and of the limits of this duty. Finally, we have applied our reasoning about the duty to rescue to the international scene, and established, at least in a tentative way, that there might indeed be an international duty to rescue, but that this duty is also limited by the relational duties of the intervening state or government to its own citizens.

per.bauhn@lnu.se

REFERENCES

Feinberg, Joel. 1987. *The Moral Limits of the Criminal Law Volume 1: Harm to Others*. Oxford University Press.

Gewirth, Alan. 1978. *Reason and Morality*. The University of Chicago Press.

Greenspan, Patricia. 2010. Making Room for Options: Moral Reasons, Imperfect Duties, and Choice. *Social Philosophy and Policy* 27(2): 181–205.
Igneski, Violetta. 2001. Distance, Determinacy and the Duty to Aid: A Reply to Kamm. *Law and Philosophy* 20: 605–16.
Kamm, F. M. 2000. Does Distance Matter Morally to the Duty to Rescue? *Law and Philosophy* 19: 655–81.
Kant, Immanuel. 1996[1797]. *The Metaphysics of Morals*. Cambridge University Press.
Mill, John Stuart. 1987[1863]. *Utilitarianism*. Prometheus Books.
Murphy, Liam B. 1993. The Demands of Beneficence. *Philosophy and Public Affairs* 22: 267–92.
Pogge, Thomas. 2005. Recognized and Violated by International Law: The Human Rights of the Global Poor. *Leiden Journal of International Law* 18: 717–45.
Quong, Jonathan. 2009. Killing in Self-Defense. *Ethics* 119: 507–37.
Rawls, John. 1999. *The Law of Peoples*. Harvard University Press.
Sen, Amartya. 1999. *Freedom as Development*. Oxford University Press.
Singer, Peter. 1972. Famine, Affluence, and Morality. *Philosophy and Public Affairs* 1: 229–43.
———. 1979. *Practical Ethics*. Cambridge University Press.

Public Reason 3 (1): 50-59

Moral Judgments, Emotions, and some Expectations from Moral Motivation

Mar Cabezas
University of Salamanca

Abstract. I first provide an analysis of the main premises involved in the core of the metaethical debate about the acceptance of an emotional or a cognitive nature of moral judgments and its implications in relation to moral motivation. In order to accomplish this, I start by sketching the main points of the argumentation of Linda Zagzebski (2003) and Kyle Swan (2004). Secondly I suggest that one of the main problematic points of the paradox detected by those authors lays in the assumption of emotions as intrinsically motivating and I develop a critic of a reduced conceptualization of motivation as well as I try to redefine the processes involved in moral motivation, as emotion and cognition, by showing the converging points of view from moral philosophy and psychology. Eventually, I conclude proposing an integral and non-compartmentalized conceptualization of moral motivation and its relation to emotions and cognition, for it could shed some light on the metaethical debate about the nature of moral judgments, externalism and internalism.

Key words: moral judgment, emotion, motivation, cognition, metaethics.

Among other metaethical debates, one of the most discussed issues is the conflict reflected by Linda Zagzebski (2003) and Kyle Swan (2004) with respect to the implications of the relation between moral judgments, motivation and emotions. Given the fact that the relation between the cognitve/noncognitive nature of moral judgment and moral motivation seems to be the core of the metaethical debates over the recent decades, the main premises involved in it will be analyzed in what follows, for they are the root of the widely discussed metaethical question.

The debate between Swan and Zagzebski and, in general terms, the debate between cognitivism and expressionism, as well as between internalism and externalism, has its root in the idea of motivation and our expectations with respect to moral judgments as morally motivating states. We are tent to believe that if we know that X is right, we would directly be motivated to act in that direction, and vice versa. Nevertheless, it seems that moral and human motivation does not work like that, as many everyday's life cases show. Hence, what does this mean? Are moral judgments not motivating? Is internalism false? Are moral judgments cognitive, affective, or both? From an externalism account, especially if it is also based on noncognitivism, the relation between the acceptance of a judgment and the corresponding action is denied. Thus, you can believe that you ought to do X, but this does not directly imply that you are motivated to act so, or that you are going to act so. That is, a judgment does not imply an intention, or an action. From an internalist account, the cases of moral *akrasia* would be explained, as Hare (1963) maintained, because either the moral agent does not really believe in those moral judgments, so she does not really think that she ought to do X, or the agent is not really autonomous to act according to what she judges to be the best reason. However, as Swan sketches, "how could it be that

merely taking the world to be certain way we are inclined to do so?" (2004, 375) These -among others- seem to be the eternal questions of metaethics today.

The problem arises by the combination of three theories referred to different moral aspects. As Zagzebski sets out, if "we expect moral judgments to be both cognitive and motivating" (2003, 105), how can we deal with the idea of moral judgments as propositions with a true value, and concretely, how can we compatibilize the idea that moral judgments are intrinsically motivating with the idea that "when we make a moral judgment we are in a cognitive state" (Zagzebski 2003, 104). As a result of this paradoxical situation, one could choose between cognitivism, non-cognitivism, internalism, externalism, rationalism or sentimentalism, obtaining the following theses, respectively:

1. Moral judgments are propositions.

1'. Moral judgments are expressions of one's attitude.

2. Moral judgments are intrinsically motivating.

2.' Moral judgments are not intrinsically motivating.

3. Emotions are intrinsically motivating.

3'. Reason/cognition is not intrinsically motivating.

At the same time, the debate between cognitivism and internalism is intersected by our assumptions about the -emotional or rational- basis or nature of moral judgments. In order to fit the pieces of morality, metaethics has had a special interest in facing the challenge that different conceptions of moral motivation, the nature of moral judgments and moral knowledge may imply. As a result of it, the debate is eventually articulated in the two following syllogisms:

Premise 1: Moral judgments are emotional.

Premise 2: Emotions are intrinsically motivating.

Conclusion: Moral judgments are intrinsically motivating.

Premise 1: Moral judgments are cognitive (genuine propositions).

Premise 2: Cognition is not intrinsically motivating.

Conclusion: Moral judgments are not intrinsically motivating.

Therefore, the problem turns to be the relation between motivation and cognition, if one assumes that moral judgments are cognitive states and proposition; and the relation between motivation and emotions, if one assumes that moral judgments are expressions of one's attitudes, feelings and emotions. The first syllogism would accept internalism but it would have to face the problem of *akrasia* and weakness of the will. The second syllo-

gism does not have to face the problem of *akrasia* since it does not expect moral judgments to be motivating, but it will have to deny internalism.

Once the basis of the debate are settled down, I would like to point out that -beyond our metaethical position about internalism, externalism, cognitivism, or sentimentalism- what is taken for granted in this debate is that emotions are intrinsically motivating as well as reason or cognition is not, in a Humean style, as the minor premises show. For this reason, it seems that a revision of the term motivation is necessary to clarify what we expect from motivation and what we mean when we refer to it. In other words, the main question seems to be, before we decide what moral judgments are like, what it means to be intrinsically motivating, and whether there is something intrinsically motivating or not.

In fact, it seems that knowing that X is true or feeling and believing that one ought to do X does not always motivate you directly to act in that direction. This can be seen as a problem of moral motivation. However, this will turn out problematic if a reduced idea of what motivation entails is assumed. In other words, this is problematic, only if we expect that if something is motivating, then it has to derive into a coherent action, but perhaps metaethical debates should bear in mind the differences between stimuli, motives, reactions and motivation itself in order to clarify the core of its debates.

For this reason, it seems sensible to maintain, as A. Roskies does, that "empirical evidence can be relevantly brought to bear on a philosophical question typically viewed to be a priori," so that we should not turn our back to scientific and rigorous researches, nor look the other way. Certainly, it is true that "empirical evidence is generally thought to be irrelevant to philosophical theorizing in a number of philosophical areas, including metaethics" as well as it is also true that "one metaethical issue often held to be immune from the empirical concerns is the relation between moral facts and moral motivation" (2003, 51), but it cannot be denied that empirical evidence could provide some clarifications as well as it could also shed some light on these debates. Needless to say, science by itself cannot solve philosophical questions, but it surely can prevent us from assuming false-proven start points. For this reason, I would like to star by clarifying what the terms motivation, emotion and cognition imply from a psychological point of view.

I. WHAT DO WE MEAN BY 'INTRINSICALLY MOTIVATING'?

What we mean when we say that something is intrinsically motivating seems to be a key question whose answer generates part of the conflict. For this reason, a general psychological approach to motivation will be given in what follows in relation to the terms emotion, reason and cognition. As J. Prinz states, "asking how one thing relates to another can lead to discoveries that we would not make if the question had not been asked" (2004, 41).

Following M. T. Sanz (2009), motivation can be seen as an indirectly observable theoretical construction about the factors that modulate an agent's behavior. Therefore, in this first sense, motivation would be a construction built in order to explain the process between motives, that is, the stimuli that trigger an action, and the action itself, so that it

would be a mistake to identify motivation and emotions, as well as it would be misleading to understand emotions as intrinsically motivating if we expect that something intrinsically motivating derives always into an action coherent to one of its stimuli, i.e. an emotion. This conception of emotions, motivation and its direct influence in action would be a reduction of the motivational process. In other words, that kind of expectation would be as wrong as identifying the stimuli and motivation, as indentifying motives and motivation, or, as the genetic fallacy states, it would be as wrong as indentifying the origin of something with that something.

In a second sense, motivation can be understood as a mental process that includes all the factors that lead an agent to act in a certain way, more or less consciously. In this sense, motivation would be an intermediate process between motives and the motivated action. Hence, if emotions were the only intrinsically motivating element of human psyche, that is, if they were the only motives, and if cognition were not intrinsically motivating, we could not act against our emotional motives, so that we could not commit to long term decisions or higher aims, such a finishing a degree, getting on a diet or sacrificing ourselves for political or religious convictions, as it happens in everyday's life.

Consequently, when the metaethical debate about moral motivation and the nature of moral judgments is approached, it should not be forgotten that many different factors convergence in motivation: basic and social needs, biological, cultural, social motives, primary and secondary motives, personality traits, etc., so that the identification of a single element, in this case, emotions, as intrinsically motivating would be -at least- a risky position, especially if moral philosophy desires to give a realistic account of moral agents, of how we develop moral argumentations and how we deliberate and act.

As a result of what have been said, it seems that it makes no sense to distinguish between intrinsically motivating motives and simple motives, since there would not be such a thing as directly motivating motives and 'second class' motives, and given that a simple motive does not directly generates an action or a behavior. The assumption of a thesis like that would imply the denial of the role of motivation in action, for it implies a reduction of our behavior and decisions into a simple effect of some causes or motives that are said to be 'intrinsically motivating' as if there were any other factors that could interact in the motivational process.

In other words, motives, as the precedents of a conduct, do not guarantee the eventual action, since an intermediate process, i.e. motivation, exists. Needless to say, emotions and affective states modulate our motivation and have an influence in our final action, and our decisions and deliberations, but if emotions were intrinsically motivating and cognition were not, we would not be able to fulfill any long term aim, as it has been previously said. Emotions, sentiments and moods can be a motivating stimulus, but we should not indentify stimuli, motivation and responses. It would not be accurate to point out emotion as an intrinsically motivating element, since, in fact, the cause of the behavior, what moves us to a certain behavior, is not emotion, but motivation. Certainly, emotion -taken into account that it involves a valuation- modules the effect of motivation, but it does not

determine it. Bearing this always in mind could shed some light on the metaethical problems mentioned before.

Besides, if we accept that motivation is an adaptive process that prepares us to act in order to survive and/or in order to guarantee the agent's wellbeing and eventually the personal development and growth, that is, if we accept that "motivation is an adaptive process that energizes and guide our behavior to an objective or aim" (Menéndez 2009, 22), and if we believe that moral judgments are intrinsically motivating due to their emotional basis, then we should be disposed to accept that moral judgments and emotions guide our behavior to an aim that provides or guarantee our survival, growth or wellbeing. Then, it would have to be demonstrated that moral judgments guarantee our survival, as well as it would have to be justified how emotions guarantee our personal growth without committing a naturalistic fallacy or without falling in moral determinism.

In this sense, especially in the internalism/externalism debate, it is important to distinguish motivation from the motivated behavior, which would be the result of the motivational process and the motivational traits, that is, the personal predisposition to act in a certain way depending of our personality's traits.

As a consequence, from a psychological point of view, neither emotions are intrinsically motivating, nor cognition is independent of the motivational process. At the same time, neither cognition, nor emotions are irrelevant to motivation, given that both of them can act as a motive and given that these two processes interact and modify each other (Damasio 2000, LeDoux 2000).

II. COGNITION AND EMOTION: RELATIONS AND INTERACTIONS

In what follows, I will try to show some evidences of the interactive relation between the emotional and the cognitive domain, for, as A. Damasio already observed, "the artificial opposition between emotion and reason has been questioned and is not as easily taken for granted" (2000, 13). In fact, it is not only the case that the questioning of boundaries between the classical dualist distinction between reason and emotion is more and more habitual, but that the distinction between cognition and reason is also necessary for moral philosophy debates.

In effect, cognition has frequently been associated to reason, not from psychological or biological approaches, but from moral philosophy ones. Nevertheless, if cognition and reason are distinguished, it could easily been understood how cognition plays an integral role in the emotional process and, therefore, also in motivation.

From a converging point of view between psychology and philosophy, cognition can be defined as the set of processes through which the agent's mind creates a representation of the environment where s/he lives in order to adjust his/her behavior to it[1]. In this

1] For an extended explanation of the distinction between cognition and reason used in this paper see Cabezas (2010, 81).

sense, the term cognition would be referred not only to "tools" as language, but also to other processes, such as attention, perception or memory (LeDoux 2000), through which human mind is able to create a mental representation of the environment. This seems to be indispensable in order to guide and adjust our behavior in a given world, which also includes moral behavior and, consequently, moral motivation. Hence, cognition would not be apart from the motivational process, since it would work as a means to search for ways of satisfying it.

On the other hand, if reason is understood as something different from cognition, it can be defined as a mental mechanism that uses consciousness in order to elaborate abstract thinking -that is, unlinked to reality, and, therefore, not necessary linked to a concrete motivation- through which we can build other possible worlds, instead of representations of the real world. Hence, with respect to moral motivation, and if we accept, as Watson states, that "one's aim in deliberation is to make a commitment to a course of action by making a judgment about what is best (or good) to do" (2007, 175), then maybe we could state that reason is supported by cognitive processes, but reason could be dissociated from motivation, since it would not operate in reality, but in abstraction.

With respect to the relation between cognition and emotions, one might sustain that they are different ways of processing information, especially if one is sympathetic to the idea that "cognitive representations are inherently neutral representations, ones that do not automatically trigger particular behavioral responses or dispositions" (Greene 2008, 40). But again this conception entails some problems. Firstly, the idea that emotions implies automatically a behavior is not accurate if we understand "automatic" as "simple" or "direct", since many regions of the brain are involved in the emotional process. If the consequences of emotions were as "automatic" as it is though -beyond universal facial expression and physiological reactions- it would be sensible to expect that we all would behave exactly the same way. Secondly, cognition, as Tucker (1981) distinguished, can generate two different kinds of cognitions, i.e., syncretic and analytic or "cold" ones. In this sense, emotions could also be understood as a type of cognitions, syncretic ones, given that they allow us to acquire a kind of knowledge referred to our inner states.

Thirdly, also in relation to emotions, and now independently of the role we can be tented to give them in morality, it is important to point out some ideas in order to clarify their relation to both, motivation and cognition.

Although as Reber states, "historically this term has proven utterly refractory to definitional efforts" (1985, 234), a general approach and definition can be reached from a psychological point of view, bearing in mind that the term emotion or emotions is referred to a multidimensional process. Thus, emotions could be defined, as Fernández-Abascal and Jiménez suggest as:

> a process that involves a variety of triggering conditions (relevant stimuli), the existence of subjective experiences or sentiments (subjective interpretation), diverse levels of cognitive procedure (evaluative process), physiological changes (activation), expressive and commu-

> nication's patterns (emotional expression), that entails motivating effects (mobilization to action) and a aim, i.e., adaptation to a never-ending changing environment. (2010, 40-41)

According to this approach, emotion would be a process whose function is not other but the analysis of meaningfulness situation for the agent, the interpretation and valuation of them in order to prepare the agent to act in a direction or another. In this sense, motivational processes may derive from emotions, but, as it has been said before, emotion and motivation cannot be equalized. Simultaneously, and although emotions entails a valuation, it is needless to say that they should not be identified with sentiments, that is, the subjective experience of emotion, for a different cognitive implication is involved in sentiments.

Another classic and inclusive approach to emotions that stresses the relation between the emotional and the cognitive domain is the Kleinginnas' one:

> Emotion is a complex set of interactions among subjective and objective factors, mediated by neural/hormonal systems, which can (a)give rise to affective experiences such as feelings of arousal, pleasure/displeasure; (b)generate cognitive processes such as emotionally relevant perceptual effects, appraisals, labeling processes; (c) activate widespread physiological adjustments to the arousing conditions; and (d) lead to behavior that is often, but not always, expressive, goal-directed, and adaptative. (Kleinginna and Kleinginna 1981, 355)

Nevertheless, it is not only the case that emotions generate cognitive processes, but that cognition is involved in the emotional process in order to elicit an emotion, as it will be shown in what follows.

Certainly, especially in relation to the metaethical problems concerning moral motivation, it cannot be ignored that there is a relevant amount of researches about the interactions between emotion and cognition -which, as it has been marked, should be distinguished here from reason-. As J. Greene maintains, "the present fMRI data support a theory of moral judgment according to which both cognitive and emotional processes play crucial and sometimes mutually competitive roles" (2004, 389).

On the one hand, emotions optimize cognition, for they modulate basic cognitive processes. It is true that the amygdale plays a central role in the emotional domain, but it is important to bear in mind that this is not the only cerebral region able to assign emotional valence, so that emotion and cognition should be considered complementary, instead of independent or antagonist processes. Thus, following Ferrandiz's thesis (1986), one could affirm that it can be affection -which is referred to the simplest element of the emotional domain, that is, to experience of pleasant and unpleasant situations- without the cognitive component, but it cannot be emotion without the cognitive element.

As a result of it, it seems that emotion and cognition maintain a relation that differs from what our traditional intuitions about them had guessed, for it has turned out to be that the complexer and opener a cognitive problem is, the more that emotion is involved. According to Forgas (1995) and his affect infusion model, the more multifaceted the problem to solve is, the more that the affective state would be incorporated and used as a

tool to find new strategies to solve it. Hence, when we face a problem that requires a high effort in order to find a solution and the strategies to find a solution are not fixed, emotionally meaningful information would be integrated in the cognitive process, given that those complex cases require a heuristic and creative system of processing information. This would be due to the fact that the emotional experience is progressively integrated according to cognitive development of the agent (Lane et al. 1990).

In fact, it seem that it is not only that emotion influences memory and attention, which are activities related to cognition, for we process information that is "affectively congruent with our mood" (Bower 1981, 129), but that a positive emotional state facilitates cognitive flexibility and decision making in complex cases, whilst negative emotional state makes us more reflexive and less daring agents (Domínguez and García 2010, 217). This is surely very relevant to approach the metaethical problem about the nature of moral judgments, motivation and their relation to emotions and cognition from a less compartmentalized perspective.

On the other hand, cognition can be understood as an integrated part of the emotional domain (Eich et. al, 2003), not only because both processes share some cerebral regions, such as the hippocampus, the orbitofrontal cortex or the prefrontal ventromedial cortex (Kringelbach 2005, Rolls 2000), but because, as Clore and Ortony (2000) stressed out, emotions occur as a consequence of the activation of the cognitive process related to the significance of the situation. This would, therefore, explain why individual differences with respect to the responses to identical situations exist, also including morally relevant situations.

III. CONCLUSION

Although evidently many consequences from the previous thesis here defended can be derived, the developing of those metaethical questions will be a challenging task for the future which unfortunately exceeds the limits and aims of this paper. Here I have presented arguments and data against a simplistic or reduced account of moral motivation, cognition and emotion in order to try to offer difference perspectives of analyzing some of the hottest debates in moral philosophy and metaethics.

As a result of what has been exposed through this paper, and in relation to the metaethical problem discussed by Zagzebski and Swan, I am sympathetic to Kennett's point of view, although there should be taken into account that when she mentions the affective domain, she is referring to the emotional domain, but not only to affects:

> The terms of the debate between rationalists and sentimentalists must be modified. Recent evidence on moral development from the social and cognitive sciences and from psychopathology does not endorse the philosophers' traditional distinction between the affective and the cognitive, or their attempt to locate morality wholly in one or other domain. (2008, 259)

Taking into account an interdisciplinary outlook from psychology and moral philosophy, and bearing in mind the initial paradox, it could be affirmed that we go through cognitive and emotional processes that occur simultaneously in our brain and, what is more, that our mind works with these two kinds of information, without establishing a fixed boundary. As LeDoux maintains, "the terms "cognition" and "emotion" do not refer to real functions performed by the brain but instead to collections of disparate brain processes" (2000, 129). In this sense, an interactive approach can be defended, given the effects of emotions in some tasks, such as memory or attention; and given that the emotional response depends on the cognitive valuation that we make of it (Eich et al. 2000).

