

PUBLIC REASON

Journal of Political and Moral Philosophy

Volume 2, Number 2, December 2010

PUBLIC REASON

Journal of Political and Moral Philosophy

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

Editors

Editor in Chief

Romulus Brancoveanu, *University of Bucharest*

Associate Editor

Thomas Pogge, *Yale University*

Editorial Team

Assistant Editor

Mircea Tobosaru, *University of Bucharest*

Laurentiu Gheorghe, *University of Bucharest*

Dorina Patrunsu, *University of Bucharest*

Editorial Board

Ovidiu Caraiani, *University Politehnica of Bucharest*

Luigi Caranti, *University of Catania*

Radu Dudau, *University of Bucharest*

Mircea Dumitru, *University of Bucharest*

Adrian - Paul Iliescu, *University of Bucharest*

Ferda Keskin, *Istanbul Bilgi University*

Valentin Muresan, *University of Bucharest*

Mihail - Radu Solcan, *University of Bucharest*

Constantin Stoenescu, *University of Bucharest*

Ion Vezeanu, *University of Grenoble*

Advisory Board

Sorin Baiasu, *Keele University*

Radu J. Bogdan, *Tulane University*

Paula Casal, *University of Reading*

Fred D'Agostino, *University of Queensland*

Cecile Fabre, *University of Edinburgh*

Rainer Forst, *Goethe University, Frankfurt am Main*

Gerald Gaus, *University of Arizona*

Axel Gosseries Ramalho, *Catholic University of Louvain*

Alan Hamlin, *University of Manchester*

John Horton, *Keele University*

Janos Kis, *Central European University, Budapest*

Jean-Christophe Merle, *University of Tübingen*

Adrian Miroiu, *SNSPA Bucharest*

Adrian W. Moore, *University of Oxford*

Philippe Van Parijs, *Catholic University of Louvain*

Mark Timmons, *University of Arizona*

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

ISSN 2065-7285

EISSN 2065-8958

Print version published by Editura Universității din București for *Public Reason*.

© 2010 by *Public Reason*

PUBLIC REASON

Journal of Political and Moral Philosophy

Vol. 2, No. 2, December 2010

ARTICLES

- Kalle Grill*
Anti-paternalism and Invalidation of Reasons 3
- Enzo Rossi*
Modus Vivendi, Consensus, and (Realist) Liberal Legitimacy 21
- Ina Kerner*
“Scales of Justice” and the Challenges of Global Governmentality..... 40
- Endre Begby*
Rawlsian Compromises in Peacebuilding? Response to Agafonow 51
- Matthias Katzer*
The Basis of Universal Liberal Principles in Nussbaum’s
Political Philosophy 60
- Michael Harbour*
Religious Toleration and Public Funding for Abortions: a Problem with
Christopher Eberle’s Standard of “Conscientious Engagement” 76
- Asger Sørensen*
Dussel’s Critique of Discourse Ethics as Critique of Ideology 84

BOOK REVIEWS

- Amartya Sen, *The Idea of Justice*.....102
Reviewed by Stefan Bird-Pollan
- Axel Gosseries and Lukas Meyer (edited by), *Intergenerational Justice*105
Reviewed by Cédric Rio

Anti-paternalism and Invalidation of Reasons

Kalle Grill
Uppsala University

Abstract: I first provide an analysis of Joel Feinberg's anti-paternalism in terms of invalidation of reasons. Invalidation is the blocking of reasons from influencing the moral status of actions, in this case the blocking of personal good reasons from supporting liberty-limiting actions. Invalidation is shown to be distinct from moral side constraints and lexical ordering of values and reasons. I then go on to argue that anti-paternalism as invalidation is morally unreasonable on at least four grounds, none of which presuppose that people can be mistaken about their own good: First, the doctrine entails that we should sometimes allow people to unintentionally severely harm or kill themselves though we could easily stop them. Second, it entails that we should sometimes allow perfectly informed and rational people to risk the lives of themselves and others, though they are in perfect agreement with us on what reasons we have to stop them for their own good. Third, the doctrine leaves unexplained why we may benevolently coerce less competent but substantially autonomous people, such as young teens, but not adults. Last, it entails that there are peculiar jumps in justifiability between very similar actions. I conclude that as liberals we should reject anti-paternalism and focus our efforts on explicating important liberal values, thereby showing why liberty reasons sometimes override strong personal good reasons, though never by making them invalid.

Key words: anti-paternalism, invalidation, Joel Feinberg, liberty, reasons.

In his argument against paternalism, Joel Feinberg states at one point that the anti-paternalist “must argue that paternalistic reasons [...] are morally illegitimate or invalid reasons” (1986, 25-26). In this article, I will develop an account of what it might mean that a reason is invalid on moral grounds in this sense, or *invalidated*, as I will call it. I will then go on to present four arguments against invalidation in the context of anti-paternalistic doctrines, and by implication against invalidation in general.

I will be concerned with the principled or deontological understanding of anti-paternalism, which tells us to disregard certain reasons as a matter of moral requirement, rather than of practical expedience. There are often practical reasons to disregard certain reasons: Lack of time or information, risk of mistake, social coordination mandating a division of responsibility, and so on. Feinberg addresses his anti-paternalism to “an ideal legislator” and claims to be on “a quest not for useful policies but for valid principles.” (1984, 4) I take it that it is on this ideal, general or abstract level that reasons can be invalidated, as opposed to excluded for practical reasons.¹ Though practical considerations and more principled arguments have as a rule been intertwined in the discussion on paternalism since John Stuart Mill set the standard with *On Liberty*, I find it preferable to keep them separated. We can fruitfully discuss paternalism either in the abstract or in some concrete

1] Joseph Raz' (1990 [1975]) account of “exclusionary reasons” as reasons not to act on other reasons is based on coordination problems and on practical constraints on deliberation. I take invalidation to be more purely normative. The ideas also differ in that it is a good thing to act in accordance with excluded reasons, though not for them, while invalidated reasons are more thoroughly emptied of normative significance.

situation or institutional setting. Anything in between is likely to fail both to account for the normative core of the problem and to provide action guidance.

Feinberg oscillates between understanding paternalism on the one hand as the counting of certain reasons as good and valid reasons for limiting liberty (as in the quote in the first paragraph of this section), and on the other as limitations of liberty with a certain (implicit) rationale. Though the latter understanding is dominant in contemporary discussion of paternalism, I believe the former is the most promising for the anti-paternalist. It is less ambitious to oppose certain reasons for limiting liberty, than to oppose certain actions or policies because they limit liberty and are supported by certain reasons.²

Feinberg's uncompromising defence of anti-paternalism on moral grounds is unusual. The standard approach to paternalism among contemporary authors is to assume anti-paternalism as a general rule and then propose exceptions to this rule. These exceptions may invoke hypothetical consent (e.g. Van De Veer 1986, 88), avoidance of great harm (e.g. Groarke 2002) or the common-sense reasonableness of preventing significant harm at limited cost to liberty (e.g. Beauchamp and Childress 2001, 185-91). There has been a lack, however, both of detailed analysis of principled anti-paternalism and of more comprehensive arguments against it. This may be due to the strong focus on practical and political circumstances and to the perceived need to decide which actions and policies are rightly called paternalistic. Recently, some authors have presented more head on critiques of anti-paternalism, most noteworthy Richard Arneson (2005) and Peter de Marneffe (2006). With this article, I hope to join that effort and take it one step further.

I will focus mainly on Feinberg's absolutist understanding of anti-paternalism according to which all paternalistic reasons are invalid. I will also, however, consider moderate versions according to which paternalistic reasons are discounted rather than invalidated, or where exceptions are made for certain types of paternalistic reasons. Throughout I discuss moral reasons. Reasons that are invalidated as moral reasons may possibly remain valid as reasons of some other sort (e.g. prudential).

Three of my four arguments are based on the fact that anti-paternalism necessarily only applies to sufficiently voluntary action. This, I argue, leads to wrong conclusions in some cases, to peculiar jumps in justifiability, and to an unwarranted disregard for the liberty of those whose actions are, according to the doctrine, insufficiently voluntary. The remaining argument is based on the fact that we can have decisive reasons to interfere with a person who is acting perfectly voluntarily, and she can accept these reasons fully as far as they concern her, yet anti-paternalism unreasonably entails that these reasons are invalid. All of my arguments concern the peculiar effects of invalidation per se. None of them depend on people being mistaken about their own good. Though I happen to believe

2] For an extensive defence of this view, see Grill 2007.

that people can be so mistaken, I will assume throughout, for the sake of argument, that we should accept people's views of their own good at face value.

I. ANTI-PATERNALISM AS INVALIDATION

I take it for granted that if an action protects or promotes some person's good (more than alternative actions) that is normally a valid reason for that action. I will from now on call such reasons *personal good reasons* rather than paternalistic reasons, since they can be invoked for actions that do not limit liberty and since, strictly speaking, reasons as such cannot be paternalistic. Personal good reasons may concern such things as a person's health, prosperity, achievement, happiness, or long-term autonomy. I will not be concerned with what exactly is good for a person, but rather, as already stated, assume that her own view on this matter should be accepted. Neither will I make a distinction between protection and promotion of good – that is between preventing harm and providing benefit. If the reader thinks that such a distinction is warranted, she may read my argument as concerning harm-prevention.

Anti-paternalism is, on my understanding, a doctrine of invalidation. We may describe invalidation as the blocking of a reason from influencing the moral status of an action (forbidden, permissible, obligatory etc.) and so what we ought to do (not do, may do etc.). To influence an action's moral status is, I take it, to have weight on the scales, to figure among the factors that should be considered when forming an all things considered judgement (under ideal conditions). A reason may have influence in this sense even if it is ultimately overridden by other reasons. We could say with John Broome (2004) that the reason figures in a "weighing explanation" of an ought fact. However, while Broome says that deontic principles replace weighing explanations (making them merely potential), I propose that some deontic principles operate by regulating the weighing, for example by making certain reasons invalid, banishing them from the scales. The picture of reasons on the scales is of course metaphorical. I admit that I do not have a theory of how exactly to derive an all things considered judgment from a set of valid reasons. I can only say, with Broome, that such a judgment should be based only on consideration of the strengths of the relevant reasons (or more Broomian – ought facts are determined by the aggregated weights of the relevant reasons) (37-38).

Invalid reasons are normatively impotent and in that sense not reasons at all. This presents a terminological problem. On the one hand, we are very tempted to talk, like Feinberg, of invalid reasons. On the other hand, it is a widespread position that the term reason should be reserved for valid reasons, for reasons with influence (e.g. Scanlon 1998, 156; Kamm 2006, 237; Parfit Manuscript). Philosophers sometimes disagree on what we should do, or at least why we should do it, though they agree on what is important. For example, non-consequentialists typically agree with hedonistic utilitarians that the fact that an action will alleviate suffering is a relevant consideration in a sense that most other consequences are not, even if they hold that this consideration does not always

provide a reason for action. So for example, Thomas Scanlon (1998) in his critique of a morality based on wellbeing readily admits: “It would be absurd to deny that well-being is important” (141). Since in the present context the very object of investigation is whether apparent reasons are valid reasons, I find it preferable to use “reason” for all considerations that are relevant in this wide sense. A reason’s strength is the influence it would have if it were valid.³ I find that this terminology clarifies the role of deontic constraints on moral reasoning, for example the constraints of anti-paternalism, and so one important difference between consequentialism and non-consequentialism.

Principles or doctrines of invalidation can be placed within a larger family of influence-regulating doctrines – doctrines that strike a wedge between the strength of a reason and its influence. The most widely recognized form of morally based influence-regulation is side constraints (introduced by Robert Nozick (1974) in the context of rights), or, in other words, absolute reasons. An absolute reason entails that an action should or should not be done, irrespective of other reasons.⁴ This a priori-like property of a reason entails that other reasons have no influence – they should be disregarded. As a relationship between reasons, invalidation is distinct from and weaker than the relationship between absolute and non-absolute reasons, since it entails only that some reason(s) have no influence, leaving other reasons to potentially override the invalidating reason. In fact, absolute reasons can be defined as reasons that invalidate all other reasons, with invalidation the more basic concept.