For all these reasons, it seems sensible to conclude in relation to the metaethical question that both, cognition and emotion may play a part in motivation, but neither one by itself determines human behavior and moral decision making.

marcabezas@usal.es

REFERENCES

Bower, Gordon H. 1981. Mood and memory. *American Psychology* 36 (2): 129- 48.

Cabezas, Mar. 2010. Autonomía y emocionalidad en el agente moral. *Factótum. Revista de Filosofía* 7: 76-85.

Clore, Gerald L. and Andrew Ortony. 2000. Cognition in emotion: Always, sometimes, or never. In *Cognitive neuroscience of emotion,* edited by Richard D. Lane, and Lynn Nadel, 24-61. New York: Oxford University Press.

Eich, Eric, John F. Kihlstrom, Gordon H. Bower, Joseph P. Forgas, and Paula M. 2000. *Cognition and Emotion*, New York: Oxford University Press.

Damasio, Antonio. 2000. A second chance for emotion. In *Cognitive neuroscience of emotion,* edited by Richard D. Lane, and Lynn Nadel, 12-23. New York: Oxford University Press.

Domínguez, Francisco J., and Beatriz García. 2010. Emoción y procesamiento cognitivo. In *Psicología de la emoción,* edited by Enrique G. Fernández-Abascal, 185-220. Madrid: Editorial Universitaria Ramón Areces.

Ferrándiz, Alejandra. 1986. Marañón: un antecedente de las teorías cognitivas de la emoción. *Revista de Historia de la Psicología* 7 (3): 3-14.

Forgas, Joe P. 1995. Mood and judgment. The affect infusion model (AIM). *Psychology Bulletin* 117: 39-66.

G. Fernández-Abascal, Enrique, and María Pilar Jiménez. 2010. Psicología de la emoción. In *Psicología de la emoción*, edited by Enrique G. Fernández-Abascal, 7-74. Madrid: Editorial Universitaria Ramón Areces.

Greene, Joshua D. 2004. The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44: 389–400.

Hare, Richard M. 1963. *The language of morals*, Oxford: Oxford University Press.

Kennett, Jeanette. 2008. Reasons, reverence, and value. In *Moral psychology. Volume 3. The neuroscience of morality: Emotion, brain disorders, and developments*, edited by Walter Sinnot-Armstrong, 259-64. Cambridge: The MIT Press and Massachusetts Institute of Technology.

Kleinginna, Paul R., and Anna M. Kleinginna. 1981. A categorized list of emotion definitions with suggestions for a consensual definition. *Motivation and Emotion* 5 (4): 345-79.

Kringelbach, Morten L. 2005. The human orbitofrontal cortex: Linking reward to hedonic experience. *Neuroscience* 6: 691-702.

Lane, Richard D., D.M Quinlan, G.E. Schwartz, P.A. Walker, and S. B. Zeitlin. 1990. The Levels of Emotional Awareness Scale: A Cognitive-Developmental Measure of Emotion. *Journal of Personality Assessment* 5: 124-35.

Ledoux, Joseph. 2000. Cognitive-emotional interactions: listen to the brain. In *Cognitive neuroscience of emotion,* edited by Richard D. Lane, and Lynn Nadel, 129-55. New York: Oxford University Press.

Menédez, Francisco J. 2009. Introducción al estudio de la psicología de la motivación. In *Psicología de la motivación,* edited by María Teresa Sanz, 2-56, Madrid: Sanz y Torres.

Prinz, Jesse J. 2004. *Gut reactions. A perceptual theory of the emotions.* Oxford: Oxford University Press.

Reber, Arthur S. 1985. *The Penguin Dictionary of Psychology*, London: Penguin Books.

Rolls, Edmund T. 2000. On the brain and emotion. *Behavioral and brain science* 23: 219-28.

Roskies, Adina. 2003. Are ethical judgements intrisically motivational? Lessons from 'acquired sociopathy.' *Philosophical Psychology* 16 (1): 51-66.

Sanz, María Teresa. 2009. *Psicología de la motivación*, Madrid: Sanz y Torres.

Swan, Kyle. 2004. Moral judgment and emotions. *The Journal of Value Inquiry* 38: 375-81.

Tucker, Don. 1981. Lateral brain function, emotion, and conceptualization. *Psychological Bulletin* 89: 19-46.

Watson, Gary. 2007. The work of the will. In *Weakness of will and practical irrationality,* edited by Sarah Stroud, and Christine Tappolet, 172-200, Oxford: Oxford University Press.

Zagzebski, Linda. 2003. Emotion and moral judgment. *International Phenomenological Society*, 66 (1): 104-24.

Public Reason 3 (1): 60-72

Ontology and the Paradox of Future Generations

Dennis Earl
Coastal Carolina University

Abstract. The following three propositions are inconsistent: (I) We have moral obligations to future generations, (II) Future generations do not exist, (III) In order to have moral obligations to X, X must exist. All three propositions are *prima facie* plausible. There are really two paradoxes here, one for obligations involving moral rights, and another for moral duties. The paper argues that (II) and (III) are true, thus (I) is false—we have no moral obligations to future generations. The paper considers the available views on the ontology of future generations, as well as various versions of Parfit's person-affecting principle, by way of defending the plausibility of (II) and (III), respectively.

Key words: future generations, future people, possible people, person-affecting principle.

I have various moral obligations to people in the present. I am obligated to see to it that my five-year old daughter is fed, loved, played with, tended to when she is upset, and so on. I am also obligated not to poke her with pins or otherwise cause her unnecessary immediate harm. I also have various obligations to presently existing people with respect to the future. I am obligated to see to it that my daughter receives her vaccinations on time, is not left in the care of adults of questionable character, and so on, where such obligations involve the prevention of future harms to a presently existing person. Furthermore, I presumably have obligations to her that involve the future of eighty years from now. I ought to plan for her college education, for instance, since that seems necessary for living the good life by age eighty. I also ought to vote for political candidates that share my views on how the future ought to go, for my daughter's sake. I also ought not waste natural resources or pollute the environment, and I should encourage others to think similarly. Such obligations are all by way of my being obligated to prevent harms from occurring to my daughter in the future of eighty years from now. This much seems uncontroversial.

Suppose my daughter will have a daughter herself. Do I have moral obligations to her, my granddaughter? *Prima facie*, it seems so—all of my obligations to my daughter of age eighty five would seem to apply to my granddaughter as well. Just as I ought not waste natural resources due to its harm to my daughter eighty years from now, it seems I have the same obligation to my granddaughter who will also exist eighty years from now. By extension, it seems I have many obligations to whole generations of people who don't exist now, but will exist in the future. But there is an important difference between my eighty-five-year old daughter in 2091 and my granddaughter in 2091: My daughter exists now, and my granddaughter doesn't. My granddaughter's nonexistence seems to discount her from my having obligations to her, one might infer, since we have no obligations to nonexistent things generally. I have no obligations to my presently existing sons, since I have none of those, and one might think that since my granddaughter doesn't exist, I have no obligations to her either. One might then infer that by extension, I have no obligations

to whole generations of people who do not exist. And that is exactly the conclusion I shall draw here.

I. TWO INCONSISTENCIES INVOLVING FUTURE GENERATIONS

We are caught in a paradox, it seems: It seems we have moral obligations to people in the future, and those people do not exist. Yet it also seems that in order to have obligations to something, that thing must exist. Put more formally, it appears that we are confronted with two different versions of what one might call *the paradox of future generations*. One of the paradoxes is formulated in terms of our having duties to future generations, and the other is put in terms of the rights of future generations. The duty-based version of the paradox runs as follows.

(I) We have duties to future generations.

(II) Future generations do not exist.

(III) In order to have a duty to *X*, *X* must exist.

Claims (I)-(III) are inconsistent, it appears, and thus at least one of them is false. Yet each of (I)-(III) looks plausible when isolated from the others. The rights-based version involves a similar collection of claims:

(I)' Future generations have rights.

(II)' Future generations do not exist.

(III)' In order for *X* to have rights, *X* must exist.

(I)'-(III)' also appear to be inconsistent, and thus at least one of them must be false as well. And just as with the first set of claims, each of (I)' -(III)' has some degree of plausibility.

The two paradoxes are independent. For instance, to some authors (I)' would present itself as a fairly obvious claim to reject.[1] But even if the second paradox has a resolution, (I)-(III) still are problematic. For one might take the view that we have duties to future generations, even if they have no rights, much like would be the case if we have duties to the natural environment itself, though perhaps it has no rights of its own.

My own thesis here is that with respect to both paradoxes, the first claim in each set is false. We have no duties to future generations, and future generations have no rights. The argument for these claims is fairly simple. The respective sets of propositions are inconsistent, and the second and third claim in each set is true. Thus the first claim in each set is false. Yet others might prefer to reject the second and/or third claim in each set, and such options are the focus of the remainder of the paper. In trying to resolve the paradoxes

1] See Macklin 1981 for a defense of such a denial, and see Partridge 1990 for criticism. See also de-Shalit 1995, Ch. 5 for discussion.

by rejecting (II) (which is the same as (II)', of course), one must defend a view of the ontology of future generations themselves, where such a view takes future generations to exist. Those options are explored below, though my contention is that all of them fail as plausible views of the ontology of future generations. Others might aim to resolve the paradoxes by rejecting (III) and/or (III)' by way of defending a different principle governing what an entity's ontological status must be in order to be the object of duties or the bearer of rights. I aim to show that one defense of the most obvious alternatives to (III) and (III)' fails to require the falsity of (III) and (III)', and thus the paradoxes are not met by that strategy. Less-than-obvious alternatives to (III) and (III)' are neither offered nor considered here, with the challenge being left to defenders of those alternatives to make their case.

Some notes by way of clarification are in order. First, different senses of the term 'future generation' will be employed below, depending on different views on the ontology of them. Nevertheless, throughout the paper the term refers only to collections of future people, however they should be characterized, where those future people do not also exist in the present. It seems that the term is used often enough to include presently existing beings too—we are members of both a present *and* a future generation, one might think—though for this paper I only speak of future generations that are remote in the sense of not overlapping with the present generation. Such remote or distant future generations do not include any of us. Second, 'exist' is meant in its straightforward, usual sense—To say that *X* exists is to say that *X* exists *now*, or that *X persists through time*, including the present. At least *prima facie*, future generations do not exist in this sense.

II. STRATEGIES FOR ADDRESSING THE PARADOXES

There are numerous ways one might address the paradoxes. One that I will not address at length is the possibility that the paradoxes arise due to some equivocation that takes place in (I)-(III) and (I)'-(III)', respectively, and thus that the sets of propositions in question are consistent after all. I set such strategies aside here, and grant that my overall conclusions might need to be rethought if such a maneuver is available. I intend to stick instead to the strategy of considering denials of one or another member of (I)-(III) and (I)'-(III)'.[2]

First option: Deny (I) and/or (I)'

Though my own preferred way to resolve the paradoxes is to deny the first claim in each set of propositions, it is worth noting that denying (I) and (I)' is to claim that we have no obligations to particular future generations and their members. (I) and (I)' are *prima facie* plausible claims, to be sure. In fact, to deny that we have moral obligations to future

2] In my concluding section, I consider drawing a distinction between two general types of duties, but this is not to suggest a solution to the paradoxes (or at least the duty-based one) based on exposing an equivocation in the statement of the paradox itself.

generations would seem to give up on the prospects for obligations to prevent or avoid bringing about various kinds of harms that will occur in the distant future. Surely there can be such obligations, so denying (I) is counterintuitive, one might think.

Perhaps one might deny (I) (and/or (I)') and still hold out for a view allowing for obligations involving harms that will occur in the distant future. For one might reject the notion of an obligation to future generations in favor of obligations to *future people* (construed as people who do not exist now, but will exist in the future), or at least to espouse a view allowing for obligations to future people without obligations to *generations* of them. But the same sort of paradox threatens. For it seems inconsistent to hold that (i) we have moral obligations to future people, (ii) future people do not exist, and (iii) in order to have a moral obligation to *X*, *X* must exist. Furthermore, given that future generations themselves are defined in terms of future people, with different notions of the ontology of future generations tied to different notions of the ontology of future people, such a strategy of denying (I) and/or (I)' fails to resolve the basic logical problem. Readers might be incredulous at this point, thinking that surely, surely we can have obligations to distant future generations that admittedly do not exist. The final section of this essay will attempt to take some of the sting out of this, while preserving the basic intuition that I have to some degree as well.

Second option: Deny (II)

The second option seems to be the most popular, and that is to deny claim (II) (and/or (II)', though for the remainder of this section I'll only make explicit mention of claim (II)). Denying (II) entails holding that future generations exist. There are a number of further options here, corresponding to different views on the ontology of future generations themselves. That is, if one aims to escape the paradox by denying (II), one must defend a view of future generations that treats them as existent things. On such a view, 'Future generations exist' or 'There are future generations' has to come out true.

The options, as I see them, are as follows. A particular future generation is either

(a) a collection of future people existing at some particular future time *t*, where none of those people exist presently, though they are nevertheless real,

(b) a collection of people potentially existing now, who will exist at a future time *t*,

(c) a collection of ontologically possibly existing people, where those people will exist at a future time *t*,

(d) a collection of imaginary people existing at some future time *t*,

(e) a "useful fiction," in the same sense as instrumentalists treat theoretical terms in the sciences, or

(f) a set, considered as an abstract entity, of "placeholders" or "offices" that are not filled by any presently existing people, but will be filled by some people at a future time *t*.

My contention is that all of options (a)-(f) fail as views that allow for denying (II). That is, none of (a)-(f) allow for future generations to exist in any meaningful sense such that they can be the objects of moral duties, nor do they allow for future generations to exist as bearers of moral rights.

Strategy (a). On the first option, one takes a static or eternalist view of time to be correct, and future generations thus can exist just as the present and the past does. Since all times are equally real on such eternalist (or four-dimensionalist) views, future generations (and their constituent future people) are as real as you and I are. Given that they are real, there is no ontological problem with respect to our having moral obligations to them, one might think—we can have duties to particular future people, and those future people can have rights.[3]

One might grant the move, and take the remainder of this essay conditionally: If eternalism is correct, then the paradox might be solved by rejecting proposition (II). Eternalism is certainly not the only view of time, of course. And some of the others, if true, are not friendly to rejecting (II) since they reject the thesis that the future is real. A dynamic view of time such as presentism is in that camp, holding that the present exhausts everything that is real. On that view, future generations neither exist nor are real in any meaningful sense of the term. Another dynamic view, such as that of Tooley (1997), holds that the past and present are real but the future is not. On such a view, once again future generations neither exist nor are real in any meaningful sense.

Yet I can set aside the dispute between static and dynamic views of time, which is obviously beyond the scope of the present project. For there is an equivocation being exploited by strategy (a). Consider 'exist' in (II). '*X* does not exist', in the sense of 'exist' in (II), expresses the proposition that *X* does not *presently* exist. Suppose I say 'My eighty-year-old daughter exists'. Say I insist this is true since I think a static view of time is true, and because I have good reason to believe that my daughter will live to be eighty five. Such a claim obviously equivocates on 'exist', for it is not the sense of 'exist' being used in (II), and by extension strategy (a) fails.

Strategy (b). On this option, future generations are sets of potentially existent people, where such potential people are currently existent. A particular future generation that

3] For example, Quine (1987) embraces this strategy: "[T]he four-dimensional view resolves the dilemma [namely our paradox of future generations]. On that view, people and other things of the past and future are as real as those of today, where 'are' is taken tenselessly as in 'Two and two are four'. People who will be born are real people, tenselessly speaking, and their interests are to be respected now and always (74-75)."

will exist at a later time *t*, then, is a collection of presently existing potential people[4] that will exist at *t*.

There are two difficulties to consider with respect to strategy (b). First, there look to be rather serious problems concerning personal identity here. It might seem unobjectionable to deny beings such as month-old fetuses the status of being actual persons, but grant them the status of potential persons. As such potential persons are nevertheless actual things, (III) does not rule out having obligations to such things. Similarly, (III)' does not rule out their having rights. One might then think that the same sort of consideration applies to members of distant future generations as well.

But for members of such future generations, who at any rate have not even been conceived yet, it is difficult to see how such beings could exist now as potential persons. The *matter* of those beings (though not in the Aristotelian sense of 'matter') is existent now, surely, as various collections of presently existing matter scattered around the universe will be organized into persons in the future. But we have no obligations to such collections of presently existing matter, and nor do such collections of matter have rights.

The second problem is that strategy (b) assumes that potential but nonactualized people would be such that we have moral obligations to them as if they were actual already. Take the rights-based version of the paradox. For resolving that paradox, the strategy assumes that potential but nonactualized people have the rights had by *actual* people. But the assumption is incorrect. I am potentially a septuagenarian. In my state, senior citizens receive a 1% sales tax discount with the presentation of proper ID—that benefit is their right to claim. I cannot claim that benefit, since I am presently under age forty. It will do no good for me to complain to the cashier that I am *potentially* a septuagenerian, and thus in some morally relevant sense I *am* a septuagenerian who has the right to the discount. Similarly, the cashier has no duties to me qua potential septuagenerian to grant me the discount, for the simple reason that I am not actually a septuagenerian. My potentially being a septuagenerian does not confer the rights had by septuagenarians on me, nor does it entail that others have septuagenerian-related duties to me. So strategy (b) fails.

Strategy (c). Perhaps the most common view of future generations (at least in the literature) is that future generations are sets of ontologically possibly existing people, where such people could exist in the future, given the present state of the world.[5]

Following Carter (2001), future generations cannot be sets of *logically possible people*, for such people need not be causally related to the present, actual world. My logically pos-

4] On potential persons, see Warren 1981, in Partridge 1981. Of course, what it is that presently exists that is potentially a person is unclear—e.g., on Aristotelian notions of potentiality, the only things that would be potential people would be those "seeds" of people that have been "planted" and undergone some sort of change. Without some sort of qualifications like these, nearly *anything* could count as a potential person. This difficulty needs resolving in order for strategy (b) to succeed.

5] The treatment of a future person as a species of a possible person originated with Parfit, at least according to Macklin (1981). The terminology of ontological possibility, and the characterization of it given here, is from Carter 2001.

sible future harems, for instance,[6] may include hundreds of members, but unless such possible people are causally related to the actual world, I have no obligations to them. In short, the mere fact that it is possible that various people will exist in the future entails no moral claims with respect to them, unless such possibilities are consequences (or at least likely consequences) of present states of affairs.

Future generations also cannot be sets of *epistemologically possible people* either, for much the same reason. An epistemologically possible person is a person who might exist, for all I know. But what might happen, or who might exist, *for all I know,* need not be causally related to present states of affairs at all. Again, without such a connection, such possibilities are irrelevant to our present obligations.[7]

Now to how strategy (c) fails. Conceiving future generations as sets of ontologically possibly existing people still seems to admit them to be nonexistent, unless one equivocates on 'exists', thus making the strategy not a denial of (II) at all. To deny (II), one must take such possible people to be existent (or real, if one draws a distinction between what is real and what is existent). Thus the strategy is committed to possibilism, where according to possibilism, there are nonactualized possible things in addition to the actual ones. The view is far from new (its most well-known proponents being Alexius Meinong and David Lewis). Criticisms of the view are also far from new—one important criticism, tuned to the present discussion, is the so-called *problem of relevance*. If the objection is decisive, then strategy (c) fails.

The events in some possible world have no relevance to the events here in the actual world, the objection begins. Even if in some possible world, there is a being that looks like me, has my name, has my characteristics, and so on, and is also the king of England, that is irrelevant to the modal facts about me. Furthermore, it seems absurd to think that we in the actual world can have duties to those individuals existing in other possible worlds. Again, there may be someone in another possible world that looks like me, has my characteristics, and so on, and that person has a quite sizable harem (and perhaps is the king of England too, for that matter). I have no obligations to the members of that harem existing in another possible world—that seems obvious. And even if they have rights, it is absurd to think that *I* can violate their rights from here in the actual world.

One might think that in such cases that really is *me* in those other possible worlds, and so I can have various duties toward members of the harem in question, and I can really violate their rights. But this is to presume that I am a transworld individual, existing in multiple possible worlds, and there are no such things. That being in another possible world that is the king of England with the sizable harem is not *me*—it is someone else, even on a possibilist view of things.

6] The example is from an unpublished paper by Stuart Rosenbaum (qtd. in Partridge 1990).

7] Again, see Carter 2001, 434.

For other actual beings (such as actual middle eastern princes and so on), such future harems might be quite likely,[8] and such appeals to possible people makes more sense. Yet even in those cases (1) they are still possible beings, not actual ones, and it is impossible that *they* will become actual. They are possible beings residing in another possible world. (2) If those possible harem members are the same people as the ones who are later actual harem members, then when members of those likely harems become actual, it would mark a point where a nonactualized possible being, existing in another possible world, *becomes* a being in the actual world. Possible beings do not become actual beings, strictly speaking, since once again, nonactualized possible people are isolated from the actual world.

One also cannot appeal to causal relations here, perhaps by proposing that various possible states of affairs are the effects (or likely effects) of the events here in the actual world. For again, on standard versions of possibilism, possible worlds are isolated from one another. A cause in the actual world cannot have an effect in a world distinct from the actual one, even if that possible world contains ontologically possible people, in Carter's sense. Thus strategy (c) fails.

Strategy (d). The next strategy takes future generations as sets of imaginary people. Though some authors have taken this tack with respect to the language used to address the problem at hand (such as Partridge (1990)), I doubt that they truly mean to say that future generations and future people are *imaginary*. First (as with the preceding strategy), imaginary things are nonexistent, and hence such a move is not to deny (II). Yet suppose one granted for sake of argument that imaginary things enjoy some sort of existence, different from the status that you and I have, but somehow existent. What sort of existence might that be? The most obvious answer is that imaginary "things" exist as mental representations of those things, and thus that imaginary "things" are mental particulars on the order of ideas. Just as Santa Claus and unicorns don't enjoy any extra-mental existence, they might be said to "exist in the mind". But to take such mental particulars as things that have rights, or as things to which we have duties, is a category mistake. My mental representations have no rights, and neither you nor I owe my mental representations anything.

Perhaps (d) might be blended with the strategy of treating future generations as being constituted by possible people, insofar as one might hold the view that identifies nonactualized possible people with imaginary people. But this runs into the difficulties just mentioned concerning treating imaginary people as mental entities, and one more: '*X* is a possible person iff *X* is an imaginary person' is false. For what is possible is not the same as what is imaginary. Some things are possible but not imaginary, due to our own cognitive deficiencies in being able to imagine them, and due to some possible states of affairs' not being represented imagistically. Moreover, some things are imaginary but not possible, as some sorts of artwork might suggest.[9]

8] See Partridge 1990, 52.

9] See Sorenson 2002 for discussion of the last point.

Strategy (e). The next option takes future generations to exist as "useful fictions," much like an instrumentalist treatment of theoretical entities in the sciences. Yet if such future generations "exist" as useful fictions, they still are fictions, and hence do not exist. This would fail to be a denial of (II), and thus would fail to address either paradox. But it seems that however it is that fictional entities might exist, it seems false that one could have moral obligations to them. Extending our obligations to fictional entities is odd indeed—one cannot violate Hamlet's rights, nor do we have any duties to preserve the genetic stock of unicorns.

A more subtle objection is that the analogy being drawn with theoretical terms in the sciences breaks down. For while it is true that theoretical language in the sciences can still be useful for making predictions even if such language fails to correspond to reality, the usefulness of the language of future generations is not of that sort. The analogy in question can be stated thusly: Electrons are to future people as predictions about the future are to statements of obligation to future people. Statements about future states of the physical world and statements of obligation to future people are indeed both about the future, but there the similarity ends, it seems. But even if the analogy is good, it still follows that future generations turn out to be nonexistent, just as electrons turn out to be nonexistent on an instrumentalist treatment of them.

Strategy (f). The last option to consider treats future generations as abstract sets of placeholders or offices not currently filled, but that will be filled in the future. The idea is that the sets of offices themselves are presently existent, even if they are not presently filled, and thus can be objects of moral obligation. The strategy is committed to our being obligated to something not constituted by things that presently have rights, or that presently are such that we have duties to them. On (f), we can be obligated to things that are not persons, and not even cognizant beings, at all. However, if one recasts the overall discussion here in terms of moral *standing*, instead of in terms of rights and duties, then the strategy is more plausible. For consider other things that would seem to have moral standing, or moral worth, or at any rate can be *harmed* in a morally significant sense, even if they are not themselves fully constituted by actual moral persons. Perhaps corporations, universities, the natural environment, the presidency, and nation-states fall into this category, along with many others. So while the strategy is unusual, to say the least, it doesn't fail on the grounds that it posits objects of obligation that are not persons.

However, (f) takes future generations to be abstract entities, and it *is* counterintuitive to think we can have duties to abstracta, or that abstract entities have rights. If some of them do have moral standing, that needs further defense prior to accepting (f) as a viable strategy for addressing the paradoxes. Numbers, properties, and propositions are all abstract entities, according to some views, and such necessarily non-spatiotemporal entities are not in the same category as those things to which we can have moral obligations.

There is a second complaint about (f). When a given future generation becomes a present generation, this would mark a transition where an abstract entity becomes a spatiotemporal one, and this seems impossible if there is to be real continuity between a fu-

ture generation and its becoming present. But such continuity is necessary if it is to be the case that when the future arrives, we will have *lived up to* our obligations to *those* particular future generations. Otherwise we will have lived up to obligations to something else, not those particular future people. Given this and the previous objection, strategy (f) is an implausible means of defending the existence of future generations.

There may be other strategies in addition to (a)-(f), but I take it that I have exhausted the available options—I take it that all views on the ontology of future generations fall into one or another of the categories above. Since all are exposed to decisive difficulties, what remains is the thesis that seemed plausible at the outset: Future generations do not exist.

Third option: Deny (III) and/or (III)'

This strategy rejects the intuitively plausible principle that a thing has to exist in order for anyone to have any obligations to it. But the principles captured by (III) and (III)' are principles that don't just seem reasonable—they seem *very* plausible, and as such would need more plausible principles to take their place if one is to reject them.

So what might be put in place of (III) and (III)'? Perhaps the following:

(III)* In order to have a duty to *X*, it must be that either *X* presently exists or that *X will* exist.

(III)'* In order to *X* to have rights, it must be that either *X* presently exists or that *X will* exist.

To be clear, what follows from this is that if it is true (now) that *X* will exist, then presently existing people can have obligations to *X*.

Let the focus here be on (III) and (III)*, since similar considerations apply to (III)' and (III)'*. What case might be made for (III)*? Insisting on its truth begs the question, and the principle is not *a priori* intuitive. So what might the line of thought in favor of (III)* be? One argument might appeal to a relatively well-agreed-upon principle in metaethics, for both (III) and (III)* are quite similar to various versions of the person-affecting principle:

(PAP1) For an action to be morally significant, it must affect persons who actually exist.[10]

(PAP2) For an action to be morally significant, it must affect persons who actually exist or who will exist.[11]

(PAP3) For an action to be morally significant, it must affect persons who actually exist or who are very likely to exist.

10] See Baier 1990.

11] See Carter 2001.

One might also consider non-anthropocentric versions of (PAP1)-(PAP3), formulated not in terms of persons, but in terms of things. For instance, it also seems intuitive that in order for an action to be morally significant, it must affect *something* that has moral standing. One might add two variations on the person-affecting principle put in terms of moral standing.

(PAP4) For an action to be morally significant, it must affect things that have moral standing.

(PAP5) For an action to be morally significant, it must affect things that have or will have moral standing.[12]

The idea is that one might find (PAP2), (PAP3), and/or (PAP5) *more* plausible or *more* intuitive than (III), and since those versions of the person-affecting principle are in conflict with (III), (III) must give way. (Principles PAP(1) and PAP(4) are consistent with (III)—accepting them doesn't assist in escaping the paradox.) What must be put in (III)'s place, one might think, is a principle that keeps the spirit of (III), for instance by ruling out our having duties to nonexistent things like unicorns, but is not in conflict with (PAP2), (PAP3), and/or (PAP5).

The difficulty with this line of thought is this. One can grant either (PAP2), (PAP3), or (PAP5) and still accept (III) after all, given that there can be some obligations that are not directed to particular individuals. It is consistent to hold that actions must affect people who do exist or will exist in order for such actions to be morally significant, while at the same time taking the view that something must exist in order to have duties to it. §3 below will address the details of this distinction more directly.

There could be other arguments in favor of rejecting (III) and/or (III)', to be sure. But whatever such arguments might be, they face the following challenges. They must be defended ultimately by premises with greater intuitive support than (III) and (III)'. They must also be defended in a non-question-begging way—i.e., they cannot unwarrantedly assert that one can have moral obligations to nonactualized future people. Finally, such suggested principles must be consistent with both the nonexistence of distant future generations and with our having moral obligations to such future generations.[13]

III. CONCLUSIONS AND FINAL OBSERVATIONS

All reasonable candidates for views of the ontology of future generations fail to support denying the second claim in each paradox. Denying the third claim in each paradox

12] There is a danger that these last formulations are not illuminating, depending on how one analyzes the concept of moral standing. For if a partial analysis is that a thing has moral standing if it can be harmed now, or could be harmed in the future, one has reached principles much like those listed above.

13] Compare my position with that of Johnson 2003. Johnson would accept (II), that future generations do not exist, as well as (I), that we have obligations to future generations. He reformulates (III) by way of reformulating the person-affecting principle, thus trying to meet the challenge laid out here.

has its difficulties as well, for it is not clear how one might defend the alternatives to (III) and (III)' that allow for moral obligations to future individuals that are presently nonexistent. So the denial of the first claim in each paradox appears to follow: We have no duties to future generations, and future generations have no rights.

Incredulous readers should take comfort in the following final observations. First, giving up on duties to particular individuals in the future, as well as denying them rights, does not entail that we have no moral obligations concerning future states of affairs. For it is possible to have no moral obligations to future generations or to future people per se, while also having obligations to do things that will have various positive effects in the future. In other words, one might draw a distinction between two types of moral obligations. There are obligations *to individuals*, and there are obligations *to do* various actions not directed at any particular individuals or groups of them. One might call them *directed* and *non-directed* moral obligations. Obligations of the first sort include obligations to oneself and to other presently existing beings. These include my obligation not to commit suicide, my obligation to behave rationally, my obligation to help the disaster victims as I am capable, and also my obligations to my daughter to alleviate any unnecessary pain she might have. Obligations of the second sort include my obligations not to waste water, fossil fuels, and energy generally, along with my obligation not to pollute the environment.[14]

Second, the three primary traditions in ethical theory allow for our being obligated to do various actions that would be much the same as those recommended by views embracing obligations to particular nonexistent future generations. Aretaic views are friendly to the notion that there are some non-directed moral obligations, for instance, for virtues such as moderation and frugality need not be defined in terms of relations to individuals. Consequentialist theories might still have us obligated to prevent waste of natural resources, etc. in order to avoid harms to presently existing individuals, while leaving aside harms to nonexistent entities such as future generations. Such theories might also have us obligated to avoid causing harms generally, as is the spirit of such theories, but not with any particular future people or future generations in mind. E.g., even if there are no presently existing future people, 'Not acting to reverse the trend of human-caused global warming will cause harms in the distant future' is still true. Adherents of deontological views might take the same approach, holding that we can be obligated to do various things that will impact the future, but with such duties being duties to presently existing individuals instead of future generations or future people.

Such observations should be comforting, and also should serve to deflect a particular sort of objection that adherents of obligations to future generations might raise. For the basic complaint against rejecting obligations to future generations, as was observed earlier, is that the move seems to give up on the possibility of having moral obligations involving the future, and surely we *do* have such obligations. Yet as I suggested here at the end, such obligations can be had without our having obligations to particular nonexis-

14] Pletcher (1981) seems to defend a similar distinction.

tent future generations, composed of particular nonexistent future people. That sort of directed obligation to future generations, as I have argued, is nonexistent.

dearl@coastal.edu

REFERENCES

Baier, A. 1981. The rights of past and future persons. In *Responsibilities to Future Generations: Environmental Ethics*, edited by E. Partridge, 171-83. Buffalo, New York: Prometheus.

———. 1984. For the sake of future generations. In *Earthbound: Introductory Essays in Environmental Ethics*, edited by T. Regan, 214-46. Prospect Heights, Illinois: Waveland Press.

Carter, A. 2001. Can we harm future people? *Environmental Values* 10: 429-54.

DeGeorge, R. T. 1981. The environment, rights, and future generations. In *Responsibilities to Future Generations: Environmental Ethics*, edited by E. Partridge, 157-65. Buffalo, New York: Prometheus.

de-Shalit, A. 1995. *Why Posterity Matters: Environmental Policies and Future Generations*. New York: Routledge.

Johnson, L. E. 2003. Future generations and contemporary ethics. *Environmental Values* 12: 471-87.

Macklin, R. 1981. Can future generations correctly be said to have rights?. In *Responsibilities to Future Generations: Environmental Ethics*, edited by E. Partridge, 151-55. Buffalo, New York: Prometheus.

Partridge, E. 1990. On the rights of future generations. In *Upstream/Downstream: Issues in Environmental Ethics*, edited by D. Scherer, 40-66. Philadelphia, Pennsylvania: Temple University Press.

Pletcher, G. K. 1981. The rights of future generations. *Responsibilities to Future Generations: Environmental Ethics*, edited by E. Partridge, 167-70. Buffalo, New York: Prometheus.

Quine, W. V. 1987. *Quiddities*. Boston: Harvard University Press.

Sorensen, R. 2002. The art of the impossible. In *Conceivability and Possibility*, edited by T. Gendler, and J. Hawthorne, 337-68. Oxford: Oxford University Press.

Tooley, M. 1997. *Time, Tense, and Causation*. Oxford: Oxford University Press.

Warren, M. A. 1981. Do potential persons have rights?. In *Responsibilities to Future Generations: Environmental Ethics*, edited by E. Partridge, 261-73. Buffalo, New York: Prometheus.

Public Reason 3 (1): 73-78

Darwall Versus Raz on Practical Authority

Mark McBride
National University of Singapore

Abstract. Recently, Stephen Darwall (2010) has offered an interesting counterexample to Joseph Raz's (1986) influential *service* account of practical authority. And Raz (2010) has replied. My aim here is, principally, to better understand the principles on which Raz's reply rests. The plan: First, I set out the particular Razian thesis to which Darwall's counterexample is offered – the *normal justification thesis* (NJT) –, and present Darwall's gloss thereon and counterexample thereto. Second, I offer a – hopefully charitable – account of Raz's reply. This account will present two key (epistemic) principles on which Raz's reply seemingly depends. I do not seek to adjudicate on Raz's reply, and in this respect my aims are modest. However, given the importance of this exchange between Darwall and Raz and given the suggestive nature of Raz's reply, simply attempting to get straight on Raz's reply should prove to be a valuable exercise.

Key words: Stephen Darwall, Joseph Raz, practical authority.

Recently, Stephen Darwall (2010) has offered an interesting counterexample to Joseph Raz's (1986) influential *service* account of practical authority. And Raz (2010) has replied. My aim here is, principally, to better understand the principles on which Raz's reply rests. The plan: First, I set out the particular Razian thesis to which Darwall's counterexample is offered – the *normal justification thesis* (NJT) –, and present Darwall's gloss thereon and counterexample thereto. Second, I offer a – hopefully charitable[1] – account of Raz's reply. This account will present two key (epistemic) principles on which Raz's reply seemingly depends. I do not seek to adjudicate on Raz's reply, and in this respect my aims are modest. However, given the importance of this exchange between Darwall and Raz and given the suggestive nature of Raz's reply, simply attempting to get straight on Raz's reply should prove to be a valuable exercise.

I. RAZ'S NJT AND DARWALL'S COUNTEREXAMPLE

Let's begin, then, with Raz's (1986, 53) formulation of the NJT:

> [T]he normal way to establish that a person has authority over another person involves showing that the alleged subject is likely better to comply with reasons which apply to him (other than the alleged authoritative directives) if he accepts the directives of the alleged authority as authoritatively binding and tries to follow them, rather than by trying to follow the reasons which apply to him directly.

1] Certain (not unnatural) readings of Raz's reply would have him making a concession to Darwall or modifying the NJT – perhaps by restricting the NJT's scope of application by ukase. However, as I do not take this to be Raz's intention, a key respect in which my account of Raz's reply is charitable is that it does not involve any such concession or modification.

Darwall (2009, n.19) notes that though "Raz says here that this is the 'normal way' to establish authority, not that it is a necessary, or even a sufficient condition...I propose to understand it [as a sufficient condition]". This proposal of Darwall's makes more sense once we see Darwall's (2009, 146) gloss on what it is, for Raz, to *accept* someone's directives as *authoritatively binding*:

> [B]y accepting an alleged authority's "directives as authoritatively binding," Raz means ...simply that the alleged subject takes the authority's directives as preemptive reasons, that is, reasons that are "not to be added to all other relevant reasons when assessing what to do," but that "exclude and take the place of [at least] some of them." (Raz 1986: 46) Thus A acquires practical authority with respect to B if B would do better in actually complying with independently applicable reasons if B were to treat A's directives as pre-emptive reasons in this sense.

This gloss enables Darwall (2009, 147) to present Raz's NJT as a sufficient condition for authority thus:

> III. If B would do better in complying with independently existing reasons were B to treat A's directives as pre-emptive reasons, then A has authority with respect to B. (normal justification thesis)

And while Raz (2010, 297) re-emphasises the non-sufficiency (and non-necessity) of the NJT being met for authority,[2] this fact is not a dialectical feature of his reply to Darwall. For present purposes, then, we can take satisfaction of the NJT to be sufficient, for Raz, for authority. We can tie the foregoing together by means of a stock example. Suppose a legislator attempts to solve a general "coordination problem" of regulating road traffic by issuing a directive that drivers should drive on a particular side of the road. Plausibly, drivers would do better in complying with independently existing reasons by treating the legislator's directive as a pre-emptive reason in this case. And, plausibly, the legislator thereby has authority with respect to drivers in this case.[3]

Now consider Darwall's (2010, 259 – footnotes omitted) counterexample:

> For example, I assume that I have prudential reasons to provide for my retirement that are independent of any obligations I might have, say, to provide for others whom I am answerable for supporting or of any obligation to support myself. It seems obvious that I do have such reasons and, moreover, that however important or valuable it might be for me to make my own choices, this latter value might not override (or sufficiently inform) the prudential reasons so that it could indeed make sense for me in prudential terms to put myself in the hands of a financial expert and simply follow her directives. Suppose, then, that I would better comply with the relevant prudential reasons if I were to do so and that there are no other reasons, or

2] Raz's (1986, 47) *dependence thesis* does not explicitly feature as a point of contention in this Darwall/Raz debate, viz.: "All authoritative directives should be based on reasons which already independently apply to the subjects of the directives and are relevant to their action in the circumstances covered by the directive." Likewise for Raz's (2006, 1014) *independence condition* on authority, viz.: "that the matters regarding which the [NJT] is met are such that with respect to them it is better to conform to reason than to decide for oneself, unaided by authority".

3] Though our focus will be on Darwall's counterexample, upon introduction of the principles on which Raz's reply seemingly depends, I will advert to our stock example (see nn. 7 and 9 *infra*).

> at any rate no sufficient reasons, for me not to do so. (We should note that the kind of case we need is not just one where I would do better to follow an expert's advice, that is, by treating her as an epistemic authority on the theoretical question of what there is reason for me to do, but where I would do better if I were to treat her as having practical authority over me, hence the standing to issue legitimate directives to me, in other words, as being in a position not just to tell me what I should do, but to tell me to do it.)

For Darwall – and he invites us to concur – this is a clear case of the NJT being met without the financial expert being a practical authority (over Darwall).[4] Now Darwall (2006) has his own reasons, rooted in the *irreducibly second-personal* nature of claims or demands of authority, for denying the financial expert the status of an authority. However, a key strength of Darwall's counterexample is that it does not rest on acceptance of his own irreducibly second-personal account of authority. Instead, it appears to have bite as a counterexample *by Raz's own lights*:

> The exercise of coercive or any other form of power is no exercise of authority unless it includes an appeal for compliance by the person(s) subject to the authority. That is why the typical exercise of authority is through giving instructions of one kind or another. But appeal to compliance makes sense precisely because it is an invocation of the duty to obey. (Raz 1986, 25-6)

And, we may suppose, correlative to Raz's authority's invocation of a "duty to obey" is the invocation of a *right to obedience*.[5] Darwall (2009, 149), however, notes, by way of reply:

> But it should be clear that [the NJT] is not a thesis that entails anything about any *right* to obedience or about any *obligation* to obey, at least as we ordinarily understand rights and obligations.