Another form of morally based influence-regulation can be found by applying the idea of lexical ordering to types of reasons. Lexical ordering or priority was introduced in moral theory by John Rawls (1971) in application to principles of justice and has later been explored in application to values (e.g. Griffin 1986). It can similarly be applied to reasons. Lexical ordering allows reasons to be absolute in relation to reasons lower in the hierarchy while standing in a weighing relationship to reasons on the same level. Invalidation is in one sense stronger and in one sense weaker than lexical ordering.

Invalidation is weaker than lexical ordering in that it is not hierarchical. Reason of a type that is invalidated by another type can still override reasons of a third type, that can override reasons of the invalidating type. For example, assume that reasons of type H are invalidated by reasons of type L. Now there may be other reasons, say of type O, which can override reasons of type L, but which can in turn be overridden by reasons of type H.

3] I aim to follow the standard use of “strength” (sometimes “weight”) employed by e.g. Raz (1990 [1975]). Joshua Gert (2007) has argued convincingly that reasons have two distinct dimensions of strength – requiring strength and justifying strength. I will disregard this complication since my argument holds for both kinds of reasons. I will however avoid talk of the “balance of reasons” and stay with the more general “influence the moral status of actions”. When I say that one reason is stronger than another, this should always be understood as concerning the same kind of strength, typically requiring strength.

4] The term from e.g. Raz 1990 [1975], 27. Such reasons are sometimes called “decisive”, but “absolute” is preferable because “decisive” other times refer to what Raz calls “conclusive” reasons – reasons that are not overridden in a certain case (though they can be).

In fact, this is exactly what anti-paternalists typically claim. H reasons (to prevent harm to a person) are invalidated by L reasons (to respect the liberty of the same person), while O reasons (to prevent harm to another person) can override L reasons, in accordance with the harm principle. Further, H reasons can override O reasons since it can be right to prevent serious harm to one person instead of preventing less serious harm to another (when there are no liberty reasons either way). In contrast, reasons of a type that is lexically prior to reasons of a second type cannot be overridden by reasons of a type that can be overridden by reasons of the second type. For example, that Rawls' basic freedoms and liberties are lexically prior to concerns of distributive justice means that there can be no type of reason that override liberty reasons while being overridden by distributive justice reasons.

Invalidation is stronger than lexical priority in that it not only makes one reason dominate another, it makes the dominated reason completely impotent. One important point of lexical ordering is to capture cases in which reasons at one level of priority balance each other out, in which case the matter is decided by reasons on the next lower level, unless they also balance each other out, and so on. If a reason is invalidated by another reason, however, it has no influence and cannot decide anything.

My account of invalidation is an attempt to explicate a central aspect of the anti-paternalist position. A doctrine of invalidation should ideally be specified to a class of actions and a class of reasons. The doctrine then says that reasons in the relevant class do not influence the moral status of actions in the relevant class. This might have the further implication that being motivated by such a reason to perform such an action, or accepting it as justification for such an action, is inappropriate or condemnable in a way that warrants disapproval or punishment, perhaps because it manifests a bad attitude of some sort. However, the more basic idea is the normative invalidity of the reason. Whether failure to see this invalidity is immoral and if so what is the appropriate response to such immorality – these are secondary questions.⁵

II. INTERFERENCE AND VOLUNTARINESS

Anti-paternalism is, on my understanding, the influence-regulating doctrine that personal good reasons are invalid for a certain class of actions. Call these actions “problematic interferences.” On most accounts, actions in this class must be liberty-limiting in some sense. Anti-paternalism typically does not entail that personal good reasons are invalid for innocuous involvement in other people’s life, such as greeting them in the street or giving them small gifts.⁶ Specifying this criterion of problematic

5] This means that I disagree with accounts of paternalism which claim that a defining criterion of paternalism is the manifestation of a certain attitude. That is if our concept of paternalism is supposed to capture what liberals are traditionally opposed to.

6] Sunstein and Thaler (2003) argue for what they call “libertarian paternalism”, postulating that “a policy therefore counts as ‘paternalistic’ if it attempts to influence the choices of affected parties in a

interferences is difficult, but I will not discuss these difficulties here. Nor will I consider what criteria there are exactly. Instead, I will focus on the voluntariness criterion for inclusion in the class of problematic interferences, and assume that whatever other criteria there are, are fulfilled.

On any reasonable account, anti-paternalism protects only choices or actions that are *sufficiently voluntary* (Feinberg 1986, chapter 20). It is not problematic interference to restrain people in rage or panic, people heavily influenced by drugs, people with severe mental disorders, or small children (unless, when applicable, this is in conflict with their previous, voluntary choice). Importantly, voluntariness cannot be determined by the reasonableness of choices or actions (Feinberg 1986, particularly chapter 20, e.g. p. 133). The very point of anti-paternalism is to protect imperfectly reasonable choices.

For Feinberg, a person acts perfectly voluntarily if she is competent, there is no coercion or duress, no subtle manipulation, no ignorance or mistake and no distorting circumstances (such as excitement, strong emotion, or time pressure) (1986, 115). Of course, hardly any choice is perfectly voluntary. Feinberg explicitly proposes that a person should be protected from paternalism if her actions or choices are “voluntary enough” (chapter 20). Given that the other criteria for problematic interference are fulfilled, anti-paternalism entails that sufficient voluntariness functions as an invalidator of personal good reasons. In other words, liberty reasons invalidate personal good reasons on the condition that the person acts sufficiently voluntary.⁷

III. FIRST ARGUMENT – VOLUNTARY CHOICE CAN LEAD TO DISASTER

In this and the following sections, I will present a series of scenarios. The scenarios include assumptions on what reasons there are and their relative strength. I will not defend these assumptions. I take it for granted that the specifics of the cases can be adjusted so that these assumptions are reasonable. Remember that the way I use the term reason, reasons need not have influence. That a reason for an action has a certain strength means only that we can promote or protect some value by performing that action. If the reader

way that will make choosers better off.” (1162) A policy of providing information upon request because this makes for wise choices would thus be paternalistic. Though it is possible to oppose even such non-interfering policies I will focus on the more plausible because more restricted anti-paternalism that targets only liberty-limiting benevolence.

7] Interference with insufficiently voluntary choice is sometimes called “soft paternalism” but is typically not opposed by anti-paternalists. Richard Arneson (2005) claims that soft anti-paternalists, accepting soft but not hard paternalism, occupy an unstable position between full acceptance of paternalism and hard anti-paternalists, who allegedly do not think that lack of voluntariness makes interference any more justified (266-68). However, even hard anti-paternalists must embrace some of the components of Feinberg’s concept of voluntariness, minimally competence and lack of duress – infants and people at gunpoint may certainly be stopped from harming themselves. The distinction between hard and soft anti-paternalism, though useful, is ultimately a matter of degree. There is no remotely reasonable end point position to occupy. (One could of course change terminology by claiming that interference with non-voluntary action is not paternalism at all, but this changes nothing of substance, see further below in the section “VI. Third Argument – Liberty for all”.)

finds that there is no possible specification under which my scenarios make sense, this must be because she has a radically different view on what has value.

Voluntariness often has an indirect effect on the moral status of actions, since a high degree of voluntariness normally makes for good decisions, which normally makes for good consequences. When things are not normal, however, a person's very voluntary choice can lead to catastrophic consequences. Consider:

The Bridge. A person tries to cross a bridge but we have a chance to stop her. We know that the person wants to live and is well informed about the general condition of the bridge, is acting in character, is calm and collected, attentive, mature and intelligent, under no duress or pressure, etc. Stopping her would interfere with her liberty to move around freely, which is a strong reason against doing so. However, being equally well informed about the condition of the bridge, and having considered its durability more thoroughly, we firmly believe that, appearances to the contrary, the bridge is unsafe. Stopping the person would therefore most probably prevent her from falling to certain death, which is a much stronger reason for doing so.

We know that the person in *The Bridge* scores high on the standard aspects of voluntariness. Presumably, she just happens to be wrong, as very able people sometimes are. Perhaps she did not bother enough to analyze the available information, perhaps she miscalculated this one time and so reached the wrong conclusion. Whatever the cause, her very voluntary choice happens to be very bad for her. Anti-paternalism seems to imply that our reason to stop the person is invalid and so we should not stop her (unless there are other reasons to do so). That is the wrong conclusion – it is morally unreasonable. We should stop the person, because otherwise she will most probably die.

In his original bridge case, from which *The Bridge* is adapted, Mill argues that the person may be coercively turned back “without any real infringement of his liberty; for liberty consist in doing what one desires, and he does not desire to fall into the river.” (1991 [1859], 107) Even if it is true that the person does not want to fall into the river, this does not mean that he does not want to cross the bridge (for a thorough argument, see Day 1970). We certainly desire some things though we do not desire their consequences (indeed we may do so even when we anticipate the consequences).

Crossing the bridge is most probably in conflict with the person's goals and values. In that sense her behaviour rests on a mistake. This mistake is small in terms of information processing, but great in terms of consequences. Voluntariness could be defined as efficient goal-satisfaction, in which case the person in *The Bridge* probably acts non-voluntarily. However, on Feinberg's concept of voluntariness, ignorance of the consequences of one's action is only one of many voluntariness-reducing factors and comes in degrees. That is why I can claim that the choice to cross in *The bridge* is very voluntary, and that is why, as

we will see in the next section, Feinberg introduces a variable standard for how voluntary is sufficiently voluntary.⁸

In general, very small imperfections in people's decision-making process can make for disastrous decisions. If the consequences of an imperfect decision are bad enough and if these consequences can be avoided without losing anything of comparable value, they should be avoided. We can save the person in *The Bridge* because, in this particular case, we understand the consequences of her actions better than she does and we have an opportunity to intervene. With all the complexities of modern life and with all the expert knowledge and the sophisticated forms of intervention available, there are many such cases.

It may seem that we can save anti-paternalism by insisting that no good can be brought about by interfering with voluntary choice. This may be a conceptual argument, claiming that what a person voluntarily chooses is always, by definition, good for her. Or the argument may be empirical, claiming that because of our great ability to make good choices for ourselves, and our great inability to help others, the outcome of sufficiently voluntary actions can never be improved by intervention. Both assumptions seem incredible, raising the standards of sufficiently voluntary action to inhuman levels and so leaving most actual choices and actions outside of the doctrine's domain. That is unless we accept a very scattered or even inconsistent idea of the good (in the conceptual case) or a very pessimistic view of human benevolence (in the empirical case).

Importantly, if it were nonetheless true that interference with sufficiently voluntary choice could do no good, anti-paternalism of the form under investigation would be saved only at the price of redundancy. If no good can come from interference, we have no reason to interfere. Then there are no reasons for anti-paternalism to invalidate, and so the doctrine has no application.

A popular moderation of anti-paternalism makes exceptions for personal good reasons that concern autonomy (e.g. Dworkin 1972; Kleinig 1983). On such a moderation, anti-paternalists could plausibly defend stopping in *The Bridge* as autonomy-preserving (if autonomy is preserved by saving a life). Most any harm can diminish autonomy in some way to some extent and so this moderation can entail that some personal good reasons are always valid, making for a substantially weakened doctrine. However, given a more narrow conception of autonomy, cases can be constructed that are analogous to *The Bridge* but where the harm does not diminish autonomy (financial ruin or disfigurement or horrific but temporary pain may be substituted for falling to certain death).

8] Danny Scoccia (2008, 358) has argued that voluntariness on Feinberg's account is about successfully furthering one's own values, or acting as one would have done if one were perfectly informed, rational and capable. If this is true, Feinberg's anti-paternalism is not about letting people lead their own lives, but rather about helping them promote their values, through either interference or non-interference, depending only on what is most efficient. I think this interpretation fails to appreciate Feinberg's strong commitment to personal sovereignty.