In sum, Darwall's case has the NJT being met *without the presence of* any (invoked) right to obedience or obligation to obey – features *taken by Raz* to be involved in any exercise of authority.

II. RAZ'S REPLY

Raz's (2010, 300-301) reply to Darwall runs thus:

> I agree with Darwall that his imagined expert has no practical authority over him. It is not entirely clear how this is meant to be a counterexample to my account...But why then does she

4] Cf. the similar *chinese-cooking* case discussed by Raz (1986, 64) and Darwall (2009, 147-8).

5] We should note that Raz (2010, 290) does not take his account of authority to be "in competition with" Darwall's irreducibly second-personal account of authority: "The sins of… [my] account... could be sins of omission. That is [it] may be right in what [it] say[s], but require supplementation. [It may] have to include the claim that only second-personal reasons can be duties, or rules made by authorities... But... I do not think that whatever its faults, my account of authority will be improved by [this] kind of augmentation." Clearly, thus, Raz's account does not rest on acceptance of Darwall's irreducibly second-personal account of authority.

> not have authority? She has epistemic authority. He should believe that if he is to invest in a pension fund he should invest in the fund she designated, and he should believe that because that is her opinion and she is an expert.[6] Suppose that Darwall believes that. In that case, she no longer meets the condition of the NJT. She believes that if he is to invest in a pension fund, he should invest in this particular one, and he believes the same. She does not know what he should do better than he does.
> This leads me to think that the explanation of Darwall's example is that the NJT is not met when the *only* reason to think that an authoritative instruction is correct is that it represent an expert view about what is good to do, a view which is not based on the fact that the expert will so instruct, or has so instructed.

Suppose, then, A is a putative authority over B with respect to Φ-matters. We can distill the following claim from Raz's reply:

> *(KNOWLEDGE-SYMMETRY)* When A instructs B with respect to Φ-matters, the NJT is not met with respect to the specific Φ-matter of A's instruction, if, on instruction, there is *knowledge-symmetry* between A and B with respect to this specific Φ-matter.[7]

For Raz, knowledge-*asymmetry* (in favour of A) is necessary for the NJT to be met: if there is knowledge-*symmetry*, it's not the case that B would do better in complying with independently existing reasons were B to treat A's directives as pre-emptive reasons than by following them directly. Instruction by A to B on a specific Φ-matter, however, can *break* a knowledge-asymmetry. How so? At this point we cannot, for Raz, validly infer *from* A's instructing B with respect to a specific Φ-matter *that* a knowledge-asymmetry between A and B can be broken on this matter. To make the inference valid, for Raz, a further condition must be met. And we can distill the required condition from the following remark from Raz (2010, 300): "However, that ... [A's] directives would be the directives one would give if one knew [everything one needs to know to be authoritative with respect to this specific Φ-matter] and made no mistakes ... is not enough to endow [A] with authority over [B]. [B] needs to be able to know that this is the case".[8] Raz, thus, can be taken to commit to:

> *(SUBJECT-KK)* In order for A to have authority over B with respect to the specific Φ-matter of A's instruction, not only must A *know* what B should do better

---

6] My note: Raz assumes Darwall's case has the expert giving a *conditional* directive, rather than the *unconditional*: "Invest in the fund I designate". To have expertise to give this unconditional directive requires ascribing to the expert a very broad range of knowledge.

7] Raz's (2010, 301) paradigm cases of instructions not-[meeting-this-sufficient-condition-and-thereby-failing-to-satisfy-the-NJT] are those involving "coordination [and] concretiz[ation] of indeterminate boundaries". Thus, for Raz, we must assume, our stock road traffic coordination example (see section I *supra*) does *not* meet *(KNOWLEDGE-SYMMETRY)*'s sufficient condition. Filling out why this is so is a good exercise.

8] Raz is here considering the *unconditional* directive mentioned in n.6 *supra*. But I see no reason why we cannot extract the ensuing general principle.

than B does with respect to this specific Φ-matter, but B must *be able to know* that A knows better.

We can best think of *(SUBJECT-KK)* as a *background necessary condition* on authority, which we can here, in the absence of defeating considerations, *assume is met (prior to instruction)* when coming to test for whether *the NJT itself is met (on instruction)*.[9] While *(KNOWLEDGE-SYMMETRY)* pertains solely to *first-order* knowledge – in Darwall's case, knowledge of the specific financial matter of instruction –, *(SUBJECT-KK)* pertains additionally to (the possibility of) *second-order* knowledge – in Darwall's case, knowledge that the financial expert has superior knowledge of the specific financial matter of instruction. And, for Raz, if *(SUBJECT-KK)* is met – and only if (something like) it is met – the foregoing inference (viz. from instruction to the possibility of knowledge-symmetry) is valid. In sum, *(SUBJECT-KK)* is a (putative) background requirement on authority; however, for Raz, if *(SUBJECT-KK)* is met – and only if (something like) it is met – we can get cases of instruction breaking a knowledge-asymmetry; and, if we do get such cases, the NJT is *not met* with respect to Φ-matters.

Spelt out: As each (further) specific Φ-matter arises, and instruction thereon is given, *modulo (SUBJECT-KK)* being met, the sufficient condition contained in *(KNOWLEDGE-SYMMETRY)* can be (repeatedly) met, such that, if so met, the NJT is not met, and we'd resultantly have no reason to think A has authority over B with respect to Φ-matters *in general*. In sum, for Raz, this is precisely what happens in Darwall's case: though the financial expert fails to be an authority over Darwall, *the NJT is not met in Darwall's case*, and we have a badly formed counterexample to the NJT's sufficiency for authority.[10, 11]

markmcbride1978@gmail.com

REFERENCES

Darwall, S. 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press.

9] This enables preservation of our treatment of the NJT as a sufficient condition for authority. Moreover, I do not see a way of incorporating *(SUBJECT-KK)* (as a necessary condition) within the NJT itself (absent significant modification of the NJT). Cf. also Raz (2006, 1025-26) for more on this *knowability* constraint on authority. And, plausibly, our stock road traffic coordination example (see section I *supra*) meets *(SUBJECT-KK)*'s necessary condition.

10] It has become fairly standard (e.g. Darwall 2010) to object that Raz's NJT is better suited to establishing the legitimacy of *epistemic* (or *theoretical*), rather than *practical*, authority. I don't take a stand on this objection. But, now that we have introduced *(KNOWLEDGE-SYMMETRY)* and *(SUBJECT-KK)*, it is clear that, given these two theses, Raz's NJT will not, without more, classify the (uncontroversial) financial expert an epistemic authority.

11] Thanks to Lee Walters for stimulating discussion, and to a *Public Reason* referee for helpful comments.

———. 2009. Authority and Second-Personal Reasons for Acting. In *Reasons for Action*, eds. D. Sobel and S. Wall, 134-54, Cambridge: Cambridge University Press.
———. 2010. Authority and Reasons: Exclusionary and Second-Personal. *Ethics* 120: 257-78.
Raz, J. 1986. *The Morality of Freedom*. Oxford: Clarendon Press.
———. 2006. The Problem of Authority: Revisiting the Service Conception. *Minnesota Law Review* 90: 1003-1044.
———. 2010. On Respect, Authority, and Neutrality: A Response. *Ethics* 120: 279-301.

Public Reason 3 (1): 79-97

David Friedman's Model of Privatized Justice

Ionuț Sterpan
University of Bucharest

Abstract. David Friedman employs mainstream economic tools to argue for a non-mainstream economic conclusion: the law and the legal system can and should be privatized. First we circumscribe Friedman's model of privatized justice within the map of the anarchist schools of thought. Second, we present how the model would purportedly work in practice. Third, we describe the assumptions of the model. Since the normative force of the model depends on their limitations, when deemed suitable, we point out ways to support them. We sketch a normative defense that owes to Richard Posner, we show why a system of justice focused on efficiency peacefully corrects initial errors and suggest why a system of justice focused of revealed preferences only, gives an unbiased treatment to agents. Fourth, we call in Harold Berman and Stringham and Zywicki's contribution to interpretating the early history of the Anglo-Saxon legal system and North, Wallis and Weingasts' framework for understanding social order. An application of their framework is double-edged. One the one hand it adds weight to Tyler Cowen's objection against the stability of the Friedman model. On the other, it makes a certain reading of history more plausible, namely, that according to which the dynamics of power between the Church and the state allowed for a system of privatized justice, albeit limited and non-transparent. If this reading is right, David Friedman's model proves less utopian than we think.

Key words: David Friedman, privatized justice, market anarchism, anarcho-capitalism.

I. THE MODEL CIRCUMSCRIBED

1.1. Privatized justice

Phrases such as "privatization of justice," "privatization of the law," and "liberalization of the market for law" only make sense if we eliminate the notion of state authority from the definition of law[1]. "Law" and "justice" are part of an effort to subject human conduct to rules (Fuller 1963, 96; Benson 1990, 11), especially under the threat of physical coercion[2]. This effort takes place in three stages: the production of law, the interpretation of law, the enforcement of law. Human efforts directed towards these aims can be conceptualised as economic services: they are "goods" - insofar as they are both desired and scarce.The "privatization of the law" or the "liberalization of the law" designates the privatization or liberalization of these three "protective services" – of producing, interpreting and enforcing the law. The anarcho-capitalist order represents an order of free enterprise and free exchange for this class of services.

1] This study is the result of an extended collaboration with Mihail Radu Solcan, Emanuel-Mihail Socaciu, Laurențiu Gheorghe, and Constantin Vică. Most of that work was carried out at workshops on the *Emergence and Evolution of Social Norms*, a project of the *Research Center for Applied Ethics*, a unit of the *Department of Philosophy, University of Bucharest*, financed by CNCSIS, code TE_61, no 22/2010.

2] The requirement of physical coercion delineates the administration of justice from that of other moral rules.

The freedom of contract for protection entails the freedom to choose not to enter a contract with any agency, no matter which. One requirement to understanding anarchism in general, notjust anarcho-capitalism, is to discard the assumption of hierarchy. Imagine three people in a triangle, where each of them is appointed to be their judge by the other two. Or think of families. Sometimes two members ask a third who is in the wrong. Each appointed judge can judge slightly differently. The assumption of legal uniformity over a territory or domain is also discarded. Individual I1 can stay under legal code A when interacting with individual I2 and under law B when intercting with individual I3. Different solutions can coexist for the same kind conflict over different pears of individuals, and for different domains of application.

1.2. The intellectual map of the anarchist schools of thought

Building on Peter Boettke's classification (2005, 207), we place David Friedman within the analytic anarchism research program, as distinguished from the utopian and the revolutionary research programs.

Utopian anarchism works with assumptions poorly justified by social and human science, about a post-scarcity world dominated by altruism. It is also impregnated with a double aversion to political power and economic inequality. The first modern anarchist William Godwin (2011) and Josiah Warren (2011), publisher of the first American journal of anarchism, sample that thinking[3].

Although revolutionary anarchism in the tradition of Mikhail Bakunin (2011) does not rely to an equal extent on the assumptions of post-scarcity and altruism, it is guided by the same double aversion. It recommends an upsetting of the system of private property *and* political power by violent means.

Another tradition, ignored by Boettke, is manifest in the moral philosophy of political obligation. We may call this "ethical anarchism" and have in mind anarchist-individualists such as Robert Paul Wolff (1998) and John Simmons (1979).

Analytic anarchism employs economic analysis of institutions, an effort to identify the effects of legal rules with the help of rational choice theory. Besides David Friedman, important examples here are Murray Rothbard (1998), Randy Barnett (1998) and Roderick Long (see Long and Machan 1998), while their common precursor is Gustave de Molinari (2011). The terms "market anarchism" (Long) and "anarcho-capitalism" (Friedman) are largely interchangeable. It is analytic anarchism that caught the attention of Robert Nozick (1992) and James Buchanan (2001). The latter formulated different arguments against the anarchist position and they were largely followed by political philosophers and economists. To this day, analytic anarchism is still a marginal research program.

3] See Doherty 2007 for an evaluation of a large number anarchist-leaning authors made from a libertarian perspective.

1.3. Intellectual sources of analytic anarchism

We can distinguish a number of different intellectual sources put to work in the analytic anarchism research program, though none of the original sources are themselves anarchistic proper.

The rejection of the idea of large-scale monocentric legislation, for reasons that have to do with an understanding of equality before the law. Alfred Venn Dicey (2005) praises the universal subjection of all classes to one law administered by the ordinary Courts.

The rejection, from an evolutionary perspective of the idea of large-scale monocentric legislation. Hayekian arguments (1945; 1973; 1988) according to which a law fit for the complexity of human interaction must benefit from local dispersed knowledge can be used in favor of anarchism (Stringham and Zywicki 2011).

The epistemic arguments of the Austrian School regarding the impossibility of rational central planning in the supply of any services (Mises 1991). Absent private property and free exchange, there are no real exchange rates; without real exchange rates, we don't know people's relative valuations of services; missing the knowledge of people's relative valuations, relative prices of centrally planned production and allocation are bound to be arbitrary. The argument could be extended to cover legal services (Solcan 2003).

The Chicago School, which is a mix of neoclassical economics and "the property rights paradigm" rooted in Ronald Coase (1991). If at zero transaction costs, property rights go to he who values them the most, then we have an argument in favor of free exchange of property rights. The argument can be used at a deeper level to argue for a free market for more fundamental legal rules, and in fact this accounts for a large part of what David Friedman is doing.

The Virginia School starting with Buchanan and Tullock (1995) and Tullock (1967). Firstly, it shows the reasons why the alternative to anarchism, the supply of legal services by the state, is inefficient – thus, the appeal of anarchism increases indirectly. Secondly, it reinforces arguments against economic objections formulated against anarchists by people in favor of a more-than-minimal state, within discussions in political science and political philosophy.

Robert Nozick. With the tools of analytical philosophy, Nozick (1992) reevaluates political philosophy from the perspective of negative individual rights and inviolability of individuals.

What sets David Friedman apart within the intellectual family of analytic anarchism? David Friedman makes less use of specific Austrian ideas, today a rather marginal methodology. A most noteworthy feature of Friedmanian anarchism is that he is employing mainstream methods to the benefit of non-mainstream conclusions. Friedman is using Coasian analysis of property rights, economic analysis of law, and neoclassical economics. As a neoclassical economist, Fredman sees economics as a study of human behaviour. He answers questions about the value of goods on the basis of their marginal utility to the consumer and relative demand. Once we base our theory of value on goods

that can actually make the object of choice at certain points in time, then we tie value to marginal (additional) units of goods, thus exploding comparisons between the totality of bread in the world and the totality of diamonds in the world. Classical intuitions that tie value to labor or to the costs of production also disappear.

II. THE MECHANICS OF THE MODEL

In order to see what the system would look like in practice, we need to employ economic analysis of law. Essentially, David Friedman identifies and estimates costs and, on the basis of rational choice theory, he makes institutional predictions.

Following a functional division of labor, the production and interpretation of a legal protocol will constitute the province of legal courts (Friedman calls them "arbitration agencies"). The legal protocol's administration and application will constitute the province of "protection agencies".

Insofar as legal services are protocols of interaction, they only make sense if produced for and consumed by pairs of individuals. Pairs of customers choose their justice providers. This does not mean that these protocols cannot be asymmetric (or bad); they can, even to the point of slavery. What it does mean, is that 'domestic' production and enforcement of law (as in 'each person with their own law') makes as much sense as using a 'private language.' The terms of a legal insurance bought by an individual form a legal code, and that legal code applies to that individual paired with others.

Pairing is the key to the Friedman model. Justice providers sell legal codes together with services of interpretation or arbitration to protection agencies and, after enhancing them with enforcement services, the latter further sell them to pairs of customers. Courts also sell the codes to pairs of protection agencies – only to be resold to pairs of customers of the two respective agencies.

In a free market for justice we can foresee three general characteristics of the model. First, we can expect more peaceful agreements than violent conflicts since the latter are more costly relative to peaceful agreements. In the case of agencies, the conflict is particularly costly, since it chases away clients towards those agencies that prove able to successfully solve comparable problems more cheaply.

Second, we can expect that the applicable codes will become tolerably similar due to pressures for cutting on bargaining costs. In the first instance, the formal realm of legal possibilities is large: it will reflect the diversity of human preferences. There are at least as many possible codes as individual pairs can think of times two, where the multiplication by two reflects the possibility of legal codes asymmetrically applied between individuals. Because of the costs of diversity, the result will nevertheless be tolerable uniformity, due to the costs of negotiation. These are costs of negotiating separately with each individual with whom one could find oneself in conflict. The costs of contracting differently with each new group of people are larger if the clauses that are applicable to the old group could not be applied to those in the new group.

Third, we should expect widespread pre-existing bilateral agreements, as pre-existing agreement reduces bargaining costs. Ad hoc bargains are costly relatively to legal insurances. Practically, protection services get to be protection services against *the eventuality* of conflict.

III. ON THE NORMATIVE CHARACTER OF THE MODEL

3.1. The prima facie case for efficiency

The attractive feature of the model of privatized justice is the efficiency of the output. The rules produced by the system are claimed to be efficient. But why would we be interested in a legal system that maximizes efficiency? Allocating rights and producing distributions of rights for those who pays the most for them on an open market is desirable for two reasons. First, the right to decide whether a particular change occurs goes to the one for whom it matters the most. Second, market allocation reveals an important information to third parties. Rates of exchanges, prices, or information about relative valuations, can only emerge out of rich networks of transactions. Third parties can use the information revealed in other transactions when setting and following their own objectives.

3.2. Value maximization

Friedman uses Alfred Marshall's method of measuring and aggregating preferences, according to which the intensity of preferences is measured by the number of monetary units that agents are willing to give up for the sake of some change in the states of affairs (money reprepresenting the value to them of some other changes). The efficient outcome is the outcome that produces the greatest net positive effect. If John would rather lose 2 dollars than confront some change C, and Mary would pay 3 dollars in order for that change to take place, the change is efficient.

This description seems burdened with the idea of interpersonal comparisons of utility. However, two things can be said to prevent objections against utilitarian suppositions. First, utility must be interpreted in a formal sense. Talk about utility is not necessarily commited to a metaphysical notion of units of happiness. If people behave consistently, their behavior can be modelled as though maximizing the expected value of something, whether they are intending to or not, whether they feel anything or not. Whatever this may mean or involve for anyone in particular context, we can call it utility without being committed to its actual existence (Binmore 2005, 65). Second, at zero transaction costs, compensations take place. At zero transaction costs the difference between a metaphysical commitment to interpersonal comparisons of utility and the methodological decision to only look at revealed preferences fades anyaway: there is nothing to prevent preferences from being reaveled. Under certain assumptions (among which zero transaction costs,

uniform valuations of the dollar across participants and an ability to pay[4]) a money market reveals quantitave information about "intensities of preferences" across agents. It measures, in dollars, how much people want something rather than something else.

A system of privatized justice founded on the ideal of value maximization or wealth maximization would focus in practice on the satisfaction of preferences registered in a market. But the concept of market is here more inclusive than it would appear. It includes both barter markets and hypothetical markets. The value of services exchanged in a barter market could be estimated by looking at the monetary price of substitutes. What are hypothetical markets and why are they included? We must remember that our normative ideal describes a zero transaction costs market. Since the real world manifests transaction costs, a court of justice set to maximize wealth would employ hypothetic market analysis and distribute rights according to their speculations on potential market transactions freed from transaction costs (Posner 1981, 61).

If courts are forced to make entrepreneurial judgements on hypothetical transactions, then we have another argument in favor of an aspect of privatized justice: free enterprise. Do we really have a discovery procedure of efficient rules better than competition? In a recent article Stringham and Zywicki (2011) make a case for a Hayekian anarchism.

3.3. Zero transaction costs

Neoclassical economists usually employ the model of perfect competition to demonstrate that some particular arrangement leads to an efficient outcome. The model of perfect competition shows why on a market with sufficiently many suppliers, competitive pressure is forcing each supplier to reduce the price close to the cost of production. The legal market has, nevertheless, a peculiarity: legal suppliers buy a set of individual agreements or consents for a particular code, which applies to those individuals paired with other individuals. The situation is that of a rich texture of bilateral monopolies. By hypothesis, in a free market for law no one can supply my agreement but myself. In the absence of competition, I will sell my agreement more dearly than it would cost me to 'produce' it. Consequently, concludes Friedman, the model of perfect competition is of little use.

Friedman then appeals to the Coasian model of bargaining for property rights (Coase 1991). The model shows that only legal changes with a net benefit take place at zero transaction costs (Friedman 1994; 1996). In the absence of costs of acquiring information about what and how strongly others desire people identify and exploit the possibilities of improving their legal insurance, and in the absence of costs of negotiation they will exploit these possibilities. These can be either changes that are in the interest of both parties, or changes that are in the interest of one but against the interest of the other party, provided the interest of the party who favors change is stronger than that of the one who prefers the status-quo. Compensations can take place for Pareto improvements.