Another form of moderation is to accept all personal good reasons as partially valid, though discounted by some factor (e.g. Groarke 2002, 219). Such versions will entail that we should stop the person in *The Bridge* if the reason to save her is strong enough and the discount factor large enough that the reason to respect her liberty is overridden. However, in similar cases, where the difference between the strength of the reasons for and against stopping is smaller, such versions will entail the morally wrong conclusion. The larger the factor (the smaller the discounting), the fewer such cases, and the less convincing. The most convincing cases must be constructed to match the exact discount factor. The larger the factor, the weaker the doctrine, and so, I think, the more reasonable.

In conclusion, any non-redundant anti-paternalism of substance will lead to morally wrong conclusions in some cases. Therefore, anti-paternalism should be rejected.

IV. FEINBERG'S COMPROMISE

According to Feinberg, what is a sufficiently voluntary choice varies with “the nature of the circumstances, the interests at stake, and the moral or legal purpose to be served.” (1986, 117) In particular, the threshold depends on the severity of the risks involved – the higher the risks, the higher the threshold – and on whether the risks include irrevocable harm (118-121). This is a very reasonable move to make for any anti-paternalist who, like Feinberg, is concerned not simply to “prevent people from acting with low degrees of voluntariness”, but rather to “prevent people from suffering harm that they have not truly chosen to suffer.” (119)

Tying the threshold for sufficiently voluntary action to risk amounts to a compromise between unmodified anti-paternalism and consequentialism. High risks normally entails strong reasons for interference. On Feinberg's account, personal good reasons do not support problematic interferences, but they do influence whether or not an action is a problematic interference. In this roundabout way, personal good reasons can override liberty reasons. However, the central structure of the doctrine remains intact: If an action is voluntary enough, personal good reasons are invalid.

The compromise makes anti-paternalism more flexible and so more reasonable. The anti-paternalist can now hold that we should stop the very voluntary person in *The Bridge* because she takes a very great risk, while we should let people who are less informed and rational go about their lives, taking smaller risks. However, the basic argument from *The Bridge* remains in force. The tiniest deficits in voluntariness can still make for disastrous consequences. Feinberg must accept that disaster is morally irrelevant as long as the degree of voluntariness is high enough to match the risk. Furthermore, high degrees of voluntariness need not correspond to great liberty interests. The compromise does not change the fact that high degrees of voluntariness invalidates very strong personal good reasons even when the voluntary choice is trivial from the point of view of liberty. This, I think, is unreasonable.

Severe risk should perhaps be accepted for the sake of important liberties, but not simply because choices are very voluntary. This point can be illustrated by comparing a well-planned philosophical suicide with a five-party game of Russian roulette. Feinberg's account entails that since the suicide is five times more risky than the Russian roulette, we should accept a much lower degree of voluntariness for the roulette. Given that the circumstances are similar in other respects, this seems absurd. What should count against stopping these activities is not so much the degree of voluntariness of the agents involved, but rather the sort of liberty at stake and its value.

V. SECOND ARGUMENT – JUSTIFIED INTERFERENCE WITH PERFECTLY VOLUNTARY ACTIONS

In *The Bridge*, we assumed that the person failed to appreciate the risks involved. Even if she scores very high on voluntariness, there are imperfections, which happen to be very important. Stopping presumably furthers her good and so does not conflict with her hypothetical, enlightened self-interest. An anti-paternalist impressed with the first argument above may restrict her doctrine to perfectly voluntary action. However, consider this case:

The Stunt. A person tries to perform a spectacular stunt but we have a chance to stop her. We know that the person is acting perfectly voluntarily. Stopping her would interfere with her liberty, which is a strong reason against doing so. Stopping her would also eliminate a substantial risk to her health, which is a relatively weak but non-negligible reason for doing so. In addition, stopping her would eliminate a small but real risk of harm to an innocent and non-consenting passer-by. This is a strong reason for doing so, but weaker than the reason against. In the aggregate, however, the two reasons for stopping are stronger than the one reason against.

Remember that the specifics must be filled in to make the assumptions reasonable. For example, the exact risk to the passer-by can be adjusted to match the assumed strength of the reason to protect her (we certainly accept some risk to innocent and non-consenting passer-bys even from very superficial activities such as driving around for fun).

The Stunt illustrates the important and, in the paternalism debate, neglected fact that there are normally several reasons for and against any particular action. That some of these reasons are strong in no way excludes the possibility that a weak reason is decisive (tips the balance). Anti-paternalism clearly implies that we should not stop the stunt (unless there are other reasons to do so), since the personal good reason is invalid. That is the wrong conclusion. We should stop the person because of the risks to herself and to the passer-by.

It might seem surprising that we can have overriding reason to stop a perfectly voluntary action. The person obviously considers the reasons for performing the stunt

stronger than those against. Since, by assumption, she also has a correct view of her own good, we cannot question her reasons for performing the stunt, nor her self-interested reasons against. However, we can disagree about her reason to avoid risking the health of the innocent passer-by. She may have little concern for the wellbeing of some passing stranger, while we (correctly) judge that this consideration decides the matter against performing the stunt.

By assumption, there must be agreement on the person's self-interested reasons for and against acting. More importantly, however, the person may very well agree with us concerning those of our reasons for stopping her that concern her. Though she values her own health higher than other people's, she may still agree with us on the exact strength of our reason to stop her (weak but non-negligible). She may also agree that this reason is valid. It might well be that she opposes our intervention only because she thinks that we exaggerate the strength of our reason to protect the passer-by. Anti-paternalism therefore implies that we should disregard as invalid a reason that concerns the good of a person, even though she is herself in perfect (and perfectly voluntary) agreement with us concerning the strength and validity of this reason.⁹

It could perhaps be argued that we have no reason to eliminate voluntarily assumed risks to start with. This is not a view about invalidation then but about what has value. It is an unreasonable view. People may regret having to choose between two risky options and still make a choice, voluntarily. If the circumstances are harsh, the risks may be great. If we can do something to lower these risks, without losing anything of comparable value, we should do so. If you voluntarily choose to risk your life driving to work by route A over risking your life by driving to work by route B, this in no way implies that I (working with city planning) have no reason to try to decrease the risks involved in driving either route.

As for moderations, substituting discounting for invalidation can entail that we should stop the stunt. If protecting the person from herself is a valid but discounted reason for stopping, it may, together with the reason to protect the passer-by, override the liberty reason against. However, there will be similar cases, with the relative strengths of the three reasons adjusted to match the discount factor, where such moderate anti-paternalism entails that we should not stop the stunt. As in *The Bridge*, the larger the discount factor, the weaker the doctrine, and the less convincing the counter-examples.

By modifying *The Stunt* we can see how moderation by discounting is in one way sharply distinct from absolute anti-paternalism. Assume that the liberty reason and the protection of the passer-by reason are equally strong. Now the case is a sort of moral dilemma if the personal good reason is invalid. However, as long as this reason has some influence, however small, we should stop the stunt. Any discounting version, regardless of the discount factor, will give the same answer. Absolutist anti-paternalists, on the other

9] This could be avoided if perfect voluntariness entailed correct judgment of the relative strength of all reasons. This, however, seems an unwarranted conflation of concepts and would only mean that anti-paternalism is redundant for perfectly voluntary action.

hand, will not like the idea that the personal good of the stunt artist should decide the case if favour of intervention.

In conclusion, we should sometimes interfere with people in order to protect or promote their good, as well as the good of others. We should do so even if the action we interfere with is perfectly voluntary and we accept people's view of their good at face value. We should do so because we have overriding reasons to, both reasons concerning the good of others, and reasons concerning the good of the person. Anti-paternalism entails that we should disregard or discount the latter reasons, even if the person in question is in informed and perfect agreement on the strength and validity of these reasons. In effect, anti-paternalism tells us to let people harm themselves in ways they would not if they were not mistaken about their reasons to prevent harms to others. This is peculiar and arguably in conflict with liberal fundamentals. Therefore, anti-paternalism should be rejected.¹⁰

VI. THIRD ARGUMENT – LIBERTY FOR ALL

The first two arguments both accuse anti-paternalism of unacceptable disregard for important personal good reasons. Somewhat paradoxically, anti-paternalism may indirectly lead to disregard also for important liberty reasons. The doctrine only protects sufficiently voluntary choices. I propose that liberty is important not only for the most informed and capable decision-makers, but also for the demented, for minors, for the ignorant, and for people under time pressure. It is not as if the value of controlling (to some extent) one's own life kicks in only at a certain degree of voluntariness. Perhaps there is some level under which people cannot choose for themselves or cannot appreciate self-determination. Liberty, however, has value for people that are well above this level but that we should nonetheless sometimes coerce in their own interest, for example people in their lower teens. With this, anti-paternalists should agree. Why then does not interference with the liberty of young teens activate the doctrine?

Three answers are possible for the anti-paternalist: First, she can insist that the liberty of young teens (and generally people acting under the sufficient degree of voluntariness) is of another type than the liberty of informed and rational adults and so does not activate the doctrine. This distinction in value is mysterious. The anti-paternalist can say that the important value is not liberty but autonomy, and that young teens are not fully autonomous. They may not be, but neither are many adults, and young teens are surely autonomous to some extent. There is no difference in kind between the self-determination of more and less capable decision-makers.

Second, the anti-paternalist can lower the threshold and claim that the liberty of young teens (etc.) should never be limited in their own interest, and in fact that their well-

10] De Marneffe (2006, 77-79) argues against anti-paternalism that since the government can justifiably and without insult substitute its judgement for an individual's for the sake of others, it can do so for her own sake too. My argument, in contrast, does not presuppose that there is no morally relevant distinction between first and third party interests in this context.

being is not even a valid reason for limiting their liberty. We have seen the problems such a view entails even for capable decision-makers. Even those who stubbornly bite the bullet and let the people in *The Bridge* and *The Stunt* kill or harm themselves, and others, cannot reasonably accept such passivity in relation to young teens.

Third, the anti-paternalist can say that for young teens, the value of liberty is appropriately reflected in the strength of liberty reasons, and so there is no need for invalidation. This is the most reasonable answer. But if this is the anti-paternalist's answer, it is entirely unclear why things should be any different for adults. Interference with more capable decision-makers generally yields smaller rewards, since there is less room for improvement. It may also be that interference with more capable decision-makers has a greater cost in terms of liberty, because more liberty (or autonomy) is sacrificed in some sense. However, none of this indicates that personal good reasons should be invalid.

The argument does not depend on the arbitrariness of any threshold between more and less capable decision-makers. Nonetheless, it may seem easier to make a distinction in value (as in the first answer above) if there is a neutral or obvious level at which to draw the line. Anti-paternalists often do make a sharp distinction between competents and incompetents, between the healthy and the (mentally) ill, between adults and children. These categories, however, depend on underlying physical properties, which vary by degree. It may be thought that legal status provides a less arbitrary basis for a threshold. This is not so. Once bestowed, legal status may admittedly make a normative difference. It is perhaps worse to limit the liberty of a person of age, because this frustrates legitimate expectations not to be so treated that are induced by the legal system. However, such legal circumstances can only reinforce an underlying moral principle, which must be spelled out in terms of non-legal, concrete physical or psychological properties of persons. It would be hopelessly vacuous to argue that the people that we must protect from benevolent interference are those that have been granted a legal right to be so protected. A moral principle like anti-paternalism should help us decide upon such matters as when people should reach lawful age, and so cannot itself depend on the answers to such questions.

In conclusion, anti-paternalism fails to address the issue of whether or not to interfere with less than sufficiently voluntary choices made by rather autonomous people such as young teens. It makes no sense that the liberty of people who make more voluntary choices should trump other concerns, while the liberty of people who make somewhat less voluntary choices has no special moral status whatsoever. Therefore, anti-paternalism should be rejected.

VII. FOURTH ARGUMENT – JUMPS IN JUSTIFIABILITY

I will now present an argument that draws on the fact that voluntariness is entirely a matter of degree (as noted by Feinberg (1986, 104)). Arneson (2005) correctly observes that Feinberg's anti-paternalism is committed to "the enormous overriding importance of the line between self-harming choice that is not quite voluntary enough and choice

that just passes the threshold of being voluntary enough.” (268) Sometimes lines have to be drawn and sometimes much depends on whether or not a threshold is reached. For example, it is very reasonable that if a bridge is safe enough we will walk over it, while if it is not safe enough we will take a long detour and lose valuable time. This is reasonable even if the threshold is somewhat arbitrary and small differences may shift the status of the bridge from safe enough to not safe enough. For another example, if we want to categorize actions into justified and unjustified actions, there will arguably be a grey area of actions difficult to fit into either category. However, in the present case we are considering two theoretical perspectives, anti-paternalism or no anti-paternalism, only one of which demands that we draw a line at all. If drawing a line leads to unreasonable consequences, this speaks against the perspective that demands that we do so.

The problem with a threshold, however, is not only its unreasonably great importance in determining what we ought to do. The threshold does not determine directly the moral status of actions, but rather whether certain reasons have influence on this status. A valid reason makes an action more (or less) justified. Anti-paternalism implies, therefore, that an action that affects a person whose choice or action is just below the threshold of sufficient voluntariness may be much more justified than an otherwise similar action that affects a person whose choice or action is just over this threshold. At the threshold the justifiability of the action takes a ‘jump’.¹¹ The line gives rise to a gap, and a very large gap if the reason is a strong one. Consider:

The Suicide. A person tries to kill herself but we have a chance to stop her. Stopping her would interfere with her liberty, which is a strong reason against doing so. Stopping her would also save her life, which is a much stronger reason for doing so.

Anti-paternalism implies that whether or not we should stop the suicide depends on the person’s degree of voluntariness (unless there are other reasons to do so). If the voluntariness is under the threshold, we have much stronger reason to stop it than not, and so stopping is clearly and with good margin justified. If, on the other hand, the voluntariness is over the threshold, we have only, as far as valid reasons go, a strong reason not to stop it, and so stopping is clearly and with good margin unjustified. I propose that this jump in justifiability is unacceptable. In fact I doubt that we can bring ourselves to even comprehend it, at least when it depends on infinitesimal differences – one further tiny piece of information, or a tiny improvement in decision-making capacity, will imply that an action that would have been overwhelmingly unjustified becomes overwhelmingly justified. Moderation by discounting would make the jumps smaller but not affect the heart of the problem. Moderation by exception might make this particular example

[1] It may be thought that such jumps are common in the law, since it is much more justified to punish a person who has barely committed a crime than one who has not. However, I propose that this is due entirely to practical considerations such as the efficiency and transparency of the legal system. Disregarding such considerations, bare crimes are not much less justified than almost crimes.

irrelevant (if suicide diminishes autonomy for example), but there will be analogous examples with the same force.

On our assumption that people's views of their good should be taken at face value, it may seem that we cannot have a strong reason to stop the suicide, since the person herself has judged that her life should end. However, the person might agree with us on the value of her survival, but still want to kill herself for some higher purpose, that we find lacks value. Just like in *The Stunt*, there may be no disagreement concerning the value of the person's survival and liberty, but only concerning the value of some other thing, in this case something for which she wants to sacrifice her life, such as her family honour, the independence of Tibet, or the glory of God (assuming these things are not part of her good).¹² It may also be that the person agrees with us that she has stronger reasons to keep living, but feels compelled to end her life, perhaps out of despair. In such cases, of course, she is not acting perfectly voluntarily.

It may be suggested that anti-paternalism does not in fact draw a sharp line between problematic interference and other actions, since there is an area in between in which we simply do not know, or where it is genuinely indeterminate, whether an action is or is not an interference. If the indeterminacy is epistemic, the doctrine is none the more reasonable for it, only more difficult to apply. If the indeterminacy is real (ontological), however, the threshold is turned into a hole – it is really sometimes indeterminate whether or not personal good reasons give valid support to an action. Consequently, some answers to moral questions are replaced by no answers. The cases discussed in previous sections (*The Bridge*, *The Stunt*) involve people of very high and even perfect voluntariness, and so are hardly indeterminate. For the *Suicide*, however, the sudden jumps in justifiability could be replaced with a twilight zone of indeterminacy. Stopping would be overwhelmingly justified on one side of this zone, overwhelmingly unjustified on the other side, and indeterminate in between. We avoid jumps by giving up comprehensiveness. To my mind, this is no improvement.

Another and better strategy for avoiding a sharp threshold is to reformulate anti-paternalism as follows: Personal good reasons for an action are valid only *to the extent* that the action is not a problematic interference. In other words, the influence of a reason for an action is some function of two variables – strength and the degree to which the action is a problematic interference (which in turn depends partly on the degree to which it is voluntary). This modification replaces the absolute moral ban on paternalism with a sliding scale of gradual acceptance.¹³ The resulting partial invalidity must be distinguished from relative weakness. If the partial invalidation only means that personal good reasons

12] Another possibility is that there is perfect agreement on the strength of the reasons involved and that these reasons determine that the person should try to kill herself and that we should try to stop her. Such agreement may entail that stopping is not liberty-limiting and so not contrary to anti-paternalism. Under perfect voluntariness, this is perhaps only possible in group cases where there are coordination problems (see Arneson (1980, 471) on enforcing a general preference through prohibition).

13] Technically, gradual acceptance is equivalent to discounting with a variable discount factor.

are not very strong relative to other reasons, then paternalism is in fact fully accepted over the whole range of partial invalidation.

Just like indeterminacy, partial invalidity would not affect the arguments in the previous sections, since the voluntariness in those cases is very high or maximal, and so the validity of the personal good reasons very low. Jumps in justifiability, however, would be replaced with a sliding scale, and so stopping in *The Suicide* would be more justified the lower the degree of voluntariness. I therefore recommend this moderation to anti-paternalists, though it comes at the price of increased complexity.

In conclusion, absolute anti-paternalism gives rise to peculiar jumps in justifiability such that one action may be overwhelmingly justified while another, very similar action, is overwhelmingly unjustified. This is a strange moral landscape. Indeterminacy does not make the doctrine any more appealing. Reformulation of anti-paternalism as gradual invalidation of personal good reasons avoids the jumps, at the price of greater acceptance of personal good reasons and greater complexity. Anti-paternalism should therefore be rejected, or, possibly, modified.

VIII. CONCLUSION

I have presented four arguments against anti-paternalism understood as a doctrine of invalidation, regulating the influence of reasons on the moral status of actions. The analysis of anti-paternalism in terms of invalidation hopefully has some merit independently of the arguments.

The fourth argument shows that in order to avoid peculiar jumps in justifiability, anti-paternalism must be reformulated as a doctrine of partial and gradual invalidation. This weakening of the doctrine will alienate many traditional anti-paternalists who are principally opposed to counting personal good reasons as valid reasons for interferences with voluntary action. The weakening, furthermore, does not make the doctrine any more resilient to the other three arguments.

The third argument shows that anti-paternalism leaves us at a loss when considering interferences with insufficiently voluntary choice. Presumably such interferences should be evaluated by simply weighing reasons for and against. If that is so, however, it is unclear why sufficiently voluntary choice should be treated any differently.

The first and second arguments show that anti-paternalism sometimes entails morally wrong conclusions. In particular, the first argument shows that the doctrine entails that we should allow very voluntary choice to cause severe harm to self, even if nothing much is at stake in terms of liberal values. The second argument shows that the doctrine entails that we should let perfectly voluntary choice cause or risk severe harm to self and others, not because the agent rejects our reasons to interfere with her for her sake, but only because she rejects our reasons to avoid harms to others.

Taken together, I find that the four arguments amount to a strong case for abandoning the form of principled anti-paternalism defended by Joel Feinberg and implicitly accepted by many others. Very often we should let people make their own mistakes and suffer the

consequences. However, there is no moral principle that forbids limiting or interfering with a person's liberty for her good. In particular, there is no moral principle which makes reasons which concern central human values such as health and happiness invalid when they conflict with liberty reasons.

Sufficient voluntariness is a factor in many other contexts than paternalism. A claim that a person has a right to something often means that she may have or do this thing on the condition that her choice is sufficiently voluntary. The right to marry freely, to vote, to enter contracts – these rights are arguably conditional on voluntariness. Annuling or failing to accept a marriage, a vote or a contract does not violate a right if the parties concerned did not act voluntarily. We may want to annul or refuse to accept these things for personal good reasons, but also for the good of other people, or for whatever other reason. The arguments against anti-paternalism can therefore rather straightforwardly be used as prototypes for arguments against most any doctrine of invalidation. Though not as straightforwardly, the arguments may also inspire arguments against other forms of influence-regulation on moral grounds, such as side constraints and lexical ordering. I believe all morally based influence-regulation should be rejected, but have not argued for this more ambitious claim.

To oppose influence-regulating doctrines is not to deny that there are intricate relationships between reasons. Things of value may be empirically related in the sense that actions tend to affect many such things simultaneously. Things of value may also be conceptually related – value may for example be disjunctive in the sense that it is of value that one of several things happen. Rather than formulate influence-regulating doctrines, we should investigate the empirical and conceptual relationships between values and so between reasons.

When Isaiah Berlin tells us that “[t]he extent of a man's negative freedom is, as it were, a function of what doors, and how many, are open to him; upon what prospects they open; and how open they are” (2002 [1969], 41); when Amartya Sen develops his concept of freedom as capability (1992, chapter three) or when Joseph Raz develops his ideal of autonomy (1986); when Mill develops the notion of individuality (1991 [1859], chapter III), and even when he briefly states that “[t]he only freedom which deserves the name, is that of pursuing our own good in our own way” (17) – in all these cases we see significant contributions to our system of values, to our views on what is important in life. When, on the other hand, Mill tells us that “the sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of action of any of their number, is self-protection” (14), we are left confused as to what may then be the value of liberty and what other values there may be, that they should be related in this way.¹⁴

kalle.grill@yahoo.com

[14] This article has been long in the writing. I am sure I have some unrecognized debts. For very helpful comments, I wish to thank especially Sara Belfrage, Alon Harel and Lars Lindblom. Two anonymous reviewers for other journals exposed several weaknesses as well as the many ways in which my argument could be misunderstood. Hopefully it is now more transparent.

REFERENCES

- Arneson, Richard. 1980. Mill versus Paternalism. *Ethics* 90: 470-89.
- . 2005. Joel Feinberg and the Justification of Hard Paternalism. *Legal Theory* 11: 259-84.
- Beauchamp, Tom L. and Childress, James F. 2001. *Principles of Biomedical Ethics* (5th edition). Oxford: Oxford University Press.
- Berlin, Isaiah. 2002 [1969]. Five Essays on Liberty: "Introduction." In *Liberty*. Oxford: Oxford University Press.
- Broome, John. 2004. Reasons. In *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, eds. Jay Wallace et. al., 28-55. Oxford: Oxford University Press.
- Day, J.P. 1970. On Liberty and the Real Will. *Philosophy* 45: 177-92.
- De Marneffe, Peter. 2006. Avoiding Paternalism. *Philosophy and Public Affairs* 34: 68-94.
- Dworkin, Gerald. 1972. Paternalism. *The Monist* 56: 64-84.
- Feinberg, Joel. 1984. *Harm to Others*. Oxford: Oxford University Press.
- . 1986. *Harm to Self*. Oxford: Oxford University Press.
- Gert, Joshua. 2007. Normative Strength and the Balance of Reasons. *Philosophical Review* 116: 533-62.
- Griffin, James. 1986. *Wellbeing - its Meaning, Measurement and Moral Importance*. Oxford: Oxford University Press.
- Grill, Kalle. 2007. The Normative Core of Paternalism. *Res Publica* 13: 441-58.
- Groarke, Louis. 2002. Paternalism and Egregious Harm. *Public Affairs Quarterly* 16: 203-30.
- Kamm, F.M. 2006. *Intricate Ethics*. Oxford: Oxford University Press.
- Kleinig, John. 1983. *Paternalism*. Manchester: Manchester University Press.
- Mill, John Stuart. 1991[1859]. *On Liberty*. In *On Liberty and Other Essays*. Oxford: Oxford University Press.
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. Malden MA: Basic Books.
- Parfit, Derek. Manuscript. *Climbing the Mountain*.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge MA: Harvard University Press.
- Raz, Joseph. 1986. *The Morality of Freedom*. Oxford: Oxford University Press.
- . 1990[1975]. *Practical Reason and Norms* (2d edition). Princeton: Princeton University Press.
- Scanlon, Thomas. 1998. *What We Owe to Each Other*. Cambridge MA: Harvard University Press.
- Scoccia, Danny. 2008. In Defense of Hard Paternalism. *Law and Philosophy* 27: 351-81.
- Sen, Amartya. 1992. *Inequality Re-examined*. Oxford: Oxford University Press.
- Sunstein, Cass R. and Thaler, Richard H. 2003. Libertarian Paternalism Is Not an Oxymoron. *The University of Chicago Law Review* 70: 1159-202.
- Van de Veer, Donald. 1986. *Paternalistic Intervention*. Princeton: Princeton University Press.