4] There are hidden assumptions that limit the normative character of efficiency and Friedman admits to them (Friedman 2001, 22). We shall discuss them in what follows.

Imagine two agencies representing their respective litigant clients. One agency is for the death penalty, the other is against it. A transaction between the two agencies occurs starting from an initial distributional point. At zero transaction costs, regardless of whether the initial distributional point specified death or no death the concrete result will be efficient. If the initial distributional point is was so that interagency conflict was solved in a pro-death manner, but now the antideath agency is able to pay enough to compensate the pro-death agency, then the efficient result (no-death) occurs. If the distributional point was already anti-death, then the result stays anti-death. The efficient result occurs regardless of the initial distribution provided that point is specified is some way so that the deal can start off.

3.4. Uniform valuation of the dollar

Maximization of efficiency is an imperfect proxy for the maximization of utility. One reason is that Marshallian aggregation, or interpersonal comparisons of utility made by means of market exchange at zero transaction costs still assume that dollars are valued uniformly by all. It might seem that the dollar has a lower marginal utility for the rich.

How could that be shown empirically, by looking at people's choices? Recall our formal analysis of utility. If we do not commit to a metaphyscal dimension of utility, then my valuing X makes sense only by comparison to my valuation of a Y. As long as there are values that are not registered in a market, it follows that the richer I get, the more worthy non-marketable values get for me by comparison to the set of values that can be bought for money. Evidence would require systematic observation of rich persons increasingly making choices of values that are not registered in a market, not even in the inclusive sense of market. This seems difficult to show.

But assmuning it is possible, Friedman's partial defense here is that when we aggregate large amounts of costs and benefits across large sets of people, the differences in valuations average out. The Marshallian method works well in evaluating legal changes that affect large classes of individuals, such that the group which does want it and the group which would rather do without it do not overlap with the rich and the poor. The objection does not have teeth in situations in which there are as many deviations in the valuation of a dollar on each side (Friedman 2001, 22). The Marshallian method still is appropriate in evaluating what Hayek called general rules.

3.5. Ability to pay

Second, the model assumes away the possibility that at some point some people can value something more than they are able to pay.

How could *that* be shown empirically? Someone might choose to spend an amount of time in jail set to be equal with a $ 10,000 fine for stealing a $3000 necklace for is wife (Posner 1981, 63-65).

If this scenario is possible, then efficiency does not have enough normative force. Efficiency or wealth maximization again fails to approximate utility maximization.

How could the criterion of wealth maximization be defended against these objections? One defense is that we cannot base a system of justice on speculations about utilities. Although these scenarios are possible, we cannot know *when* they are true. These scenarios are burdened with an element of speculation.

From a methodological point of view, if individuals are separate opaque worlds (I can't feel your pain), we refrain from speculating on their interior. From a normative point of view, we refrain from violating these worlds through unwanted interventions for redistributive (deliberately inefficient) purposes. Limiting the purpose of justice to efficiency is the consequence of accepting these normative and methodological contraints. In traditional analysis of happiness, or in traditional talk about utility maximization, what matters is the passive, subjective experience. Analysis of wealth anchors valuation in intersubjectivity. Methodologically, we can easily observe transactions. A focus on exchange is a focus on the interesction between manifest preferences (Posner 1981, 66). Normatively, what matters is the active character of producing wealth. Successfully offered services are let out freely from an opaque world and are freely received in another.

Of course, we cannot hope for a philosophical settlement here. But the advantage of a legal system that concentrates on minimizing transaction costs weakens the objection from the ability to pay: such a system encourages the creation of kinds of services that build upon services. An vast array of possible labor niches for everyone facilitates inclusion in the network of exchanges enabling everyone to pay.

3.6. Two aspects of transaction costs

Transaction costs hinder an efficient output. There are two aspects of transaction costs that the model would confront in practice. The first side is that given a distributional starting point, the cost of a rectifying trasaction is naturally positive. This adds friction to the mechanics of the model. The other side is that the initial distributional point may be itself unstable, unclear or under dispute. This difficulty is more like one that does not allow the mechanical process to start off.

Regarding the first problem, a reasonable claim is that transaction costs have a better chance to be overcome if bilateral bargains are allowed, and that is what privatizing justice means. Private law is efficient because given any distribution between two agencies, if agencies realize any efficient possible change, they will make it happen, with the winner compensating the loser (Friedman 1994). If pairing is the key to understanding the mechanics of the Friedman model, the key to understanding why the output of that model tends to be efficent is that at zero transaction costs, if pairs imagine an improvement, they will realize that improvement.

Bilateral bargains with resolutions that have zero external net effect (especially those that hardly affect third parties at all) will have a real chance of producing an efficient outcome.

Bilateral bargains whose resolutions have a non-zero effect on third parties will have lesser chance of producing an efficient outcome. The more widespread the externalities, the larger the transaction costs and the lesser the chances that the outcome will be efficient.

To understand this, let us take intellectual property rights as an example (Friedman 1996). An efficient fully privatized market for intellectual property law has to deal with three relevant costs. On the one hand, there is the cost of a decrease in intellectual production. Let's say the perspective of such a cost pressures in favor of a rule of intellectual property. On the other, there is the cost of not using sufficiently that which has already been produced plus the increase in the cost of law enforcement. In the anarcho-capitalist model, if John agrees to pay Mary for her intellectual product (their relationship will be governed by the rule of intellectual property), then Mary will produce as long as she is paid. If the products are beneficial to third parties who do not pay, the outcome will probably be inefficient because it does not take into account the valuations of third parties. An efficient rule would have to minimize the sum of all these costs. Why so? Let us presume that the cost to Mary to produce an extra work is over 2 dollars. The work in question would be valued by John and other two people at 1 dollar each. But if only John is a party to the transaction the work is not being produced.

The second side of transaction costs is "the baseline problem" (Friedman 1996). Theoretically, agencies could return at any time to that initial distributional point – a result of a bilateral monopoly bargaining process which includes mutual threats of refusal to agree on any arbitrator. The claim is however that changing the status-quo with aggressive means is relatively costly.

3.7. Levels of rules

A market for legal rules assumes mutual recognition of the baseline and since that baseline is itself describing a rule, we arrive at the idea of several levels of rules.

In an attempt to clarify the idea of privatized justice let us distinguish at least three levels of rules. Transacting rules (or rights) at the first, surface level is not unfamiliar. Let's say according to the official, formal distribution given by the judicial system I have the right to phonically pollute my neighbor but I am selling that right to her, as I am currently in need of cash and she is conducting lucrative sessions of psychotherapy in her apartment. This surface allocative level is today privatized, since my neighbor and I really can choose to re-allocate differently such rights set for us in our initial distributional point. The freedom of transacting at the allocative level is obviously conducive to efficiency. In most cases today, starting from the distributional points offered by the current legal system, parties reach agreements outside courts and reach efficient solutions (Friedman 1996).

We have distinguished two levels of rules, the first, the allocative level, and the second, the formal distributional level. According to this second layer, in my example, I am the one holding the right to decide when I make noise. The official distributional point is fixed today but we can imagine that my neighbor and I could choose among different providers of default distributional points. Choosing legal codes on a market for justice means choosing formal distributions of rights that form starting points in subsequent allocative transactions. Today this level is not privatized but we can imagine it could be. In Shasta County the official distribution by state law defining whether the *range* is open or closed (whether ranchers' cattle can move freely in neighboring farms) is immaterial (Ellickson 1991). Neighbors have created and are following a different distributional level. Ellickson claims the farmers from Shasta County manage to reach a more efficient agreement than through state law.

However, transactions of legal codes at the second level presuppose a more fundamental recognition of rights, they presuppose yet a more fundamental distributional point that describes the terms allowing for that business. Actors operating on the free market for justice must themselves operate within some framework of mutually recognized rules. Thus, a third level must stand as a baseline responsible with maintaining and formulating the second level. This most fundamental (often informal) distributional level represents the power distribution that defines the initial bargaining positions of the very actors operating on the free market for justice.

3.8. The initial bargaining position can incorporate inefficiencies

Recall that example of the disputed property right to decide when your silence can be broken by my singing. If the initial distribution entitles you to silence and you value silence more than I value my own singing next to you, the concrete result will be efficient – silence. If in the initial distribution, I have the right to pollute you with noise, but you value silence more than I value my own singing next to you, the concrete result will be efficient – silence– since at zero transaction costs we exchange property rights accordingly. There will be silence whenever silence has a social value (aggregate value) greater than my singing, regardless of initial distribution of property rights over silence and my own singing).

Let us assume that in the initial (second level) distribution the right to decide lies with the singers, that the social preference lies with silence, and that real world transaction costs are positive. Then the outcome itself (at the first level) is inefficient, since the efficient allocation of property rights is hindered.

To be sure, the number of transactions that might be avoided should be the main factor in the selection of an efficient initial distribution, at any level. But to the extent that the initial distribution and initial bargaining positions are the result of relative force and

rectifying transactions are costly, then inefficiencies at the base level plague the entire legal edifice.

Moreover, the initial bargaining position at the second level incorporates inefficiencies to the extent to which the second level is inevitably the outcome of a balance of power (the outcome of a credible threat game), which is not necessarily the outcome of valuations manifest in market exchange.

3.9. A privatized system of justice corrects in time initial inefficiencies

Should we then go back to the fundamental distributional level and make forceful reparations? The problem with that proposal is that in the absence of a market we lack a method of determining what a reparation is. How can we recognize an overlapping between power and valuations? Imposition of force would exclude the market. The new power elites would be incapable of showing rationally that the new fundamental distribution represents a more exact overlap with valuations. This is a problem not only for any anarchist radical revisionism but indeed, for radical revisionism of any kind.

An authentic correction of initial inefficiencies can take place under an institutional arrangement arrangement that, starting from the status-quo, would reduce transaction costs for all legal transactions, at the first two levels. The market would correct in time the inefficiencies embedded in the fundamental level.

The first reason why we should expect corrections is that favorable but inefficient rules that some obtain for themselves are generating litigation after litigation until the rule is being corrected. The first is that favorable but inefficient rules which are obtained in some way or another generate a series of litigations, to the point that the rule corrects itself (Friedman 2001, Ch.1). If the rule is inefficient at the outset, those who stood to lose had more to lose than those who gained had to gain. The rule corrects itself because the resources of former losers come into play, and given the assumption of the ability to pay, these resources are larger by definition.

The second reason is that a rule which is favorable to John but inefficient will shortly generate corrections, by increasing the price of transactions between John and the disenfranchised actors in other parts of the economic system. If, for instance, I should be so lucky as to obtain a rule that permits me to kick my tenant out sooner than in common practice, the price my tenants will be willing to pay should reflect this risk and be lower than the average market price. If, on the contrary, I am a tenant and I obtain a favorable rule, according to which the owner is required to allow me more time than it is commonly the case in the free market, soon the prices at which owners would be offering me accommodation will turn higher than the average market price. In a tolerably free market, the boomerang effect gets to discourage efforts to secure favorable but inefficient arrangements.

IV. ON THE UTOPIAN CHARACTER OF THE MODEL

4.1. Tyler Cowen's objection from instability

In 1992 Tyler Cowen (2007) formulated an invisible hand objection against the stability of the model. His conclusion is that incentives in the legal sector push (back) towards cartelization.

According to David Friedman all protection agencies get to have contractual agreements with other agencies that establish the terms in which potential conflict of interest between their clients are to be adjudicated. From the customer's perspective, the more worked out the pre-arrangement, the better.

But once an infrastructure for transactions is established between all operating firms, asks Cowen, what could prevent the members of such a network to constitute into a cartel with the intention of dividing the market into exclusive territories, raise prices and even institute taxation? Each firm belonging to the network would agree not to deal with upstart firms, or with firms that violated the common agreement to monopolize. Could they do it? Cowen does think so because of the following constructive dilemma.

First, brute force allows them: if members of a network can act collectively to such an extent that they can use force to put down outlaw agencies that do not accept their higher-order arbitration, then they are also able to put down agencies that do not adhere to a collusive agreement and thus enforce their cartel. If we assume a functional system of privatized justice, then we assume that competitive market forces prevent an outlaw firm from increasing its business by promising not to turn over its guilty customers for imprisonment or trial. But then, the infrastructure for collective action is already provided for, regardless of the number of agencies operating at one point. If the members of a network have sufficient resources to collectively force a truly outlaw agency (illegitimate by more largely shared standards of legitimacy) to turn over their outlaw clients to some adjudication then why don't they have sufficient resources to impose the terms of a cartel to outside agencies? Won't the majority of the network members also have the power to close that network, enforce the cartel terms on renegade agencies - rebel agencies that initially used to be a part of the network? The network can enforce compliance, as the network could threaten to cut off all relations with the renegades. Nonconforming firms will lose market share and go out of business because they cannot promise peaceful adjudication to customers.

Second, economic forces allow them too: the special network-nature of their industry makes it hard for new entrants and renegade colluders to obtain market share, if the cartel refuses them access to its arbitration network. Think of mobile telephone companies and the fact that intra-network conversations are cheaper. The more users a network has, the more useful it becomes. The larger network N1 and the smaller N2, the more costly it is for potential customers of N2.

Cowen's conclusion is that the network develops the attributes of a state: finance through taxation, claim of sovereignty, ultimate decision-making, authority, prohibitions on competitive entry (2007, 268).

Friedman's answer comes in two parts. First, members of the cartel may compete on the kind of service provided for intra-network clients, so there is some small room for competition. Or, at the very least, special insurance packages can be enforced by the cartel over different pairs of customers, which means we don't have uniform law codes across the cartel.

Second, as in any industry, members of a cartel can secretly sell more at a lower price (chisel on the cartel price) and thus serve unsatisfied left-out demand (Friedman 1994). Sometimes they could find it in their interest to settle some of the issues in ways prescribed by the non-member agency. This can happen when the non-member agency is willing to secretly pay for the change in a way in which both litigants are satisfied with the result.

We may add a third defense. If more or less incognito, out-of-network firms exist at all, they can also provide legal services between out-of-network members. The out-of network firms can establish terms of agreement with one another and thus start to form themselves a new network, perhaps incurring lower costs of production. The larger this second network the more incentives for customers of agencies who are part of the original cartel to give up on the all-inclusive package and to start buying parts of their protection services from agencies who are not members of the original cartel. Some people choose to have two or even three mobile phones and thus be members of different networks.

The first defense does not seem strong; after all, it does allow for quality collusion. The second defense seems stronger but it only secures a rather unofficial competiton, where firms operate more or less incognito. Competition depends on the degree to which firms feel pressured to conceal their dealings. The third argument also seems to revolve around the optimal degree of transparency of unofficial "renegade" dealings. Can we determine the optimal degree using economic analysis of law?

4.2. The actual foundations of the model rest with the distribution of power

What is the degree of transparency that agencies unabiding to the terms of the cartel can afford? This level depends on the relationships with peers and on the relationships with potential clients. The relationship with peers is characterized by such things as the risk of fewer transactions with peers and the risk of confronting their withdrawal of support in various collective actions. The relationship with potential clients depends on the transaction costs that unabiding agencies confront in reaching and concluding dealings with potential customers. Direct routes between competitors and clients may be hindered by other powerful agencies who seek to maintain monopoly. Potential customers' ability to pay may be low. If with anything, the utopian character of the model rests with these aspects. If we inquire into the utopian aspect of the model, we must inquire into possible

changes in these two factors continuously affecting the definition of the initial distributional points between agencies, and between agencies and their clients.

Let us pause the analysis of the factors that determine the fundamental distributional level for a methodological observation about our inquiry. As we can see, economic analysis is useful in determining that optimal degree of concealment. However, we were forced to go deeper than Coasian analysis can take us, since Coasian analysis only applies to transactions, and transactions always start off from some given initial distributional point. It is this precise point that is now in question. Indeed, the economic approach we entered goes in the direction of an institutionalism belonging to a family of authors like Mancur Olson (1993) and North, Wallis and Weingast (2009) who focus on the role of organizations more than on the role of transactions.

Coming back to the main line of inquiry, we said that the formal distributional point at the second level assumes the existence of some distributional point at the third (and fundamental) level. Could we also infer something substantial about the third level from data describing the second level? One conjecture is that if we observe a functional set of more than one independently operating protection agencies on a free market for justice, we can correlate the existence of that set with a balanced distribution of power between independent actors at the third and fundamental level. The argument for that correlation is not difficult. We have seen in the section on zero transaction costs that if the distributional level is mutually recognized, rectifying trasactions occur for an efficient result. In real life however agencies can actively obscure property rights and pretend that the distributional point was in their favor. They can do that because real options include violence besides just paying when prompted and the costs of violence for the two respective agencies may be different. If when calculating the relative costs of conflict, one agency repeatedly realizes that it is much weaker than the other, the other simply prompts and collects payment in all cases. After enough such repeated interactions weak agencies would disappear. The argument also works if formulated form the perspective of customers. Assuming the existence of a single center of power, freedom of enterprise in law offers escape routes to pairs of customers. But as the flow of rents towards that center weakens, so is the power necessary to sustain it.

The idea of correlating a certain freedom of legal choice with polycentricity of power finds an illustration and an empirical support in the first centuries of the Anglo-Saxon system of common law.

According to a strand of research on the development of common law, local courts of justice, manorial courts, mercantile courts, universities' courts and above all ecclesiastical courts, were independent from the king's courts. Officially, from the 12th century on, each of them was limited to a domain. However, for at least some legal areas such as contract law, these providers of justice were able to circumvent their limitations to effectively compete for customers (Berman 1983; Stringham and Zywicki 2010).

Even under an official cartelization between the Church and the Emperor (think of the Concordat of Worms in 1122 which established that the jurisdiction of the ecclesiasti-

cal courts should be limited to matters concerning the soul), competition between the ecclesiastical and the king's courts still stayed in place.

This was because in practice, it was the power to interpret the official limits that determined real jurisdictions; the power to interpret was continuously and silently disputed as a consequence of a larger dynamic. The justice providers (courts and systems of courts) could claim jurisdiction over domains with an unclear interpretation by making use of procedural fictions (Stringham şi Zywicki 2010, 11). Pairs of customers seeking legal insurance or adjudication of potential disputes from ecclesiastical courts could effectively choose that if they both swore an oath to God. A breach of contract could have been interpreted as an impiety against God. Impieties against God would fall under the jurisdiction of the Church. A set of such legal fictions came to constitute an unwritten basis of private justice. These customs endure until the 17th and 18th century, when the king's courts absorb all other courts and effectively take over the whole legal sector (Strigham and Zywicki 2010, 9, 11).

On the one hand, the availability of legal fictions is crucially important in determining whether the legal system was indeed privatized. The practical availability of legal fictions to common folk meant that customers were practically free to choose their legal provider. To the extent to which choice belonged to customers, justice was privatized.

On the other hand, the un-written, unofficial, or 'fictional' character of legal fictions is important in determining how transparent legal competition was. The number and complexity of functional legal fictions capable to help common people evade formal jurisdictional limitations, depended on the two factors mentioned above: the power support enjoyed by each of the legal courts or networks, and the transaction costs with customers.

North, Wallis and Weingast's (2009) and Wallis and North's (2010) idea that individuals use organizations as tools (sometimes organizations that they themselves did not create) in their confrontation with other powerful organizations offers a conceptual framework of interpretating connections between formal distributions of rights and fundamental distributions of power. It throws light on the reason why the power distribution (the third level of rules) is important for the practicability of privatized justice (i.e. the freedom to provide second-level distributional points plus the freedom to choose among providers). Courts in the ecclesiastical network stand in a client-patron relationship with the ecclesiastical political organization that forms and supports them. They can use it as a tool in their confrontation with powerful organizations such as the secular state. If the distribution of power between the two elite organizations, the Church and the state, is balanced, then pairs of customers are able to choose an ecclesiastical court as a legal provider. This is because, at the other end of the transaction, there is a powerful actor. The ecclesiastical political organization, which, as a matter of fact is able to choose them as customers, backs up by force its ability to deal with them.

At this point we can make another methodological observation about the usefulness of the Olson – North, Wallis, Weingast institutional economics for our inquiry in the stability of the Friedman model. In the chapter about stability in the *Machinery of Freedom*,

judging from the efficient size and number of present-day police forces, Friedman takes the guess that the number of protection agencies in the United States would be closer to 10,000 than to 3. He conjectures that if the legal system was privatized, a protection agency protecting as many as one million people would be far above the optimum size. The guess implicitly assumes that the transaction costs confronted by independent operating firms and their customers would be the same as the transaction costs confronted by present day police stations and citizens spread throughout the US[5].

Police stations today don't have the possibility to hinder other police stations' transactions with new clients for the purpose of extracting rents, since they do not compete. They are all units dependent on a central power elite. Independent economic agents would have, on the other hand, the incentive to hinder competitors' transactions with customers. This is clearly happening in all domains of activity in all history. Appealing to power resources (such as lobbying the state legislature) with the purpose of extracting rents is ubiquitous. Even if not transparent, rent seeking activities are effective, and this always shows on the number of independently operating firms. In the historical example we gave there were indeed many courts across England but they were not independent. They were caught in small number of operating networks. The number of those independent networks was actually closer to 3 than to 10,000.

At this point one might think the mechanics of power is reduced to a clear number of patron-client vertical networks: an ecclesiastical power elite with a set of client-courts and a secular power elite with its set of client courts. This would mean that the number of operating networks is 2. But if we look closely, that number was closer to 3 than to 2. There were also other legal networks. One of them was the network of merchant courts operating Law Merchant. Common people's access to merchant courts was granted as usual by a set of customary legal fictions. For example parties entering a contract who wished their future matters to be adjudicated in merchant courts, would, instead of giving an oath to God, "wet the bargain" by having a drink together at the moment of contract (Stringham and Zywicki 2010, 2, 14). The existence of a third independent network is curious because we cannot readily find the supporting patron.

Merchant courts are examples of organizations that emerge under the shelter of competition between the two power elite organizations, the Church and the State. The independence of the merchant legal network is explained by the closely balanced character of the distribution of power thereof. The close competition for power between the two main dominant coalitions forces each of them to bid for the support of third parties. Less powerful agents who formed Merchant Courts grew their freedom to form them as a consequence of a continuous bidding process for their support.

5] The size of the firm, or the existence of organizations, is explained by appeal to how costly market transactions would be by comparison to intra-firm relationships. This may be misleading for activities for which there are no preexisting market transactions at all and such no price tags to compare (Solcan 2002). This is the case with the provision of law: power elites have by and large prevented the existence of market transactions.

To conclude the section, if the historical data provided by Berman and Stringham and Zywicki is correct, then none other than Western history offers a lesson on the non-utopian character of private justice. Using the frame inspired by North, Wallis and Weingast, we can impute commoners's freedom of choice for legal distributional points to a particular power distribution at the fundamental distributional level. A fundamental distributional point characterized by a consistent balance of power between independent power elite organizations operating in the same territory stands at the origins of Western history. The difference between the actual history of legal competition in the first centuries of Anglo-Saxon common law and the David Friedman model of privatized justice is one of degree, not of nature. The difference rests in how limited, how unofficial and how opaque was freedom of legal choice, but nevertheless, the choice was there.

Although the Anglo-Saxon precedent based system of law is commonly defined as law applied by the king's courts (Hogue 1986), the content of the doctrines was, and to a certain extent still is, the result of a competitional dynamics (Strigham şi Zywicki 2010, 23). On the one hand, there was an incentive for the production of legal fictions in order to attract cases from other competitor-courts, and on the other, each of these alternative legal systems had the incentive to adopt the other system's successful legal innovations. The evolution of these concepts reflects the evolution of the legal system as a whole. Richard Posner (1992) discovers that the legal output of the Anglo-Saxon precedent based system is more efficient than the output of legislatures. If there is indeed a tendency of privatized justice to produce efficient rules, then one conjecture is that the present Western status quo contains fossilised efficiencies to the extent that it is itself the result of a private market for justice.

ionut.sterpan@cadi.ro

REFERENCES

Bakunin, Mikhail. 2011. *Stateless Socialism: Anarchism*. Anarchy Archives. http://dwardmac.pitzer.edu/anarchist_archives/bakunin/stateless.html (accessed June 20, 2011).

Barnett, Randy. 1998. *The Structure of Liberty*. Oxford: Oxford University Press.

Benson, Bruce. 1990. *The Enterprise of Law. Justice without the State.* San Francisco, California: Pacific Research Institute for Public Policy.

Berman, Harold. 1983. *Law and Revolution. The Formation of the Western Legal Tradition.*
Cambridge, Massachusetts: Harvard University Press.

Boettke, Peter. 2005. Anarchism as a Progressive Research Program in Political Economy. In *Anarchy, State and Public Choice*, edited by Edward Stringham. Northampton: Edward Elgar Publishing Co.

Buchanan, James. 2001. *The Limits of Liberty. Between Anarchy and Leviathan*. Liberty Fund. http://www.econlib.org/library/Buchanan/buchCv7.html (accessed June 20, 2011).

Buchanan, James, and Tullock, Gordon. 1995. *Calculul Consensului.* Translated by Paul Fudulu. Bucuresti: Expert.

Coase, Ronald. 1991. The Problem of Social Cost. In *The Market Economy. A Reader,* edited by James Doti, and Dwight Lee. Los Angeles: Roxbury Publishing Co.

Cowen, Tyler. 1994. *Rejoinder to David Friedman's Law as a Private Good.* http://www.gmu.edu/centers/publicchoice/faculty%20pages/Tyler/TYLER.pdf (accessed May 20, 2011).

———. 2007. Law as a Public Good: The Economics of Anarchy. In *Anarchy and the Law: The Political* Economy of Choice, edited by Edward Stringham. New Brunswick: The Independent Institute.

Dicey, Albert Venn. 2005. *Introduction to the Study of the Law of the Constitution.*
http://www.constitution.org/cmt/avd/law_con.htm (accessed May 20, 2011).

Doherty, Brian. 2007. Radicals *for Capitalism.* New York: Public Affairs.

Ellickson, Robert C. 1991. *Order Without Law: How Neighbors Settle Disputes.* Cambridge, Mass: Harvard University Press.

Friedman, David. 1973. *The Machinery of Freedom: Guide to a Radical Capitalism.* La Salle: Open Court Publishing Company.

———. 1986. *Price Theory.* http://www.daviddfriedman.com/Academic/Price_Theory (accessed May 20, 2011).

———. 1992. *Less Law than Meets the Eye.* http://www.daviddfriedman.com/Academic/Less_Law/Less_Law.html (accessed May 20, 2011).

———. 1994. *Law as a Private Good. A Response to Tyler Cowen on the Economics of Anarchy.* http://www.daviddfriedman.com/Academic/Law_as_a_private_good/Law_as_a_private_good.html (accessed May 20, 2011).

———. 1996. *Anarchy and Efficient Law.* http://www.daviddfriedman.com/Academic/Anarchy_and_Eff_Law/Anarchy_and_Eff_Law.html(accessed May 20, 2011).

———. 2001. *Law's Order: What Economics Has to Do with Law and Why It Matters.*Princeton: Princeton University Press.

Fuller, Lon. 1963. *The Morality of Law.* New Haven and London: Yale University Press.

Godwin, William. 2011. *Enquiry Concerning Political Justice and Its Influence on Morals and Happiness.* Anarchy Archives. http://dwardmac.pitzer.edu/anarchist_archives/godwin/PJfrontpiece.html (accessed May 20, 2011).

Hayek, Friedrich. 1945. The Use of Knowledge in Society. In *The American Economic Review* 35 (4): 519-30.

———. 1973. *Law, Legislation and Liberty, Vol. 1 Rules and Order.* Chicago: University of Chicago Press.

———. 1988. *The Fatal Conceit.* London: Routledge. 1992.

Hogue, Arthur. 1986. *Origins of the Common Law.* Indianapolis: Liberty Fund, Inc.

Molinari, Gustave de. 2011. *De la production de la securite.* Wikimedia.
http://fr.wikisource.org/wiki/De_la_production_de_la_s%C3%A9curit%C3%A9 (accesed April 27, 2011).

North, Douglass, John J. Wallis, and Barry Weingast. 2009. *Violence and Social Order.* Cambridge: Cambridge University Press.

Nozick, Robert. 1992. *Anarchy, State and Utopia.* Oxford and Cambridge: Blackwell Publishers.

Olson, Mancur. 1993. Dictatorship, democracy and development. *The American Political Science Review.* September, 87 (3): 567-76.

Posner, Richard. 1981. *The Economics of Justice.* Cambridge, Massachussetts and London: Harvard University Press.

———. 1992. *Economic Analysis of Law.* Boston: Little, Brown & Co.

Rothbard, Murray. 1998. *The Ethics of Liberty.* New York: New York University Press. http://mises.org/rothbard/ethics.pdf (accessed June 20, 2011).

Simmons, John. 1997. *Moral Principles and Political Obligation*. Princeton: Princeton University Press.

Solcan, Mihail Radu. 2002. Institutiile si teoria formala a actiunii (Institutions and the formal theory of action). In *Institutii in tranzitie* (*Institutions in transition.*), edited by Adrian Miroiu. Bucuresti: Punct.

———. 2003. *Freedom, Minds and Institutions*. Bucharest: Bucharest University Press. http://www.ub-filosofie.ro/~solcan/LMI/lmi.pdf (accessed May 20, 2011).

Stringham, Edward Peter and Todd Zywicki. 2010. Rivalry and Superior Dispatch: An Analysis of Competing Courts in Medieval and Early Modern England. *Springer Science+Business Media*. http://www.law.gmu.edu/assets/files/publications/working_papers/1057RivalryandSuperiorDispatch.pdf (accessed May 20, 2011).

Stringham, Edward Peter and Todd Zywicki. 2011. Hayekian Anarchism. George Mason Law & Economics Research Paper No. 11-06. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1744364 (accessed May 20, 2011).

Tullock, Gordon. 1967. The Welfare Costs of Tariffs, Monopolies, and Theft. *Western Economic Journal* 5 (3): 224–32.

Wallis, John and Douglass North. 2010. Defining the State. Mercatus Center, George Mason University. http://mercatus.org/sites/default/files/publication/WP1026_Defining%20the%20State.pdf (accessed May 20, 2011).

Warren, Josiah. 2011. *Equitable Commerce. A New Development of Principles*. New York: Burt Franklin. http://tmh.floonet.net/pdf/jwarren.pdf (accessed May 20, 2011).

Wolff, Robert Paul. 1998. *In Defense of Anarchism*. Berkeley: University of California Press.

Public Reason 3 (1): 98-102

Rawlsian Compromises in Peacebuilding: A Rejoinder to Begby

Alejandro Agafonow
ESSCA School of Management, LUNAM Université

Abstract. Begby's response neither offers a clarification of what he meant by the sort of political institutions that he claimed are provided by the idea of human security, nor ventures a word of defense for his unsatisfactory account of political representation. In this rejoinder, I provide textual evidence that shows what Begby is missing when he asserts either that political liberalism applies only to well-ordered societies, or that an overlapping consensus cannot be applied to relatively stable forms of political cooperation. In addition, I advance further considerations in order to dispel any doubts about what is at stake in this debate; to my mind, Begby risks standing for "comprehensive liberalism," while I emphatically stand for "political liberalism."

Key words: human security, comprehensive liberalism, political liberalism, liberal peace, John Rawls.

My reply to Begby and Burgess (2009) was built upon an account of political representation as part of the remedy for conflict-torn societies. The account provided by Begby and Burgess (2009) is unsatisfactory, as it appears back-to-front, with political representation as a secondary goal preceded by another kind of political institutions that they do not clearly define. In my view, this lack of definition stems from their inductive approach to peacebuilding, neglecting the big picture that the deductive approach of Kantian constructivism offers (cf. Agafonow 2010, 79). Begby's response, unfortunately, neither offers a clarification of the sort of political institutions which they claim are provided by the idea of human security, nor ventures a word of defense for their unsatisfactory account of political representation. In his response, Begby (2010, 52) confines himself to affirming that "one of the defining aims of liberal peacebuilding is to assist in the creation of a political institutional framework capable of dealing equitably and peacefully with ethnic or religious tensions as well as other sources of conflict," a statement too general to satisfactorily deal with my point.

It is important to clarify that ours is not a debate between a liberal, Begby, and a communitarian, myself. Instead, this is a debate between someone who risks standing for "comprehensive liberalism," Begby, and someone who stands for "political liberalism," yself.[1] Therefore, I do not doubt that liberal peace and its rights-centered agenda can alto-

1] Following John Rawls, in Agafonow (2010) I used metaphysical or epistemological liberalism as opposed to political liberalism. Here, I have decided to follow Waldron (2004) in referring instead to comprehensive liberalism as opposed to political liberalism, which I think conveys a more straightforward meaning. Waldron (2004, 91) defined them as follows: "The political liberal insists that the articulation and defense of a given set of liberal commitments for a society should not depend on any particular theory of what gives value or meaning to a human life. A comprehensive liberal denies this. He maintains that it is impossible adequately to defend or elaborate liberal commitments except by invoking the deeper values

gether successfully integrate "the sorts of compromises of moral and political principle which might be required in order to construct stable political institutions in societies emerging from conflict" (Begby 2010, 52). What I do doubt is that the idea of human security, at least as it is set forth in Begby and Burgess (2009), might be able to deal with the potential conflict between individual and community claims (cf. Agafonow 2010, 78).

In order to provide a meaningful settlement of this conflict, the idea of human security needs a qualification provided by political liberalism. This qualification concerns "political cooperation," which can be conceived on a continuum between a full overlapping consensus or perfect political agreement at one extreme, and no consensus whatsoever or complete political disagreement at the other. Indeed, Rawls not only provided a full description of how a well-ordered society would look if it complied with the ideal conditions of Kantian constructivism, but he also touched on the intermediate stages that a society would go through before fully embracing political liberalism. As shown below, Rawls considered "relatively stable forms of political cooperation" to be initially characterized by contending parties that temporarily and reluctantly agree on a political constitution, i.e. a modus vivendi, pending a final settlement that will eventually move the society from "simple pluralism" toward a "reasonable pluralism":

> [A]t the first stage of constitutional consensus the liberal principles of justice, initially accepted reluctantly as a modus vivendi and adopted into a constitution, tend to shift citizens' comprehensive doctrines so that they at least accept the principles of a liberal constitution. These principles guarantee certain basic political rights and liberties and establish democratic procedures for moderating the political rivalry, and for determining issues of social policy. To this extent citizens' comprehensive views are reasonable if they were not so before: simple pluralism moves toward reasonable pluralism and constitutional consensus is achieved. (Rawls 1996, 163-64)

Furthermore, Rawls referred to the opposite extreme, i.e. no consensus or complete political disagreement, in the case of belligerent and unreasonable comprehensive doctrines that prevent a political constitution with a full overlapping consensus. For such actors with deeply conflicting interests, it is impossible to meet the requisites of a mature liberal democracy:

> [I]f the liberal conceptions correctly framed from fundamental ideas of a democratic public culture are supported by and encourage deeply conflicting political and economic interests, and if there be no way of designing a constitutional regime so as to overcome that, a full overlapping consensus cannot, it seems, be achieved. (Rawls 1996, 168)

These passages illuminate what Begby (2010, 54) is missing when he asserts either that political liberalism applies only to well-ordered societies, or that an overlapping consensus cannot be applied to relatively stable forms of political cooperation—provided, of course, that we take care to use an appropriate qualifier with "overlapping consensus." In this respect, Begby overlooks the fact that I used "incomplete" as a qualifier here. Begby's

---

and commitments associated with some overall or 'comprehensive' philosophy."

oversight might have to do with his own conception of value and the good, perhaps one that stems from a specific political culture. This might explain why he finds counterintuitive the analytical outcomes of the method devised by Rawls and others (i.e. the simulation of negotiations in an original position, behind a veil of ignorance, to agree on a society's basic structure), which allows the inputs of other idiosyncrasies. There is no reason, for instance, to exclude a priori a feminist insight. I myself find Rawls too constrained by a specific political mindset, which prevented him from broadening the idea of overlapping consensus in a way I claim is coherent with the principles of Kantian constructivism. In this respect, his *Law of Peoples* is the work I find most problematic, as it sets forth a series of new propositions that do not clearly comply with the constraints derived from an original position and a veil of ignorance. Scholars like Abdel-Nour (1999) and Pogge (1994) have pointed out several related problems.

Furthermore, although Rawls' work contains the seed of a more applied perspective of Kantian constructivism, which nonetheless remains to be developed, it is doubtful that there is any actual well-ordered society deserving the status of full overlapping consensus. I gave several examples in my reply, stressing that not even some of the most enduring polyarchies neatly match this ideal (cf. Agafonow 2010, 80-1). If we have to avoid applying it to relatively stable forms of political cooperation, what entitles us to proclaim, for instance, Norway as a full overlapping consensus? As Begby (2010, 57) suggests, one would expect oppressive practices against women to be prevented in a full-fledged overlapping consensus. As it happens, Norway has one of the OECD's highest rates of physical violence against women allegedly committed by their own current or former partners.[2] Among the 28 countries included in the data, Norway has the seventh highest rate, exceeded only by Iceland, the UK and Switzerland among European countries.[3] However, this does not change the fact that in other instances, Norway largely resembles what we normally think of as a full overlapping consensus. A resemblance, however, is not a perfect match.

The picture that emerges from applying the method of Kantian constructivism is static; we need a theory that can also handle the dynamics behind political consensus building. In fact, this theory already partly exists in the literature dealing with democratization, particularly the institutionalist view of democracy, which shows how the causal order that explains democracy as a consequence of cultural factors appears to be mistakenly back-to-front. In other words, the emergence and consolidation of democracy can be better explained if political culture is understood as a consequence of certain constraints

2] As a general rule, reported domestic violence accounts for less than 2% of the population, with 1.16% in Norway. However, it is widely known that the extent of domestic violence is underestimated due to a greater reluctance to report it. In 2005, a nationwide mail survey carried out by the Norwegian Institute for Urban and Regional Research, with a response rate of 59.4%, showed that 27.1% of women in Norway experienced violence or the threat of violence in current or former relationships (see Office of the United Nations High Commissioner for Human Rights).

3] See the OECD Family Database, under the "Structure of the Family" heading, indicator SF3.4, "Family violence."

provided by institutional arrangements. If these arrangements are provided for—and here political representation is very important—we can expect internal liberalization to become self-propelled.[4] I find it surprising, therefore, when Begby (2010, 56) affirms that "[there] is no empirical support for the thesis that liberalization will, as a matter of fact, occur as the result of such compromises, nor that such compromises constitute the best or most reliable method of encouraging liberalization." If we cannot expect liberalization to emerge from compromises internal to the communities themselves, the Western imposition of foreign-led post-war reconstruction in the not-too-distant past begins to take on a certain inevitability.

Finally, we have to bear in mind that Begby's response addresses a reply of mine which, by definition, has a limited scope. Although some of his counter-arguments go beyond the intended scope of my reply, that which concerns the lack of a predictable time frame for the emergence of liberalization from within the local political culture itself particularly deserves comment. Begby (2010, 57) demands a "sharply defined time frame," at the end of which the institutional system is expected to have triggered initiatives of self-determination to tackle oppressive practices left to the internal jurisdiction of groups. But to expect that appropriate political representation shows results according to a sort of deadline is to overestimate, in a positivist fashion, the possibilities of social sciences.[5] It also confronts us with the moral problems of intervening in a sovereign country that has a stable but incomplete overlapping consensus.[6] And, one might think, recognizing sovereign non-liberal states is a far less pressing concern than preventing non-reasonable comprehensive doctrines from oppressing the individual.

alejandro.agafonow@essca.fr

REFERENCES

Agafonow, Alejandro. 2010. Human Security and Liberal Peace. Some Rawlsian Considerations. *Public Reason* 2(1): 77-84.

———. 2011. Deliberative Safeguards and Global Governance: A Market Based

4] Some representative scholars of this view are Juan J. Linz, Arend Lijphart, Fred W. Riggs, Alfred Stepan, and Arturo Valenzuela.

5] This, coupled with Begby's dismissal of what he terms "speculative psychological claims," ignores the sound reputation and extensive use of counterfactual theories in philosophy, not to mention political science or economics. Begby's response seems to rely on positivism's discredited sole reliance on external sensory observations.

6] By a stable but incomplete overlapping consensus, I don't mean a dictatorship, no matter how benevolent the ruler might be. Instead, I mean a country with a competitive political system, where politicians cannot push through a political agenda which they otherwise could in the absence of other politicians trying to match their own agendas with citizens' preferences, within the limits of the specified public matters that all contending groups are concerned with as a whole. Here, the qualifier "incomplete" means that some controversial issues would be left to the sole internal jurisdiction of specific contending groups, pending a final settlement in an open-ended future." For a similar point see Agafonow (2011).

Approach to Address Garrett W. Brown's "Deliberative Deficit" within the Global Fund. *Theoria. A Journal of Social and Political Theory* 58(128): 40-54.