Modus Vivendi, Consensus, and (Realist) Liberal Legitimacy

Enzo Rossi

Social Ethics Research Group, University of Wales, Newport

Abstract: A polity is grounded in a modus vivendi (MV) when its main features can be presented as the outcome of a virtually unrestricted bargaining process. Is MV compatible with the consensus-based account of liberal legitimacy, i.e. the view that political authority is well grounded only if the citizenry have in some sense freely consented to its exercise? I show that the attraction of MV for consensus theorists lies mainly in the thought that a MV can be presented as legitimated through a realist account of public justification. Yet I argue that, because of persistent ethical diversity, that realism problematically conflicts with the liberal commitments that underpin the very ideas of consensus and public justification. Thus, despite the interest it has recently attracted from critics of political liberalism and deliberative democracy, MV is not an option for those wishing to ground liberal political authority in some form of consensus. So if realist and agonistic critiques are on target, then the fact that modus vivendi is not an option casts some serious doubt on the viability of the consensus view of liberal legitimacy.

Key words: consensus, legitimacy, liberalism, modus vivendi, realism.

A framework for the exercise of political power is grounded in a modus vivendi when its main features can be hypothetically presented as designed and adjusted over time through a virtually unrestricted bargaining process between the competing individuals and groups that make up the society. In this paper, I consider whether a political framework¹ grounded in a modus vivendi could or should be appealing to theorists who subscribe to the consensus view of liberal legitimacy, i.e. to the view that liberal political power is well grounded and properly exercised only if it is in some appropriate sense acceptable to those subject to it. I argue that – *pace* a number of theorists – the idea of modus vivendi is not a viable account of the hypothetical agreement at the core of the consensus view. Moreover, I contend that the genuine appeal of modus vivendi to consensus theorists is symptomatic of a deep structural flaw in the consensus view of liberal legitimacy.

Crudely put, the attraction of the modus vivendi-approach for liberal consensus theorists can be understood as lying mainly in the idea that the agreement it envisages may be seen as more legitimate than the heavily moralized idea of an overlapping consensus of reasonable doctrines, in so far as it allows the consenting parties to resort to their actual values and commitments and to express them without restrictions. In the parlance of contemporary consensus theories of legitimacy, then, we may say that modus vivendi could be presented as enjoying legitimacy through public justification.

In this paper I discuss exactly how one may construe a conception of public justification that can present a modus vivendi – based political framework as legitimate – a move that aspires to retain the voluntaristic elements of the consensus view alongside

1] By “political framework” I mean what Rawls refers to as “the basic structure of society”: “constitutional essentials and basic questions of justice” (1993, 10).

a more realistic understanding of what constitutes a political consensus. However, I will argue that those desiderata cannot ultimately be reconciled: despite the pressing need to accommodate realist instances within liberal legitimacy theory, the idea of *modus vivendi* does not offer a viable *internal* corrective for consensus-based accounts of the foundations of liberalism. So the interest that *modus vivendi* has recently attracted from liberal realists, agonistic deliberative democrats and other critics of political liberalism is misplaced: *modus vivendi* is not an option for those wishing to ground liberal-democratic political authority in some form of consensus.

The paper is in four sections. Section I provides some background and anticipates some of my conclusions, in order to clarify the place of my argument in the wider context of current debates on the foundations of liberalism. I then provide a new, detailed account of the idea of *modus vivendi*, and of how it can be deployed to construe to what I will call a ‘realistic’ conception of public justification (section II). I argue that, however, *modus vivendi* is unable to satisfy the desiderata of a liberal theory of public justification, and hence is not a promising option for liberal consensus theorists (section III). These points lead to some more general remarks about the significance of my critique of the *modus vivendi*-based account of liberal consensus for the overall prospects of the consensus view of liberal legitimacy (section IV).

I. LIBERAL REALISM, AGONISM, AND MODUS VIVENDI

The consensus view of liberal legitimacy is familiar, but a brief characterization of it will be necessary to set the stage. The paradigmatic advocate of the view is, of course, John Rawls. In *Political Liberalism* he writes:

Our exercise of political power is *fully proper* only when it is exercised in accordance with a constitution the essentials of which all citizens *as free and equal* may reasonably be expected to *endorse* in the light of principles and ideas acceptable to their common human reason. This is the liberal principle of legitimacy. (1993, 137, emphasis added)

The basic idea is that a well-grounded political framework need not just embody liberal values; for political power to be properly exercised we also need a freely developed consensus. That consensus is established through the ideas of public reason and public justification:² publicly justified principles are acceptable to reasonable citizens (on an adequate characterization of reasonableness), and thus can be presented as enjoying the free hypothetical consent of the citizenry, which in turns confers legitimacy upon liberal political arrangements.³ That is the sort of (hypothetical) consent relation I discuss here.

2] In line with much of the literature, I will use “public reason” and “public justification” interchangeably (unless otherwise specified); to be more precise one may say that the use of public reason is a necessary condition for achieving public justification. The idea of public reason is, of course, rooted in a Kantian approach to political theory. This point has been insightfully articulated by Onora O’Neill (1986).

3] Other prominent proponents of a similar approach are Gerald Gaus (1996), Charles Larmore

Finally, by “free consent” – as opposed to consent simpliciter – I understand a form of consent that is based on the exercise of the consenting individuals’ personal autonomy. “Personal autonomy” should be understood as an umbrella term here: rather than in its original Kantian sense,⁴ I use the term as a placeholder for all the typical foundational commitments of contemporary liberalism (a conception of persons as free and equal, a notion of human flourishing, etc.). Those commitments explain how a consensus can carry normative force, and are thus used by different theorists to motivate their adherence to the consensus view of liberal legitimacy.

In order to understand the relevance of the problem at stake in this paper within the contemporary debates on the foundations of liberalism we may start from the idea that the liberal political framework is best grounded simply as a morally neutral medium for the adjudication of disputes between competing interests and conceptions of the good – a thought that seems to lie at the heart of many critiques of Rawlsian approaches to liberal legitimacy. In fact many theorists defend *modus vivendi*-like accounts of the foundations of political authority as a reaction to what they regard as an excessive distance between Rawls’ heavily moralised prescriptions for the conduct of political deliberation and the actual political practice of most contemporary liberal democracies.⁵ These complaints register a certain dissatisfaction with the restrictions Rawls places on the deliberative process in order to safeguard the liberal normative commitments of freedom, equality and autonomy; in other words, they accuse Rawls of ‘rigging’ the deliberative process to ensure a liberal-friendly outcome. This worry has been poignantly formulated by Thomas Nagel:

Part of the problem is that liberals ask of everyone a certain restraint in calling for the use of state power to further specific, controversial moral or religious conceptions – but the results of that restraint appear with suspicious frequency to favor precisely the controversial moral conceptions that liberals usually hold. (1987, 216)

That is what one may call the ‘agonistic’ or ‘radical democratic’ critique of political liberalism and other similar accounts of deliberative democracy.⁶ For instance, while remaining committed to a view that grounds liberal democratic legitimacy in the participation and allegiance of the governed, Chantal Mouffe writes:

If both Rawls and Habermas, albeit in different ways, aim at reaching a form of rational consensus instead of a simple *modus-vivendi* or a mere agreement, it is because they

(1996, 121-151), Stephen Macedo (1991), and Lawrence Solum (1993).

4] The other important source for the value of autonomy in the liberal tradition is, of course, J.S. Mill, though he tends to use terms such as “individuality” and “spontaneity”.

5] For example, defenders of ‘procedural’ (liberal) democracy often use similar arguments in their attacks on ‘deliberative’ democrats such as Rawls. See Cohen (1996).

6] See, for instance, Mouffe 2000.

believe that, by procuring stable grounds for liberal democracy, such a consensus will contribute to securing the future of liberal democratic institutions. (2000, 9)

Yet, she argues, they are mistaken because securing allegiance to a liberal democratic polity “is not a matter of *rational justification* but of availability of democratic forms of individuality and subjectivity. By privileging rationality, both the aggregative and the deliberative perspective leave aside a central element, which is the crucial role, played by passions and emotions in securing allegiance to democratic values.” (2000, 10) However my argument will show how Mouffe’s agonistic position misses the point that the moral and epistemic restrictions placed by Rawlsian political liberals on the deliberative process do not just aim “to establish a close link between liberal values and democracy”, but they also – and more importantly – aim to safeguard the very values that motivated the adherence to a consensus or allegiance-based account of legitimacy. So in a way Mouffe is right in criticising Rawls (and Habermas) because “their move consists in reformulating the democratic principle of popular sovereignty in such a way as to eliminate the dangers it could pose to liberal values” (2000, 3). That (serious) problem may be solved by adopting a looser, *modus vivendi*-like approach to public deliberation (as Mouffe in fact recommends); yet, as we will see, doing so would also jeopardize the voluntarism that underpins any consensus-based account of the legitimacy of a liberal polity – hardly an overall improvement.

Similar lines of argument can be found in the debates about the role of religious arguments in political deliberation:⁷ critics of Rawls often argue that democracy cannot fully deploy its normative efficacy if we place (normative) constraints⁸ on political activity to the extent that some individuals and groups will be prevented from deliberating about certain policy matters by referring only to values which, according to their conception of the good, are directly and sometimes pre-eminently relevant to an assessment of the issues at stake. Summarising and elaborating on some such views, Philip Quinn writes:

So I am skeptical about there being any assured real (as opposed to merely possible) costs associated with being guided by Perry’s inclusivist ideal rather than the Rawlsian ideal of public reason. And if there are none, the inclusivist ideal is more attractive than its rival because, being less restrictive, it allows all citizens to express themselves and their deepest values more fully in the political sphere [...] Adams suggests that “Rawls underemphasizes the combative aspects of a democratic polity and tends to overestimate the level of theoretical agreement in political ethics needed for an attainably just society.” I concur. Of course we let ourselves in for something more like debate than like dialogue on many issues if we adopt the inclusivist ideal, but I consider that no bad thing when there is disagreement in a pluralistic democracy. (1995, 35)⁹

7] For an extensive introduction to these debates, see Audi 2000.

8] I use the term “normative constraints” to indicate that Rawls does not support political constraints such as restrictions of free speech, even though he does say that unreasonable citizens should be “contained” (1993, 64n; cf. Quong 2004).

9] Also see Perry (1993).