Abdel-Nour, Farid. 1999. From Arm's Length to Intrusion: Rawls' "Law of Peoples" and the Challenge of Stability. *The Journal of Politics* 61(2): 313-30.

Begby, Endre. 2010. Rawlsian Compromises in Peacebuilding? Response to Agafonow. *Public Reason* 2(2): 51-59.

Begby, Endre and and J. Peter Burguess. 2009. Human Security and Liberal Peace. *Public Reason* 1(1): 91-104.

OECD Family Database, http://www.oecd.org/document/4/0,3746,en_21571361_38039199_37836996_1_1_1_1,00.html (accessed June 5, 2011)

Office of the United Nations High Commissioner for Human Rights, Special Rapporteur of the Permanent Mission of Norway in Geneva, http://www2.ohchr.org/english/issues/women/rapporteur/Norway.pdf (accessed June 5, 2011)

Pogge, Thomas W. 1994. An Egalitarian Law of People. *Philosophy and Public Affairs* 23(3): 195-224.

Rawls, John. 1996. *Political Liberalism*. New York: Columbia University Press.

Waldron, Jeremy. 2004. Liberalism, Political and Comprehensive. In *The Handbook of Political Theory*, eds. Gerald F. Gaus and Chadran Kukathas, 89-99, London: Sage.

Public Reason 3 (1): 103-113

MacIntyre on Personal Identity

Lia Mela
University of Patras

Abstract. MacIntyre's interpretation of human life is influenced by the medieval conception, which considers the human to be substantially in via. What allows MacIntyre to maintain that there is a specific "narrative unity of human life" is the conception of life as "quest" or "journey". Thus, the self is tied up to a character and his/her unity is given as the unity of the character which is demanded by the narration. The construction of personal identity revolves around the question "what is a good life?", adopting a structured narration as a basis for the formation of personal identity. Nevertheless, although the unity of life demands an end, the recognition of the fact of pluralism averts MacIntyre from any effort to define positively what is good. However, given the fact that the possibility of understanding life is intertwined with a specific time, place and civilization, the quest does not blindly move in a vacuum, but within traditions, since we are bearers of a specific story. I will argue that (a) as long as the notion of the end remains bereft of content, its capability of solving conflicts of values becomes weaker, (b) in this context, the problem of the coexistence of individuality and social determination is raised.

Key words: personal identity, practice, narrative, moral tradition, modernity, individualism, autonomy, community, teleology.

MacIntyre's theory is part of a wider discussion about the content, the limits and the perspectives of Modernity, which acquires the character of a conflict that revolves around conceptual pairs such as substantial/procedural ethics, ideal of good life/rules of fair coexistence, teleology/deontology, communitarianism/liberalism. Vis-à-vis the question that concerns the possibility of inducing a universal core out of the pluralism of valid forms of life, the principal tendency of modern ethics argues that the claim of universality requires detachment of the contents associated with the ethos of specific communities. The questions about life are posed as a matter of responsibility of the socialized individuals, judged through the perspective of each one of them.

MacIntyre's updated account of the virtues works as a critique of the abstract character of modern ethics in favour of an orientation towards the communal ethos. First it is a critique of individualism, which conceives the individual, that is a physically discernible and psychologically continuous rational and autonomous subject, as the basic ethical unit, ultimate source of value and bearer of justification. For MacIntyre, the human individual has an inherent complexity which the individualistic point of view is unable to conceive. The modern self cannot be held responsible in the way that pertains to a moral subject, as s/he faces alternative ways of life from an external point of view, provided that ex ypothesi it does not have any commitments. Nevertheless, its conflicting desires are unable to comprise a base of choice and, as a consequence, it cannot form a rational history in its movements from one state of moral commitment to another. The subject, if no definitive teleology is presupposed, i.e. a notion of what is good, definable by communal or species needs, lacks a guiding thread. The capability for practical reason is not abstract, is not understandable outside a social (*After Virtue*, *Whose Justice? Which Rationality?*) or biological

(*Dependent Rational Animals.Why Human Beings Need the Virtues*) framework. In the present article I am going to discuss the early historical-social approach of the philosopher.

According to MacIntyre, moral theory is today in a state of disorder, as it emerges as a mass of conceptual fragments that have survived from the past, though detached from the wider viewpoint with which they were connected and**,** hence, lacking the context from which they derived their meaning (MacIntyre 1985, 2, 110-11, 256, 257).[1] Modern culture suffers from the loss of a normative order, a loss that the philosophy of the time reproduces in theory and, as a consequence, the solution of moral problems on the level of social practices is rendered impossible**.** MacIntyre's diagnosis concerns Modernity as a whole, given the fact that the confusion does not concern just theoretical, philosophical approaches, but life as a whole, individual and collective, firstly around matters of ethics and secondly around matters that concern truth and rationality. Post-enlightenment culture is deflected towards emotivism on the theoretical level and towards bureaucracy on the level of social-political practice.[2] Under that second aspect, modern society constitutes an exemplary form of alienation.[3]

MacIntyre argues that in terms of modern and contemporary ethics, moral judgments cannot be characterized as being true or false, as it is accepted that these express simply subjective feelings or attitudes of the speaker, contrary to the judgments over facts, which have an objective meaning, as they are empirically verifiable. Since moral judgments cannot be justified by rational criteria, they are faced as objects of an arbitrary choice. Thus, the relation of the self to his/her ends lies in a free act of will, like transportation from one moral point of view to another. This is an approach of cosmopolitanism without roots, in the centre of which MacIntyre tracks down an impoverished conception of the self and the society, a homeless individual which is potentially everything and in reality nothing (MacIntyre 1988, 339). And he holds that in this capability of the self to avoid any necessary identification with any particular moral point of view or social role, modern philosophy detects mistakenly the essence of moral subjectivity (MacIntyre 1985, 31). MacIntyre suggests that the shadowyfigure of the person, as sketched by Modernity,

1] See MacIntyre's "disquieting suggestion", *op.cit.,* 1-5. The understanding of modern ethics in terms of fragments first appeared in E. Anscombe 1958, pp. 1-19. With regard to MacIntyre's view, see Begley 1995, 221-23, Garcia 2003, 94 ff., Haldane 2005, 92-93, Kelly 2005, 128.

2] The term emotivism in the work of MacIntyre functions as a wider hypothesis, so it does not coincide with the specific metaethical theory of the early 20th century. So, when MacIntyre puts under that flag different philosophical stances, from Hume to Nietzsche, Stevenson or Sartre and Weber, he is not referring to a specific theory but to the current incapability of rational justification in ethics, which is a necessary product of the acceptance of the basic thesis that evaluative propositions are matter of preference. For MacIntyre this is not the point of view of some specific philosophical school, because the liberal individualism, which is responsible for this, has affected even the philosophical programs that are supposed to be contrary to it, like Marxism.

3] The surface of emotivist society is about freedom. But on a deeper level there are functioning manipulative relations, since the individual is under the power of elites formed by the three central modern characters: the rich Aesthete, the Manager and the Therapist.

is deprived of crucial qualities which are traditionally connected with the idea of moral personality, receiving thereby an abstract and pale character and indeed in a way that puts into question the continuity and the identity of the subject. As a matter of fact, it is formed through the loss of traditional bonds that used to structure the social identity and the view of human life as an order towards a specific end (MacIntyre 1985, 33-34).

According to MacIntyre, the Enlightenment inherited a complex combination of ancient Greek elements, in conjunction with assumptions of medieval Christendom, but the teleological aspect, the teleological understanding of human life**,** was lost in the historical process. Nevertheless, the teleological worldview of the lost ancient and medieval world incorporated a unity of metaphysical, epistemological and ethical presuppositions. MacIntyre aims at the reactivation of the ethical content, without, however, retaining its overall epistemological and metaphysical presuppositions. He substitutes the notion of History for the notion of Nature**,** aiming at the construction of an objective teleology, which will not be loaded with a specific conception of human nature or essence and will not presuppose an agreement on a common conception of the good.

MacIntyre develops the meaning of "virtue" in three stages (MacIntyre 1985, 186-87): (1) "practice", (2) "narrative unity of human life", (3) "moral tradition". Centered upon a specific "practice" and the "internal goods" that determine it, the participants form stories that give meaning to their life.[4] These "narrations" weave themselves together, in order to form a concept of a human end, which is the driving force of a "tradition", that is a wider cultural narration which functions as a basis of justification on the individual and the social level, so that the separate actions are judged according to their cohesion to or their deviation from this wider complex of values.

MacIntyre's interpretation of human life is influenced by the medieval conception, which considers the human to be substantially in via. What allows MacIntyre to maintain that there is a specific "narrative unity of human life" is the conception of life as "quest" or "journey". The unity of life does not lie upon the chronological arrangement of facts but is rather characterized in terms of a literary work, it has a dramatic form. The unity of the self is based upon the unity of a narration which connects birth to life and to death, as in the context of a dramatic play the beginning is connected to the middle and to the ending (MacIntyre 1985, 205; 1990, 198). Human beings have a responsibility for the history of which they are the authors, whereas human expression renders itself comprehensible only when it finds its position within a specific narrative (MacIntyre 1985, 209). Thus, it should not be assumed that the person occupies a specific static or fixed position, but just that s/

4] "Internal goods" of a practice are the goods that cannot be identified and attained in any other way than by the experience of participating in the practice in question. On the contrary, "external goods" are those goods externally and contingently attached to a practice, that is there are always alternative ways for achieving them. For example, strategic capacity is an internal good to the practice of chess playing. Fame or wealth on the other hand is external goods, because there are always alternative ways of achieving such goods and their achievement does never depend only on the engagement in that particular kind of practice (MacIntyre 1985, 188 ff.).

he is situated in a specific point of the journey towards the attainment of her/his end. The narrative unity of life is teleological, since it presupposes the ability to evaluate in terms of success or failure the search for a conception of the good, the history of which surpasses the individual (MacIntyre 1985, 34).

The human in practice, as in his/her fictions, is in fact a "story-telling animal". I can answer the question "what to do?" only if I have previously answered the question "which history or histories I am part of?." According to MacIntyre, listening to stories about bad step-mothers, lost children, deceived kings, children learn the meaning of relevant roles, they recognize the characters of their personal drama and they understand the world. Through the myths and the fairy-tales of childhood, the child finds a way of handling a chaotic world. The weight of the myth is later taken over by philosophy. According to MacIntyre, there is no way of comprehending any society, including ours, if we do not try to do it through its initial dramatic sources (MacIntyre 1985, 216; 2006, 7-8).

The subject functions as a writer and an actor at the same time. What we can say and do is deeply affected by the fact that we are never something more than co-writers of our narrations. Only in our imagination do we live the story which we like. We do not begin ab initio, we are inserted in a stage of the story which we have not planned, and we become part of an action which we did not initiate. We are restricted by the actions of others and by the social environment (MacIntyre 1985, 215). Each one of us is the basic character of his/her own drama and plays secondary parts in the dramas of others, while each drama sets limits to others. Thus we are not autonomous writers, since we are protagonists of our own story, but at the same time, part of the stories of others. We are characters of a parallel series of intermerging narrations, some of which are often integrated in others, as for example, it can happen with Mary Stewart's life in that of Elizabeth's I. In that way the narrative conception of the self is formed. The subject is what s/he is justifiably considered by others to be for the duration of the story that he/she unfolds from his/her birth to his/her death (MacIntyre 1985, 213, 215, 217).

The conception of the narrative unity of life involves, first, an ontological dimension, in the context of which priority belongs to the society, since the character is defined by the social construction of individuals, their self-understanding within a specific social structure. Second, a normative dimension, in the sense that it is constitutive not only of the experiences, but of the person as well, of his/her identity and continuity as a bearer of these experiences. And third, an historical-social dimension, which implies the continuity of life in time and is integrated in a wider story, in the sense that it is constitutive of the social structure and the groups that are formed in its frame, elements which also have an inherently narrative character.

However, what does the unity of individual life consist in? MacIntyre does not try to base personal identity upon the psychological continuity of the self**,** nor its lacking upon his/her discontinuity. Personal identity is the identity that is demanded by the unity of the narration's character, since, without unity of this form, there can be no subjects with narrationable stories. Thus, the self is tied up to a character and his/her unity is given as

the unity of the character which is demanded by the narration. According to MacIntyre, this is about the unity of a narration incarnated in a single life (MacIntyre 1985, 217-18). The construction of personal identity revolves around the major question "what is a good life?" adopting a structured narration as a basis for the formation of personal identity. In this context, MacIntyre's critique of emotivism is about the loss of an end, which provides life with an order and a goal, setting limits, social frame and standards**,** so that life can be evaluated. In post-enlightenment culture, moral judgements, unable to be graded according to a specific supreme end, simply express equally valid preferences.[5]

By the introduction of the concept of "narrative unity of life", MacIntyre seeks to solve the problem of choice among different "practices". As the "internal goods" of various practices are not ordered according to their value, the choice among them is made according to the conception of the end of human life as a narrative unity. Otherwise, we would face the problem of an arbitrary choice, without any criteria, which characterizes emotivism. The understanding of human life as a narrative unity renders its confrontation possible, not as an expression of arbitrary will, but as a field of self-expression that provides elements regarding its own content.

Thus MacIntyre blames modern culture because, by abandoning teleology, it is unable to indicate ends that aim towards the moral perfection of man, understood as self-realization. MacIntyre's work is an effort to approach the human subject in relation to the achievement of a specific end, which gives to human practices a perfectionist character, by forming the criterion according to which human life will be judged as a whole**.** The self emerges as a subject of a narration, which is bound together by a sovereign idea of the good, aspiring to authenticity, to the fulfillment of the truth that is connected with the achievement of the human end, and implies the moral perfection of the self, by means of exercising the proper virtues.

The notion of authenticity sets limits to the continuous re-definition of identity, which are dictated by the integration of the subject into a specific precedent moral frame, which is not of his/her own making. MacIntyre's critique is centered upon the thesis that

5] According to MacIntyre, Aristotelian teleology is based on a crucial distinction between "man as it is" (ἐνεργεία) and "man as it could be if his telos is realized" (δυνάμει). Aristotelian ethics presupposes (MacIntyre 1985, 53): (a) a conception of the human nature as it is in its untutored state, (b) a conception of the human nature as it would be when its "end" (τέλος) is realized, (c) a conception of practical reason and ethics as a means of the transition from the first to the second state. Ethics is the science which can understand and guide this transition, because it signals the road which leads to the realization of true human nature, e.g. the realisation of the human telos through the cultivation of virtues (op.cit., 52). Due to the abandonment of teleology during modern times, the concept of the good (ἀγαθόν) is deprived from any factual information and is relegated to a simple sentiment. But it is precisely the concept of the human telos that justifies the movement from "factual" to "evaluative" judgments, from "is" to "ought," that is the transition from empirical premises to evaluative judgements (op.cit., 57). With his theory of man as a "functional concept," MacIntyre aims at confronting the problem of the naturalistic fallacy. Functional concepts concern things that due to their construction or beings that due to their nature are destined to perform a specific function. Since their destination is the fulfilment of purposefulness, they will be judged in terms of this purpose which constitutes their destination.

the critical liberation which is connected with the ideal of autonomous choice leads to a loss of moral orientation. In this context he recognizes the capability of reason, but he holds that this presupposes a precedent complex of commitments for an evaluation of right and good to be possible. Thus, it cannot be regarded as a capability of detachment from the historically fixed practices of evaluation.

In this way a wider cultural background of common knowledge and historical memory is formed, from which the subject draws objective moral criteria and the elements that allow the formation of his/her personal story. The very stories that s/he forms, inevitably bound up with the frame of the social union to which the subject belongs, in a way that the historical and the social identity eventually coincide. So, when moral language collapses to the conditions of disorder linked to emotivism, man can neither find common meanings on an individual level, meanings that would allow him/her to get individualised, nor can the political community express the moral community of citizens in a certain common program on a collective level.

By the introduction of the notion of tradition, criteria of excellence are being defined, which are meant to function as markers in the narrative process of the subject, so that s/he can manage to handle the various particular goods of the practices that have been historically formulated**.** Thus, whereas on the one hand the understanding of human excellence is open to changes, on the other hand it is defined by the authorities of a specific historical time. That means that, first, this understanding does not concern the unfolding of a specific context of human nature, of some substance, and it does not stem from a timeless definition of this nature. Second, it is not a matter of an arbitrary choice, in the manner of emotivism. For this reason, the rejection of traditions constitutes a major problem for liberal individualism.

Individualism ignores the social integration of the human end and for this reason it refers to isolated choicemakers, incapable thus of giving a coherent ideal of the good life. From this point of view, I am what I choose to be. I can always, if I want to, put into question these elements that are regarded as merely random social characteristics of my existence (MacIntyre 1985, 220). On the contrary, for MacIntyre, the conception of the good has a social dimension, since it is inherently and indissolubly linked with the nature of a given historical society, which limits the way in which the individual conceptions about what is good are formed**.** The self, in the final analysis, consists in the social roles which s/he inherits.

Thus, the definition of the identity of the subject includes elements descriptive of the roles s/he plays: "But it is not just that different individuals live in different social circumstances; it is also that we all approach our own circumstances as bearers of a particular social identity. I am someone's son or daughter, someone else's cousin or uncle; I am a citizen of this or that city, a member of this or that guild or profession; I belong to this clan, that tribe, this nation. Hence, what is good for me has to be good for one who inhabits these roles. As such, I inherit from the past of my family, my city, my tribe, my nation, a variety of debts, inheritances, rightful expectations and obligations. These constitute the

given of my life, my moral starting point. This is in part what gives my life its own moral particularity." (MacIntyre 1985, 220)

What happens, though, when the same person is part of more than one conflicting practices and traditions? And on what basis can there be a quest for the social good in societies which are fragmented and conflicting? Modern ethics, by transferring the emphasis upon the rules, tries to confront the problem of human coexistence in the context of an "open", pluralistic society, when, that is, people cease to define themselves according to a common conception of the good and their society is not any more held together through the strong bonds of a community. The demand for autonomy liberates from the hierarchical structures of the past, no matter whether they stem from the secular or from the religious power. It substantially concerns the guarantee of a frame, within which everyone seeks his/her good, a good which should remain indefinable from a public point of view. Provided that he recognises it, MacIntyre should too confront the same problem.

MacIntyre has to show how the notion of the end, which differentiates itself from person to person, can be connected with a wider sense of the end for human beings, so that the constitution of a narrative unity can become possible and the life as a whole can have a direction. Nevertheless, although the unity of life demands an end, a predefined way of life that determines its conditions, the recognition of the fact of pluralism averts MacIntyre from any effort to define positively what is good. However, we notice that as long as the notion of the end remains bereft of content, its capability of solving conflicts of values becomes weaker, as it is rendered conventional in a frame of heterogeneous practices and traditions, and as a result the concept of the quest remains accordingly undefined and, hence without a guiding line, especially with respect to the probability of the function of conflicting and possibly evil practices and traditions.[6]

The systematic posing of the question "what is the good life?" receives different answers, which probably diverge from the true good or the common good. Thus, there is an indetermination of the quest, the exact nature of which remains undefined, since its end is not defined.[7] Under which conditions, though, is it possible to construct a teleology without defining the notion of an end? And how can we know that our end is true?

First, we have to recognize that MacIntyre locates a real problem when he delineates the question of a meaningful life in Modernity, suggesting that the ideal of autonomy has no meaning, if it concerns an empty notion of freedom. The commitment to autonomy, which entails that there can be no external imposition of a specific conception of the good on the individual, can function only as a necessary, but negative, condition of existence. The free individual remains without a guiding line in life. It is indeed necessary to assure both the cognitivist claim in the field of ethics and the possibility to articulate a discourse

6] See the paradigm of Nazism in MacIntyre 1985, 179-80. The problem of evil practices is posed by several critics, see, between others, S. Feldman (1986, 310ff.). Especially from a feminist point of view, see Frazer & Lacey 2005, 273ff. and *passim* with further bibliographical references.

7] About the need to determine the concept of the end, see also Frankena 1983, 585, Downing 1984, 43 ff., Kitchen 1997, 81 ff. Cf. Miller 1984, 55-56.

about the content of life. His approach is flawed as far as its positive side is concerned, as it hesitates to identify positively the conception of the good. Of course the requirement is not a final theory of the good. To demand that MacIntyre gives us a final conception of the good means that in some way we ask him the impossible. In any case an independent definition of the good would conflict with the principle of autonomy. But nevertheless we need criteria according to which rival conceptions of the good can be assessed. The condemnation of Modernity, towards upon which MacIntyre calls us, based on the argument that Modernity does not offering any specific conception of the good cannot be justified exclusively on negative arguments. We therefore needed a more positive identification of good, which allows us to judge the beliefs of the communities on the basis of which the positive part of MacIntyre's alternative in relation to the program of Modernity will eventually be judged and, with it, his critique of Modernity.