Quinn and Perry's claims clearly resonate with the realist idea that we should keep normative theory closer to actual politics.¹⁰ And it seems plausible to interpret those claims as relying at least in part on the normative force of a process of deliberation that allows for the unconstrained will of the citizenry to be instantiated in the political framework – a way of claiming that the liberal political framework is legitimate because it is the object of a consensus.¹¹ Those liberals may or may not explicitly or unambiguously defend a *modus vivendi* account of consensus-based legitimacy. But I maintain that some of their criticisms of political liberalism commit them to such a view: in sections 2 and 3 below I will argue that there is no middle ground between *modus vivendi* and the Rawlsian view, as any significant attempt to lighten the moral restrictions Rawls places on the hypothetical deliberation process turns the consensus into a *modus vivendi*, which defeats the purpose of invoking a consensus in the first place.¹²

More generally, the appeal of *modus vivendi* as an alternative to mainstream consensus-based liberal legitimacy resonates with 'realist' political theorists such as John Gray, John Horton, Glen Newey, and Bernard Williams, many of whom explicitly embrace the idea of *modus vivendi*, and all of whom lament the conceptually unsustainable precedence afforded to morality over power dynamics by contemporary liberal legitimacy theory – the foremost example of this problem being, of course, Rawlsian political liberalism.¹³

This points to a feature of real political stability which, I believe, has been seriously neglected by Rawls and other deliberative democrats who seem to think that a law or policy will necessarily be more acceptable to its opponents if it is the outcome of process of political deliberation conducted in accordance with public reason. [However] In some circumstances it is more important that the outcome be seen on all sides as a rough and ready compromise in which all the parties have been given something and each has made concessions. Rawls regards this kind of process as 'political in the wrong way'. [...] However, a *modus vivendi* need not be an arrangement entirely devoid of a moral dimension. It does not have to be understood, as Rawls presents it, exclusively as an unstable balance of forces (Rawls, 1996: 432-3); rather it is a mix of morality and power. (Horton 2003, 22)

10] Cf. Newey (2001, 22-30).

11] Of course not all arguments in favour of the appeal to religion in political deliberation rest on the consensus view of legitimacy. For example, for David Hollenbach the use of religious arguments in the public sphere should not be curtailed because doing so would jeopardize a nation's "civic unity", and, ultimately, "the common good" (1993, 890).

12] Though it is of course possible to argue for fewer restrictions than Rawls envisages on the grounds that the use of religious reasons does not constitute a violation of fellow citizens' autonomy: see Eberle 2002, Gaus and Vallier 2009. But those arguments do not concern us here, as they accept Rawls' moralized conception of consensus and just disagree on the interpretation of some of the values underpinning it.

13] For an overview of the realist stream in recent political theory (including a discussion of these theorists' contribution to it) see Galston 2010. Also see Gray 2000, Horton 2003, Newey 2001, and Williams 2005.

So, by analyzing how the idea of *modus vivendi* can be deployed (as an alternative to Rawls' "overlapping consensus") to correct what is often considered a weakness of the legalism and moralism of political liberalism, I will provide a new angle from which to systematize and make sense of those critiques.

Of course, some proponents of *modus vivendi* will not be interested in deploying that idea to ground liberalism; rather, they will see it as an *external* corrective to the liberal moralism.¹⁴ In other words, they will ground political authority by replacing liberal legitimacy theory with an account of *modus vivendi*, perhaps because they maintain that prudence should take precedence over morality when assessing the normative status of political authority. My argument does not directly engage with that position – nonetheless it clarifies the differences between those two takes on the idea of *modus vivendi*.

II. MODUS VIVENDI AND PUBLIC JUSTIFICATION

The idea of *modus vivendi* has come to the fore through Rawls' discussion of it in *Political Liberalism*. The rough characterization of *modus vivendi* I offered at the outset of the paper is roughly equivalent to Rawls' widely known definition of *modus vivendi*: "A consensus on accepting certain authorities, or on complying with certain institutional arrangements, founded on a convergence of self- or group interests." (1993, 147) Rawls contrasts prudentially-motivated *modus vivendi* with the idea of an overlapping consensus, which is an agreement underpinned by moral (primarily, but also by epistemic) reasons that are in line with an appropriately specified conception of liberal citizenship. To spell out this point, I propose a more precise characterization of *modus vivendi* as an agreement establishing a stable political framework in which at least one – but typically most – of the parties (i.e. a sizeable minority) participate for non-moral and non-epistemic reasons, i.e. (typically) prudential reasons, or for reasons that are not compatible with a liberal conception of citizenship.¹⁵

That characterization of *modus vivendi* could strike some as rather different from what is perhaps the most articulate such characterization in the literature, namely the one recently proposed by Gerald Gaus:

Agreement X is a *modus vivendi* between agents A and B if and only if:

- 1) X promotes the interests, values, goals etc. of both A and B;
- 2) X gives neither A nor B everything they would like;
- 3) The distribution of the gains of the compromise (how close X is to A or B's maximum reasonable expectation) crucially depends on the relative power of A and B;
- 4) For both A and B, the continued conformity by each to X depends on its continued evaluation that X is the best deal it can get, or at least that the effort to get a better deal is not worth the costs. (2003, 59)

[14] As I will discuss in section 4, John Horton's latest work arguably falls within this camp.

[15] This type of motivation for accepting a settlement should not be confused with that of somebody who accepts a political framework for liberal moral reasons – respect for others as free and equal citizens, say – without morally endorsing its contents.

However, the differences lie more in the definiens than in the definiendum, as both definitions pick out roughly the same sort of agreements; the differences, then, are given by the fact that while Gaus focuses more on the goods and interests promoted by a *modus vivendi* agreement and on its being affected by the differences in bargaining power between the parties, I focus on the reasons the parties have to subscribe to a *modus vivendi*.¹⁶ The focus on motivation is important if, like Rawls, we are interested in the view that the agreement (rather than the values and interests it promotes) confers legitimacy.

To present the idea of *modus vivendi* less abstractly, we may simply think of a number of parties with different preferences: they reach a *modus vivendi* whenever they reach an agreement (leading to a stable political framework) without constraints on the sort of reasons that can motivate the parties to accept the terms of the agreement. The idea is that a *modus vivendi* maximizes the satisfaction of the preferences of each party, subject to the constraints given by the relative bargaining power of the parties. It is also important to note that the political framework should satisfy the condition of political stability (understood in the ordinary sense – a condition of social peace enabling reliably regulated social cooperation – rather than in the morally laden, Rawlsian sense of ‘stability for the right reasons’): if this constraint were not in place, on my definition just about any political arrangement could be characterized as a *modus vivendi* (thus making “*modus vivendi*” almost synonymous with “the course of history”), for any given situation, including one of anarchy, could be presented as the outcome of hypothetical bargaining between the parties involved in it.¹⁷

Now how exactly can *modus vivendi* connect to the idea of public justification? The term ‘consensus’ employed in Rawls’ definition suggests an answer to that question. Public justification is the liberal consensus theorist’s response to what one may call the challenge of persistent ethical diversity: under the conditions of freedom brought about by the political frameworks naturally favoured by consensus theorists, citizens develop an extensive array of diverse and diverging conceptions of the good. What is more, this diversity is persistent to the extent that it becomes virtually impossible to identify, or hope to identify in the foreseeable future, a justificatory account of the basic structure of society

16] An actual (yet not crucial) difference is that for Gaus all parties to a *modus vivendi* participate in the agreement for non-moral reasons, whereas I maintain that it is enough that at least one party do so. The thought is that if one agrees to a settlement for moral reasons but those reasons are not reciprocated, then we have something less than what would count as a Rawlsian overlapping consensus, or even as a constitutional consensus (because even the latter requires the parties to consider the institutions and principles agreed upon as good in themselves: Rawls 2003, 158ff. I return to the idea of a constitutional consensus in section 2.).

17] By identifying *modus vivendi* ex post the stability constraint also addresses Gaus’ point that *modus vivendi* cannot count as a publicly justified agreement because its persistence is subject to private judgments on the shifting balance of power (2003, 63-64). That is because a measure of stability (in the ordinary rather than the Rawlsian sense) is built into my definition of *modus vivendi*. At any rate consensus theorists’ claims about the superior stability of their envisaged consensus have been convincingly undermined by Sterling Lynch (2009).

that is (directly) acceptable to all – or even most – of the citizens’ private¹⁸ normative standpoints. The project of public justification, then, aims to overcome the challenge of diversity without renouncing the goal of consensus; it does that by offering an alternative standpoint, a public standpoint that enables us to reach some kind of consensus about what the basic structure of society should look like. The idea is to construe public reason in such a way that it can be shown that citizens have reason(s), from the standpoint of their private outlooks, to adopt a public outlook when deliberating about the basic structure of society. Of course, divergences arise among public reason theorists as to what kind of reasons there might be for adopting the public standpoint: roughly, on what I call an *idealistic* conception of public justification they will be mostly moral and epistemological reasons, whereas on a *realistic* conception they will be predominantly pragmatic and prudential – hence the connection with realist and agonist accounts of liberal democracy.

So *modus vivendi* can be understood as a consensus-based account of the foundations of political power if we construe a conception of public justification that produced the sorts of agreements described by the definition of *modus vivendi* offered above. I say that it can be understood as a version of the consensus view, rather than saying that it is a version of that view, because one can of course also defend a *modus vivendi*-grounded political framework on the basis of considerations other than the foundational role of the relation of consent between government and governed. The idea here would still be that the basic structure of society should be shaped by a consensus; the difference, however, is that this arrangement is not required in order to characterize a particular relation between government and governed, but rather because those political arrangements safeguard and promote certain goods (e.g. social peace, stability, human rights understood as interests, and so on).¹⁹ This may well be a promising line of argument, but exploring it would be beyond our scope here. What we need to show, on the other hand, is how a *modus vivendi* account of political legitimacy can be connected to a version of the realistic conception of public justification.

To answer that question, let us put the points we have just seen in Hobbesian terms (for, as shown by David Gauthier’s work on public reason, those terms are especially appropriate when presenting a realistic account of public reason²⁰): the use of private reason in deliberation about the political framework leads to social conflict (a point famously stressed by Hobbes: crudely, left to their own devices, people conflict), making

18] “Non-public” would be more a more accurate term, as “private” here does not refer to the private sphere of the early liberal theorists (most notably Constant): it is not the domain of life where the state has no right of interference. Rather, it is the cultural background that is exclusive to particular citizens or groups of citizens.

19] E.g. Haldane 1996, Neal 1993.

20] See Gauthier 1995 and Ridge 1998. For a historical overview see Iverson 1997. The controversy between Gauthier, Ridge, and Gaus revolves primarily around the issue of how exactly we can distinguish between public and private reasoning – an issue with no direct bearing on this paper’s main argument. Gauthier and Ridge do not discuss the implications of their views for liberal legitimacy theory; however both their positions could be subsumed under my general account of realist public reason.

it impossible to identify stable terms of political cooperation and peaceful coexistence. In societies with persistent ethical diversity the justification of the political framework will have to be a public one, i.e. one which is able to transcend the divergences of citizens' private conceptions of what values should inform the design of a political framework. Hence the Hobbesian spirit of the realistic conception of public reason: public reason is just a way of reasoning we adopt for the purposes of agreeing on how to live together in a political system. The content of the ideal of public reason, i.e. the rules for the adjudication of political controversies, does not really matter much, as long as it enables peaceful political coexistence. Anything goes, as long as it secures agreement and stability. In other words, on this strategy we are taking citizens as they really are, in the sense that we construe public reason on the basis of general normative commitments that are *actually* available to the citizens.

To better identify the defining traits of the realistic conception of public reason, contrast it with a more idealistic, Kantian, one. The idea here would be to take a standpoint that ought to be accepted by all. Private standpoints are not like that (because of our moral and epistemological idiosyncrasies), thus we should strive to find a public one. This is true both at the epistemological level (e.g. "use reasons that are intelligible to others/comply with epistemological standards that are acceptable to others") and at the moral level ("make on others only claims that can be justified to them"). Often the two levels are combined in a set of mixed epistemological and moral criteria.²¹ In contrast to the realistic strategy, the idealistic strategy takes citizens as they would be, were they committed to certain (moral and/or epistemological) values, i.e. it takes citizens as they should be: public reason is construed on the basis of normative commitments the citizens *should* have, regardless of whether they are actually available to them.²²

Of course, the realistic and the idealistic conceptions of public reason as I have just presented them are pure types; the actual conceptions found in the literature often try to include elements of each type. But what this taxonomy shows is how each pure type embodies one crucial desideratum for the project of establishing free hypothetical consent through public justification: the realistic conception embodies the pragmatic desideratum, whereas the idealistic conception embodies the moralistic desideratum. The *pragmatic* desideratum requires that hypothetical consent to a publicly justified set

21] In Fred D'Agostino and Gerald Gaus' language, this is "the epistemological-moral view" of public reason (1998, xiii).