The definition of what is good in *After Virtue* is formal; good life is what we live while seeking the good life for man. Leaving aside the question about whether this definition constitutes a vicious circle, a question arises about how the definition acquires any content. In *After Virtue* and *Whose Justice? Which Rationality?*, this happens through the connection with tradition. The notion of tradition comes to cover the void that is created by the disappearance of the ancient Greek city-state and of the medieval community, aiming at the preservation of the teleological framework, as it is associated with the defense of a culture of belonging, bringing to the surface the historical-social framework of the subject's constitution. The subject is, in terms of values**,** integrated in a specific wider historical-social horizon, whereas the conditions of his/her existence are intertwined with the collectively consolidated background of this horizon. Since MacIntyre rejects not only the possibility of an empirical understanding of the world (Humean tradition) but also the perspective of an a priori understanding (Kantian tradition), he turns to the possibility of a dialogue between the subject and the wider value framework inside which s/he is integrated and on the basis of which s/he is self-defined.

Nevertheless, neither the concept of tradition is given a precise definition by MacIntyre. The concept rather emerges through examples, in a way that does not convey a unified conception.[8] First, there is the problem of delimiting the tradition on the basis of which we are self-determined, given the fact that we belong to more traditions or even to the same tradition, which, nevertheless, we interpret differently. Is the geographical, the racial, the national, or the religious element important? Which degree of cohesion is indispensable for the recognition of a tradition and how many deviations can it endure? Finally, which is the value and on what basis is our primary commitment to a certain tradition justified?

8] See also Annas 1989, 389, Porter 2003, 38-39. About the problem of the pluralism of traditions and their individualization, see also Feldman 1986, 314-15, Paden 1987, 135, 138 ff., Wallace 1989, 346-347, Haldane 1995, 95.

An answer to these questions is required, particularly since by invoking the powerful bond of a specific tradition a slippery path probably opens up towards authoritarianism.[9] Given the fact that the possibility of understanding life is intertwined with a specific time, place and civilization, the quest does not blindly move in a vacuum, but within traditions, since we are bearers of a specific story. In this context, the problem of the coexistence of individuality and social determination is raised. If the sense of individuality is not meant to stand in a contradictory relation to the sense of belonging, a provision should be made for the preservation of the bonds of the subject with the social world while also allowing space for a critical attitude towards commitments.

Man cannot be assumed to be a simple epitome of the historical-social condition, because the more the social determination is strengthened the more the margins for freedom of the will are limited. S/he can neither be considered as a completely free individuality, since that would amount to a slip towards emotivism. The more profound beliefs and commitments of the person are not dealt with as the object of choice, but as an expression of belonging to a given form of life. The value of belonging comprises a reorientation of thinking, under the aspect of which reason does not aim at the critical detachment from the historical and social conditions, but faces them as what determines the way we move inside the form of life to which we belong.

We notice that the problem appears to be double: (a) since the good does not have a predefined form, it will be defined through the quest, and this means that it can even be in conflict with the given forms of what is good, which we have inherited. From this we should conclude that the role of the community is indeed binding for the individual, albeit not in such a way that the individual is eventually absorbed by the community, because MacIntyre recognizes the possibility of setting oneself apart by adopting a critical stance (MacIntyre 1985, 221).[10] The very understanding of the end as a quest leaves room for the development of pluralism. Nevertheless, the rules of critique appear to be historically and communally integrated, which means that we inevitably move within traditions, even when we criticize them. Consequently the problem of authority, which is additionally attached to the existence of power structures within the tradition, does not arise, since MacIntyre does not give a dogmatic description of the human end or the fundamental characteristics of a specific tradition, but this is interwoven exactly with the relativisation of the criteria within their context.

9] The existence of authoritarian characteristics is posed by J. Porter (2003, 62-66), and G. Graham (1995, 169). That Macintyre forms a conservative form of traditionalism which is incompatible with (a) the Aristotelian-Thomistic tradition, see Coleman 2005, 80 ff., Nussbaum 2007, 1 ff. (b) the Hegelian tradition, see Jurist 1992, 167-69 ff. MacIntyre is judged as neoconservative on the basis of Habermasian arguments by Paden 1987, 125, 138, 141 and *passim*.

10] J. Horton & S. Mendus (2005, 13) are referring to the face of Janus in MacIntyre's philosophy: one the one hand there are elements of tolerance connected to the recognition of pluralism similar to those of liberal theories. On the other hand, he recognizes an authoritative element residing within tradition. This problem is also posed with reference to the concept of social roles in Macintyre's theory by L. Downing (1984, 43 ff., 51-52).

(b) Since practical reason is defined by tradition, without there being objectively recognized and valid common criteria for the weighing up of heterogeneous traditions that are considered to be incommensurable, how can we choose between them during the formation of the narrative unity of our life, and which commitments can put up limits to the actions of people who are not tied up to each other by common practices or traditions? The epistemological relativism of MacIntyre leads up to cultural relativism, if it is accepted that there is a multitude of traditions and it is up to us to attach ourselves to a specific one of them or to abandon it.[11]

It is clear from the above that MacIntyre strives to move between the Scylla of emotivism and the Charybdis of traditionalism. He poses a demand for reframing reason, so that the problems of the universal abstract reason of Modernity are remedied, by trying at the same time to assure the possibility of inside-the-framework criticism. We notice, however, that this vacillating position of his, struggling on the one hand to preserve conditions that allow for criticism in the framework of wider practices and on the other hand fighting to reassure their authoritative character, by emphasizing the value of co-identification, as a necessary condition for the constitution of a tradition, he creates a certain internal tension in his theory. If tradition retains elements of authority, then it bears an authoritarian element. If we are to reduce the authoritarian elements to ensure the possibility of criticism, how far are we from the Enlightenment Project?

liamela@yahoo.com

REFERENCES

Annas, Julia. 1989. MacIntyre on Traditions. *Philosophy and Public Affairs* 18 (4): 388-404.

Anscombe, Elizabeth. 1958. Modern Moral Philosophy. *Journal of Philosophy* 33 (124): 1-19.

Begley, John. 1995. The Virtue Theory of Alasdair MacIntyre. *Pacifica* 8: 220-28.

Coleman, Janet. 2005 [1994a]. MacIntyre and Aquinas. In *After MacIntyre: critical perspectives on the work of Alasdair MacIntyre*, edited by J. Horton, and S. Mendus, 65-90. Cambridge: Polity Press.

Downing, Lyle. 1984. After "Telos": The Implications of MacIntyre's Attempt to Restore the Concept in "After Virtue." *Social Theory and Practice* 10 (1): 39-54.

Feldman, Susan. 1986. Objectivity, Pluralism and Relativism: A Critique of MacIntyre's Theory of Virtue. *The Southern Journal of Philosophy* 24 (3): 307-19.

Frankena, William. 1967. The Naturalistic Fallacy. In *Theories of Ethics*, P. edited by P. Foot, 50-63. London: Oxford University Press,.

Frazer, Elizabeth, and Nicola Lacey. 2005 [1994a]. MacIntyre, Feminism and the Concept of Practice. In *After MacIntyre: critical perspectives on the work of Alasdair MacIntyre*, edited by J. Horton, and S. Mendus, 265-82. Cambridge: Polity Press.

11] In this sense, it is indeed argued that MacIntyre's thesis collapses to the emotivist thesis; see Schneewind 1982, 659-66, Wallace 1989, 336-337, Hinchman 1989, 651, and Meilander 2006, 6.

Garcia, J.L.A. 2003. Modern(ist) Moral Philosophy and MacIntyrean Critique. In *After MacIntyre: critical perspectives on the work of Alasdair MacIntyre*, edited by J. Horton, and S. Mendus, 265-82. Oxford: Polity Press.

Gordon, Graham. 2005 [1994a]. 'MacIntyre's Fusion of History and Philosophy'. In *After MacIntyre: critical perspectives on the work of Alasdair MacIntyre*, edited by J. Horton, and S. Mendus, 161-75. Cambridge: Polity Press.

Haldane, John. 2005 [1994a]. 'MacIntyre's Thomist Revival: What Next?'. In *After MacIntyre: critical perspectives on the work of Alasdair MacIntyre*, edited by J. Horton, and S. Mendus, 91-107. Cambridge: Polity Press.

Hinchman, Lewis. 1989. Virtue of Autonomy: Alasdair MacIntyre's Critique of Liberal Individualism. *Polity*, 21 (4): 635-54.

Horton, John, and Mendus Susan (eds.). 2005 [1994a]. *After MacIntyre. Critical Perspectives on the Work of Alasdair MacIntyre.* Cambridge: Polity Press.

———. 2005. Alasdair MacIntyre: After Virtue and After. In *After MacIntyre: critical perspectives on the work of Alasdair MacIntyre*, edited by J. Horton, and S. Mendus, 1-15. Cambridge: Polity Press.

Jurist, Elliot. 1992. Recognizing the Past. *History and Theory*, 31 (2): 163-81.

Kelly, Paul. 2005. MacIntyre's Critique of Utilitarianism. In *After MacIntyre: critical perspectives on the work of Alasdair MacIntyre*, edited by J. Horton, and S. Mendus, 127-45. Cambridge: Polity Press.

Kitchen, Gary. 1997. Alasdair MacIntyre. The Epitaph of Modernity. *Philosophy and Social Criticism* 23 (1): 71-98.

MacIntyre, Alasdair. 1985 [1981a]. *After Virtue. A Study in Moral Theory.* London: Duckworth.

———. 1988. *Whose Justice? Which Rationality?* London: Duckworth.

———. 1990. *Three Rival Versions of Moral Enquiry.* London: Duckworth.

———. 1999. *Dependent Rational Animals. Why Human Beings Need the Virtues.* London: Duckworth.

———. 2006. *The Tasks of Philosophy. Selected Essays, vol. 1.* Cambridge: Cambridge University Press.

Meilaender, Gilbert. 1999. Review Essay of Dependent Rational Animals: Why Human Beings Need the Virtues and The MacIntyre Reader. *First Things* 96: 47-55.

http://www.firstthings.com/ftissues/ft9910/articles/meilaender:html (accessed October 20, 2006).

Miller, David. 1984. Virtues and Practices. *Analyse und Kritik* 6 (1): 49-60.

Murphy, Mark (ed.). 2003. *Alasdair MacIntyre. Contemporary Philosophy in Focus.* New York: Cambridge University Press.

Nussbaum, Martha. 1989. Recoiling from Reason. *The New York Review of Books,* 36 (19): 36-41.

Paden, Roger. 1987. Post-Structuralism and Neo-Romanticism. Or Is MacIntyre a Young Conservative? *Philosophy and Social Criticism* 13 (2): 125-43.

Porter, Jean. 2003. Tradition in the Recent Work of Alasdair MacIntyre. In *Alasdair MacIntyre*, edited by M. Murphy, 38-69. Cambridge University Press.

Schneewind, Jerome. 1982. Virtue, Narrative, and Community: Alasdair MacIntyre and Morality. *The Journal of Philosophy* 79 (11): 653-63.

Wallace, Jay. 1989. Review of After Virtue: A Study in Moral Theory, Whose Justice? Which Rationality? *History and Theory* 28 (3): 326-48.

Public Reason 3 (1): 114-117

Book Reviews

Gillian Brock, Global Justice: A Cosmopolitan Account, Oxford: Oxford University Press, 2009, Pp. 288, ISBN 0199230943.

A cosmopolitan account of global justice should naturally be at the forefront of any theory concerned with the issues and problems of justice in the international arena. But what kind of cosmopolitan account is involved? And how such an account can be defended? In particular, how can it respond to the charge of being utopian? And does it drive states to the periphery and bring individuals to the centre? These are the challenges that Gillian Brock attempts to meet in this book.

Brock argues for a cosmopolitan model of global justice which takes the moral worth of individuals seriously, without ignoring the role that a defensible form of nationalism can play in this regard. In this, she aims to address two kinds of scepticism. One is the argument that cosmopolitanism interferes illegitimately with the defensible scope of nationalism and leads to the weakening of national values, such as authentic democracy, national self-determination and state sovereignty. The second argues that cosmopolitanism is unrealistic, unfeasible, a merely theoretical vision of the world. Brock's aim is to prove these two kinds of scepticism wrong.

Brock's response to skepticism about feasibility is based on formulating certain public policies, in the middle chapters of the book, which range from monitoring the protection of basic liberties and regulating humanitarian intervention to implementing a system of global taxation, immigration policies and global economic arrangements. She declares at the outset that she intends these policies and the theory of global justice on which they are based to meet Nagel's challenge to political theorists – that of formulating "workable ideas" of justice for global and international institutions "in the long run" (Nagel 2005).

Cosmopolitan accounts of global justice are inspired by Rawls's theory of justice, which is essentially meant to apply to the basic structure of a society and function within the bounds of a state. However, some liberal egalitarian philosophers have painstakingly extended Rawls's principles of justice to the global arena. Brock is certainly one of those, however, with some reservations. Some cosmopolitans have imagined a global basic structure arranged according to these principles, and argued that cosmopolitans should endorse a global difference principle and a principle of global equality of opportunity. However, Brock argues against such proposals and formulates an alternative *needs-based minimum floor* principle.

But, why assume that Brock's alternative principle is the optimal one? To argue for it, Brock conjures up a Rawlsian–style normative thought experiment regarding a global conference of delegates selected to choose the principles that regulate the fair terms of cooperation among the world's inhabitants. Brock claims that her thought experiment differs from Rawls's idea of the original position, since hers is not a device of representation. But, if the delegates are instead to be randomly selected, as Brock insists, then one wonders what legitimacy their decisions could have for other people, who have not selected these delegates. Nonetheless, as with Rawls's original position, Brock's thought experiment puts delegates behind a veil of ignorance, which denies them information about the demographics of world population, how powerful people are and the natural wealth and size of territories. Two primary rights are the concern of

the delegates: that everyone should enjoy some equal basic liberties, and that everyone should be protected from risks of serious harms, their basic needs being met. Brock argues for a list of freedoms that delegates would opt for, including freedom of dissent, conscience, speech, and the freedom to exit a society as well as minimum guarantees against assault, torture, imprisonment without trial, or extreme coercion of various kinds. The other claim which Brock makes is that delegates would not choose a world government, contrary to what cosmopolitans might propose. According to Brock, delegates would choose to retain states for two reasons. First, they would find it prudent to be risk-averse given the gravity of the situation, and so they are more cautious about making their decisions. And second, they would be concerned about the possibility of an overwhelming disastrous result if world government turned out to be bad.

Brock thus argues that delegates would not go beyond those social and political arrangements that guarantee our basic needs. She then asks, do we want more? She answers in the negative, and argues that global difference and equal opportunity principles would not be chosen in the thought experiment. First, regarding the difference principle, she appeals to the empirical work done by Norman Frohlich and Joe Oppenheimer, who conducted experiments regarding the choice of principles of distributive justice under conditions of impartiality. They offered four principles to participants: the two key ones being Harsanyi's principle of maximizing the average income and Rawls's difference principle, and the others being the principle of maximizing the average with a floor constraint, or specified minimum income, and the principle of maximizing the average with a range constraint, or specified maximum difference between the poorest and the richest. These experiments were repeated in different countries to ensure generality. The outcome was that the majority in all countries chose the principle with the guaranteed floor constraint. Around "78 per cent chose the floor constraint principle, 12 per cent chose to maximize the average income, 9 per cent chose the range constraint principle, and 1 per cent chose the difference principle." (55) The most revealing point about the experiments was that people under conditions of impartiality would not, as Rawls argued, choose to maximize the situation of the worst off. What they were more concerned about in their negotiation for a fair principle is that the minimum floor should be set so as to reflect both entitlements and incentives. Although Rawls did not intend his difference principle to apply at the global level with which Brock is concerned, these experiments obviously tell against the principle even at the national level.

Regarding the extension of the fair equality of opportunity principle to the global level, Brock argues that it also faces significant problems. Most notably, it is not sensitive to cultural diversity, since certain positions may be valuable to a certain culture, but not to another. She argues that recent attempts to revise this principle have been unable to address the differences of power that will result from equalizing standards of living. Brock's conclusion is that what is most important is not that people have an equal set of opportunities, but that they have a *decent* set of opportunities.

Brock then argues that her needs-based approach is similar to Amartya Sen and Martha Nussbaum's capability approach and that Nussbaum's list of human capabilities depends heavily on needs. Although one can see the point Brock makes about the resemblance between the two approaches, the capability approach is a wider notion that includes needs along with a set of fundamental entitlements held to be necessary for human dignity.

Having thus responded to skepticism about the feasibility of her cosmopolitan-

ism, Brock turns to the second doubt about cosmopolitanism, that raised by liberal nationalists such as Yael Tamir and David Miller, who argue that cultural membership is the precondition for individuals' choice and thus their autonomy, and that cosmopolitanism interferes with this. Brock focuses on two theses shared by liberal nationalists regarding obligation and rights: first, The "associative obligation" thesis that national membership is important to people's well-being as it creates ties, which in turn generates obligations; and second, the right of fellow members to be helped outweighs those of non-members when the gap between their needs is not too great. The right of fellow members to be helped can be outweighed, based on moral responsibility to all humans, only if the needs of non-members are far greater than those of fellow members. For Brock, the problem with the liberal nationalist account of moral responsibility is that it is based on a model of personal identity which is based on a sense of community and belonging which fits with the idea of a moral duty to help those who are in dire need. Liberal nationalists therefore attempt to affirm the moral worth of all individuals and our duty to assist the distant needy, while at the same time restricting this duty subject to national and communal cultural ties, which makes their account of responsibility inconsistent.

Brock's argument against liberal nationalists is based on her normative account of global justice, according to which delegates in the thought experiment would not choose a principle that favours fellow members over non-members, since one can end up in a resource-poor community and in need to be assisted by non-national members. She also argues that nationalist differentiation is unable to deal with global problems, since global organizations are more likely to be effective in dealing with matters of injustice, poverty and suffering.

Following this critique of liberal nationalism, in the penultimate chapter Brock states that all "forms of cosmopolitanism have in common a commitment to our equality" (298). Having rejected a global principle of equality of opportunity for the sake of the idea of a decent set of opportunities, she explains that the kind of "equality" that is most consistent with her cosmopolitan account is Elizabeth Anderson's version of democratic equality, since both are committed to making capabilities, in the sense of a person's freedom to achieve valued functionings, equally available for everyone. Anderson's fundamental thesis of democratic equality is that people should stand in relations of equality with each other, and to "live together in a democratic community," in forms of "collective self-determination by means of open discussion among equals." This means that "one is entitled to participate, that others recognize an obligation to listen respectfully and respond to one's arguments, that no one need bow and scrape before others or represent themselves as inferior to others as a condition of having their claim heard" (Anderson 1999, 313). Although Brock clearly realizes the difficulties of extending the model of democratic equality to her account of global justice, she argues that these can be surmounted.

In setting out her cosmopolitan account of global justice, Brock calls her version of cosmopolitanism "quasi-institutional," which aims not at radically "reconstructing" current global institutional orders, as "institutional cosmopolitans" would have it, but at "renovating" them. At the heart of her cosmopolitan account –and, in fact, of all cosmopolitans– there lies a commitment to the equal moral worth of human beings *as such*, and this is what makes her account lean towards moral cosmopolitanism. In a strong formulation of her needs-based account, she claims that all human rights set

out in the Universal Declaration of Human Rights can be explained in terms of this account (72). Although Brock engages deeply, in the book, with public policy and institutional changes, it is her needs-based approach, which she insists is the same as Nussbaum's capability approach that makes little space for a practice or institutional-based approach to global justice.[1]

It goes without saying that Brock's book makes a serious contribution to the theme of global justice as she comes up with many original insights. It is her academic passion for the subject that she puts forward compelling arguments that are hard not to be convinced by.

Dara Salam
LUISS University, Rome

REFERENCES

Anderson. Elizabeth. 1999. What is the Point of Equality. *Ethics* 109 (2): 287-337.
Beitz, Charles. 2009. *The idea of Human Rights*. Oxford: Oxford University Press.
Nagel, Thomas . 2005. The Problem of Global Justice. *Philosophy and Public Affairs* 33 (2): 113-47.

1] For a critique of Nussbaum's capability approach to human rights, see Beitz 2009, 62-8.

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 brancoveanu@publicreason.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 office@publicreason.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 brancoveanu@publicreason.ro.

Permission: *Public Reason* holds the copyright of all published materials. Requests for permission to reprint material from *Public Reason* should be sent to brancoveanu@publicreason.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.ubfilosofie.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

ARTICLES

BOOK REVIEWS

http://publicreason.ro

ISSN 2065-7285
EISSN 2065-8958