22] Note that those reasons are not primarily the reasons that may be deployed when debating the design of the political framework. Rather, they are the reasons a citizen may appeal to when deciding to accept an agreement about the design of the political framework. As a further point one may note that, if we do not place any restrictions on the reasons one may have for accepting an agreement, one may very well also not restrict the types of reasons and arguments that may be deployed when deliberating about the political framework. The connection between these two levels of unrestrictedness of available reasons is not a logical one, nonetheless it appears to be supported by strong pragmatic reasons: if we want the full bargaining power of the parties to influence the deliberative process, we have reason to allow them to defend their positions in the way they deem to be more effective.

of principles regulating the political framework be a concrete possibility – something feasible given the citizenry’s actual motives, beliefs, and desires. On the other hand, the *moralistic* desideratum stresses that, for consent to retain its legitimating force, it should be given without violations of the personal autonomy of the consenting individuals.

That shows how the idea of *modus vivendi* connects with the realistic conception of public reason (and how it cannot connect with the moralistic conception). The (Hobbesian) idea here is to construct public reason in such a way that it requires us to transcend or bracket elements in our private system of values, in our private reason, to the extent where we can find enough common ground (probably through our ‘self- or group interests’) to come to a settlement guaranteeing peaceful political coexistence. The medium – indeed any medium – that enables the citizens to take the standpoint leading to such a consensus is indeed the Hobbesian public reason. There are no restraints on this process of transition from private to public reason. The consensus is still hypothetical (actual consent is chimerical, as most contemporary consensus theorists recognize), but it is just, as it were, one layer of hypotheticalness, in the sense that the consenting individuals are not the actual citizens but their counterparts (because the actual citizens will typically not be directly asked for their express agreement), yet they are not idealized to the point where the sources of their normative commitments may be fundamentally different from those of the actual citizens. So, if a *modus vivendi* over liberal institutions can be found, hardly anybody will be excluded from the consensus.

These considerations also show in what sense we can say that constructing public reason through a *modus vivendi* is a way of responding to the pragmatic desideratum of theories of public reason: the thought is that a *modus vivendi*-based agreement will ensure that the hypothetical consent secured by the agreement will be the consent of the actual members of the society, not of their epistemically and/or morally idealized counterparts. However, the achievement of meeting the pragmatic desideratum comes at a rather high cost, as I shall argue in the next section.

III. AGREEMENT, AUTONOMY AND (LIBERAL) LEGITIMACY

In this section I will discuss what I regard as a serious problem affecting the *modus vivendi*-based approach to public reason. As we have seen, the approach obviously has advantages, and indeed not just advantages: it addresses a vital concern of the consensus theory of liberal legitimacy, namely the need for pragmatism. However, in a nutshell, the serious problem is that the sort of consensus reached through a *modus vivendi* is incompatible with a crucial desideratum of the consensus view of legitimacy, i.e. grounding political power in a way which is respectful of personal autonomy. On the *modus vivendi* approach to public justification, all that counts is the fact of stable agreement to a set of rules regulating the political framework. But then any concern for how that agreement came about (at gunpoint, under the effect of propaganda, and analogous situations) becomes secondary, if not entirely irrelevant. Now that, of course, is a problem because

liberal consensus theorists are – with good reason – interested in *free* consent (i.e. consent which preserves the consenting individual's personal autonomy), not just any kind of consent.²³

But the problem is that the *modus vivendi* strategy is introduced precisely because, under conditions of persistent ethical diversity, there is no agreement of that kind. That is to say, it is introduced in order to address the pragmatic desideratum. The *modus vivendi* strategy tries to produce that agreement by relaxing the standards of what is legitimate, i.e. what counts as freely consented to, but that does not seem a good move for someone committed to the consensus view of liberal legitimacy. As we have seen, the appeal of the consensus view lies in the fact that it allows the grounding of political power in a way which is respectful of personal autonomy (hence the natural link with liberalism) – an idea which has been well expressed by Jeremy Waldron: “If the rule is one that the citizen has agreed to, surely little that is important in relation to liberty is lost if it is subsequently enforced against him.” (1987, 133) But surely if it is not guaranteed that consent results from of a free (albeit hypothetical) choice, its appeal for liberal consensus theorists quickly melts away.

So it seems that, while the combination of *modus vivendi* and the realistic approach to public reason could *prima facie* appeal to those realist or agonistic consensus theorists who want to “take people as they are”, it is the very pragmatism of this approach which, in a context of persistent ethical diversity such as that characterizing modern liberal societies, condemns it to sanctioning outcomes or procedures those same theorists cannot consider up to the standards of liberal legitimacy.

To further clarify that point it is worth noting how, in my reading of *Political Liberalism*, Rawls' rejection of *modus vivendi* is motivated by considerations similar to the ones I offer here: *modus vivendi* does not guarantee a legitimate political framework because it may sanction agreements that violate citizens' personal autonomy. This interpretation may appear somewhat controversial. In fact, many commentators²⁴ have thought that Rawls' argument against *modus vivendi* is a pragmatic one (something along the lines of “*modus vivendi* is the product of a contingent balance of power between competing parties, therefore it is bound to collapse sooner or later”); but in my view it is not – there might be a minor pragmatic side to the argument, but surely it is not its crux. Rather, it is a moral argument, because Rawls uses “stability” in a moral sense (he talks of “stability for

23] The point here is not the truism that unrestricted bargaining does not necessarily yield liberalism; rather, the point is that even when it does yield liberalism it yields a liberalism that cannot be considered legitimate qua object of a consensus. But that argument does not amount to a defense of mainstream consensus theories of liberal legitimacy: as I will argue, the issues identified by realists and agonists are very pressing, and if they cannot be met then we should seriously question the viability of consensus-based accounts of liberal legitimacy.

24] See, for example, Scott Hershowitz (2000, 222): “Rawls's reason for requiring stability for the right reasons as opposed to accepting a *modus vivendi* rests on his belief that a *modus vivendi* cannot provide enduring stability”.

the right reasons”, i.e. a situation in which citizens are motivated to comply with the norms regulating the political framework by appropriate considerations of political morality that do not infringe on their “status as free and equal citizens” – or, in my terminology, on their personal autonomy). Rawls, as a liberal consensus theorist, is not interested in mere agreement. He is interested in an agreement between free and equal citizens, which cannot take place at gunpoint or in any circumstances curtailing the autonomy of the consenting parties. So, if my reading of *Political Liberalism* and my critique of modus vivendi-based accounts of liberal legitimacy by consensus are correct, it follows that Rawls is right in maintaining that nothing short of an overlapping consensus is needed to meet the desiderata of the consensus view of liberal legitimacy. But that should by no means be taken as a defence of Rawlsian liberal legitimacy theory against other, broader uses of the idea of modus vivendi. The point here is that those who wish to embrace modus vivendi to reject Rawls’ moralism and legalism should also abandon the measure of (hypothetical) voluntarism that is constitutive of the consensus view of liberal legitimacy.

As anticipated in the introductory section, in a number of debates within liberal-democratic theory the claim is often advanced that Rawls’ account of what counts as a free consensus is needlessly restrictive and too morally laden. To have a closer look at a good example of this kind of position, and one of the most explicitly worked out ones, let us consider Claudia Mills’ critique of Rawlsian political liberalism. Mills is explicitly committed to both liberalism and to the view that the source of political legitimacy is an agreement or a consensus between those subjected to the exercise of political power. Yet she also takes issue with the sort of moral demands Rawls places on what counts as a free consensus (i.e. as the proper form of endorsement of the principles characterizing the political framework):

I argue that if we look at what Rawls wants for liberalism compared to what he thinks we get from a modus vivendi, we will find that he can get what he wants more easily than he thinks. In fact, Rawls himself provides a persuasive story for how the kind of endorsement he wants for liberalism can grow out of a modus vivendi, without any invocation of an overlapping consensus. Where he goes wrong, I argue, is in overestimating the importance to stability of a shared allegiance to principles and in underestimating the importance of a shared history of living together. (2000, 192)

One of the problems here has to do with our earlier discussion of ‘stability for the right reasons’: like many other critics of political liberalism, Mills does not fully acknowledge the moral dimension of the Rawlsian notion of stability. It is certainly possible that principles that make peaceful coexistence possible will in a sense be endorsed by the citizenry (“Our endorsement of the rules was based first and foremost on the pragmatic consideration that *they worked*. [...] We then value the principles in large part because they make it possible for us to live together” (2000, 201-2)), yet it is not the sort of endorsement that should be considered appealing by those committed to the consensus view. To be fair, Mills is aware of the fact that her position abandons the voluntarism that characterises the consensus view:

It may be that Rawls downplays allegiance to history, culture, and place and lays stress instead on shared allegiance to principles because he believes that the latter can be *voluntary* in a way that the former cannot and that social contract theories as a group seek to establish the way in which our consent to governmental authority is *free*. (2000, 202)²⁵

Nonetheless she maintains that stability can make up for that loss: “But while history, culture and place do not fit well with the voluntarism typical of social contract theorists, they do serve well to establish the kind of stability Rawls claims to be seeking.” (2000, 203) Except that, as we have seen, that is not the stability Rawls seeks. More importantly, Rawls has good cause to seek “stability for the right reasons”, for if the agreement tasked with grounding legitimacy is not a free one why would one think that it had any (or enough) normative force? If liberals commit to the view that citizens’ consent is the source of legitimacy it must be because they think that consent can be based on the exercise of the citizens’ freedom and autonomy – hence the inescapable need for the sort of restrictions Rawls places on what counts as consent. And, as we shall see below, invoking the substantive virtues of stability cannot make up for a lack of voluntarism within a consensus-based account of legitimacy.

At this point it is worth considering some lines of reply to the argument I advanced so far. Could a *modus vivendi* consensus theorist not reply that all her theory needs in order to become immune to my criticism is a simple and innocuous restriction on the deliberative processes sanctioned by *modus vivendi*, such as a rule prohibiting the use of coercion? Surely, she may argue, one does not need thick and controversial moral notions in order to have a deliberative process respectful of the citizens’ personal autonomy. However, it is my contention that this line of reply is not satisfactory because, if we consider carefully enough what is needed in order to cash out such a restriction in a way that will prove strong enough in order to safeguard the citizens’ personal autonomy now and in the foreseeable future (as required by the moralistic desideratum of the consensus-based view of liberal legitimacy), we will come to realize that we need a set of normative commitments of comparable weight (i.e. moral ‘thickness’) to those embodied – for example – by Rawls’ notion of “reasonableness”. Rawls does not explicitly make this claim or provide an argument for it, but it is possible to supply one by noting that is not enough, for the moralistic desideratum to be satisfied, that free consent be *possible* – it has to be *guaranteed*. It may very well be the case that in some societies, as a matter of fact, conditions are such that a free consent-friendly *modus vivendi* is possible for the time being. But on the consensus view of liberal legitimacy the citizens’ personal autonomy cannot be left hostage to the circumstances (e.g. a critical increase in the popularity of intolerant ethical outlooks, and the like). A good theory of liberal legitimacy needs strong constraints in order to make sure not only that our deliberation procedure guarantees free consent given

25] Rawls does offer a *causal* story of how an overlapping consensus may arise from a constitutional consensus (1993, 158ff), but that story is irrelevant to the *normative* status of the overlapping consensus.

the present level of ethical diversity, but also that it will continue to do so in the foreseeable future. Successful normative legitimation should be immune from contingent shifts in political leverage. The idea is that if we look at a liberal, freely consented to, regime *now* and conclude that it is legitimate because it is a *modus vivendi*, we will be at a loss of arguments to denounce it as illegitimate if at some point *in the future* changes in the equilibrium of power or the level of diversity within it (say) will yield changes in the design of the political framework to the extent that the basic structure of that society will stop enjoying the free consent of the citizenry. But if we do have to go down this path of thick procedural constraints, surely the inclusiveness of *modus vivendi* (and hence its ability to satisfy the pragmatic desideratum) will be lost.

To reinforce that point, recall how Rawls maintains that only the consent of reasonable citizens is needed in order to secure legitimacy. In this way he restricts the deliberative process, ensuring that it will be conducted in a way that is respectful of the citizens' personal autonomy, for reasonable citizens are indeed "persons engaged in social cooperation among equals", and they "desire for its own sake a social world in which they, as free and equal, can cooperate with others on terms all can accept" (1993, 48-50). In my view, this sort of restriction makes it rather unlikely that political liberalism will satisfy any sensible formulation of the pragmatic desideratum of the consensus theory of liberal legitimacy, for it is far from clear that the boundaries of reasonableness are not set arbitrarily – but that is not our concern here.²⁶ We should instead note that Rawls' restrictions on the deliberative process are deliberately engineered to be as 'thin' as is compatible with ensuring the safeguard of citizens' personal autonomy (hence, in short, the well-known distinction between moral and political values, and the scope restrictions on the bindingness of his prescriptions, which only apply to public discourse on the basic structure). That shows that, as I have been arguing, should the *modus vivendi* consensus theorist try to ensure the autonomy-friendliness of the deliberative process, she would have to put in place rather severe restrictions; and as a result of those restrictions the *modus vivendi* would indeed morph into an arrangement not very different from Rawls' overlapping consensus.

Here one might object that *modus vivendi* could at least be seen from a historical point of view as an instrument for eventually bringing about the sort of consensus envisaged by the liberal legitimacy view.²⁷ If we cannot have the overlapping consensus

26] I explore this point in Rossi 2008. In a nutshell, Rawls presents the challenge of liberal legitimacy by asking "How is it possible that there may exist over time a stable and just society of free and equal citizens profoundly divided by reasonable though incompatible religious, philosophical, and moral doctrines?" (1993, xx) He replies that it is possible, as we only need the consensus of reasonable citizens, who in turn are characterized as committed to seeking fair terms of cooperation between free and equal citizens, i.e. as liberals (1993, 48-54). But grounding the legitimacy of a liberal consensus in the normative force of reciprocity and other liberal values makes the consensus redundant. Yet the voluntarism of the consensus was supposed to remove the need for a substantive defense of those liberal normative commitments, which now appear groundless or arbitrary.

27] This point could be presented as a modified version of Mills' argument from "shared history"

right now, the objection goes, let us support a *modus vivendi*, as it will eventually lead to the emergence of a liberal overlapping consensus between free and equal citizens.²⁸ In other words, *modus vivendi* may not be legitimate as such, but it is the path to legitimacy. In light of our analysis we could respond in at least two ways. First, this objection is somewhat off the mark: we are after all considering simply whether *modus vivendi* arrangements could count as legitimate according to the consensus view of liberal legitimacy, not how legitimate political arrangements might arise. Second, and more importantly, it is far from clear whether any *modus vivendi* is likely to evolve into a form of hypothetical agreement of the sort sanctioned by the consensus view of liberal legitimacy. With regard to this, Richard H. Dees convincingly showed that “the story that Rawls tells about the emergence of overlapping consensus from *modus vivendi*] is sketchy, but its unspoken optimism belies the deep problems that such a transformation involves.” More specifically, Dees argues that for the transformation to take place the parties to the conflict need to come to regard toleration as a value *per se*; historical examples, however, show how that is by no means guaranteed to happen, or indeed even likely: “whether toleration can be justified in a way that the parties to such deep conflicts can accept will depend crucially on contextual features” (1999, 667-68),²⁹ which are difficult to pick out. Thus, if consensus theorists want to defend *modus vivendi* as the path to legitimacy, they cannot do it abstractly: they face the arduous task of providing a case-by-case account of how, in a given context, it is likely that the *modus vivendi* will evolve into a consensus between free and equal citizens.

I conclude this section by considering another line of reply that could tempt those wishing to combine *modus vivendi* and the consensus view of liberal legitimacy. At least, they may argue, securing a political settlement produces stability, social peace, and so on.³⁰ What is more, they may add, at least in some circumstances these values are to be prioritized. This claim may be read in two ways: on the one hand one may argue that peace and stability are goals that actual citizens desire (simply because peace is necessary for the pursuit of most other goals one may have), thus securing them through a *modus vivendi* amounts to obtaining a consensus. But that argument is still open to the earlier critique that it does not safeguard the foundational commitments (such as autonomy) that underpin the recourse to the consensus view. On the other hand the argument could rest on the intrinsic appeal of peace and stability. However that would be a departure from the concerns of the consensus view, for it grounds political power entirely on the value of stability, leaving no role to play for consent. The same would be true in the case of appeal to

we discussed above.

28] This is envisaged by Rawls in his discussion of the transition from a constitutional consensus to an overlapping consensus (1993, 158-68).

29] So Dees shows that the crucial – and very difficult – step is the one from *modus vivendi* to constitutional consensus (i.e. when toleration starts to be seen as good in itself). That is why I do not discuss constitutional consensus here: once we secure it, we can agree with Rawls that we are on the path to an overlapping consensus (1993, 164ff.).

30] For this line of argument see Dauenhauer 2000, Haldane 1996, Horton 2006, Neal 1993.

the argument that the possibility of an agreement is likely to track the appropriateness of certain set of rules in a given context, and so on. On those views, then, *modus vivendi* would ground the exercise of political power through substantive considerations of justification rather than relational considerations of consent: political power would be grounded solely in the fact that the political framework possesses certain valuable features (stability and the like), rather than also in a hypothetical relation of consent obtaining between the government and the governed. Such a substantive justification-based approach might very well be worthwhile in its own right, but it does not help the cause of the consensus view of liberal legitimacy.³¹

IV. CONSENSUS, REALISM, AND LIBERAL LEGITIMACY

The analysis of the idea of *modus vivendi* I carried out here focuses on its prospects as the core of an account of how a political framework could satisfy the desiderata of the consensus view of liberal legitimacy, which is dominant in contemporary liberal theory. This way of looking at *modus vivendi* may perplex some: after all, many proponents of *modus vivendi* as an account of the normative foundations of liberalism do not – at least explicitly or intentionally – present their view as a consensus-based account of legitimacy. However, as I have shown, the language usually adopted by these theorists often does imply, or at least allude to, a commitment to the consensus view. Conversely, some realist or agonistic versions of the consensus view do not explicitly propose a *modus vivendi*, but I maintain that at least some of those views can be subsumed under my account of realist public justification. Yet we have seen that, while a *modus vivendi*-based theory of legitimacy certainly does satisfy the pragmatic desideratum of the consensus view, it does so at the cost of jeopardizing the moralistic desideratum: given a level of persistent ethical diversity such as that characterizing most contemporary liberal societies, we cannot have a genuinely inclusive hypothetical consensus on the political framework while at the same time ensuring that everybody's consent will be free in any normatively salient sense. And that is why liberal consensus theorists – or perhaps all consensus theorists – should not rest any hopes on the idea of *modus vivendi*.

But there is a broader and more important issue – which however can only be briefly canvassed here – that these considerations draw attention to. Even though *modus vivendi* cannot deliver what consensus theorists need, the issue it tries to address (i.e., to put it crudely, the need to achieve a broad hypothetical consensus grounded in reasons actually available to the citizenry) is a genuine concern for the consensus view of liberal legitimacy (embodied, in fact, by the pragmatic desideratum, which echoes the concerns of agonistic and realist critics of Rawlsian liberalism). And it is far from clear that it is

31] So, to return to the earlier discussion of Mills' position, if one is interested in grounding liberal legitimacy in a form of consensus adducing such substantive considerations to make up for a lack of voluntarism will not do.

possible to address this issue adequately without falling short of the equally important goal of ensuring that the consensus will be respectful of the citizens' personal autonomy. Perhaps viability concerns such as those briefly hinted at here have recently led a theorist like John Horton (2009) to retain the ideas of *modus vivendi* and of a consensus-based account of legitimacy at the expense of the commitment to liberalism (which, on his view, would only be legitimate subject to favourable background conditions). Exploring this interesting new view would take us beyond the scope of this paper; yet the analysis offered here would caution against grounding the exercise of political power in a sort of agreement if one is not also prepared to ensure that the agreement is a freely and autonomously undertaken one. As seen above, Waldron pointed out that voluntarism can preserve freedom despite the exercise of political power; one does not need to invoke pragmatic contradiction arguments to also see that voluntarism does not in and by itself ground the exercise of political power unless it also safeguards freedom and autonomy.³² The problem, however, is that under conditions of pluralism the safeguard of autonomy has to take the form of restrictions (such as Rawls' criterion of reasonableness) on what should count as a normatively justificatory consensus – and those restrictions are, in turn, difficult to justify in light of the initial commitment to a measure of voluntarism.³³

These considerations suggest a working hypothesis for a critique of the consensus view of liberal legitimacy: have consensus theorists set themselves an impossible task, given the persistent ethical diversity that characterizes liberal polities? If pluralism creates an irreconcilable drift between the moralistic and the pragmatic desideratum, then the prospects of the consensus view of legitimacy as a viable model for the construction of a political framework, rather than a mere regulative ideal, are rather bleak. A related and somewhat less pessimistic line of inquiry may simply seek to establish what – if any – are the empirical conditions under which the consensus view will be feasible. However, on the consensus theorists' own account of the connection between liberal institutions and persistent diversity, those conditions seem unlikely to obtain in modern liberal democracies. Thus the critical hypothesis I just sketched may be supplemented with the observation that the consensus view may owe its deficiencies to its historical roots, in the sense that it is only designed to accommodate the relatively low level of diversity found in early modern European societies.

The demise of the consensus view of liberal legitimacy, however, need not coincide with the demise of liberalism. To sketch an even broader research agenda, I would suggest that if modern persistent pluralism makes liberal moral commitments incompatible with the voluntarism of the consensus view, then perhaps liberals would be better off abandoning the idea that legitimacy requires a somewhat voluntaristic consensus. The

32] One may say, though, that voluntarism grounds power in so far as it is instrumental to effective social cooperation, or peaceful coexistence, and the like (and that might be another way of reading of Horton's position). That is a plausible view, but it is a departure from the consensus view of legitimacy, as the normative work would be done by the substantive values of social cooperation and peace.

33] This point could be considered an instance of "the paradox of positive liberty". See Carter 2008.

liberal tradition features many teleological exponents,³⁴ and thus clearly not univocal as to the need of such a consensus. Securing a free consensus on liberal values requires very demanding procedural restrictions. Indeed, such a consensus-based approach is not merely problematic in its own right; it is also detrimental to liberalism, in so far as it strips liberal values of their justificatory force by turning them into seemingly arbitrary constraints on a consensus-based legitimation process, while neglecting the crucial task of making a direct, substantive case for those normative commitments. In fact, the consensus' view attitude to substantive justification of liberal political practice is more than just neglect – it is a proscription, insofar as any attempt at explicit justification of the normative commitments informing the procedural restrictions would expose their partiality and controversiality, which the consensus view is bound to deny. Thus a shift away from the consensus view would arguably reinforce our ability to make the case for liberalism.³⁵

enzo.rossi@newport.ac.uk

REFERENCES

- Audi, Robert. 2000. *Religious Commitment and Secular Reason*. Cambridge: Cambridge University Press.
- Carter, Ian. 2008. Positive and Negative Liberty. *Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/liberty-positive-negative/>. (accessed March 1, 2009).
- Cohen, Joshua. 1996. Procedure and Substance in Deliberative Democracy. In *Democracy and Difference*, ed. Seila Benhabib, 95-119. Princeton: Princeton University Press.
- D'Agostino, Fred, and Gerald F. Gaus, eds. 1998. *Public Reason*. Aldershot: Ashgate.
- Dauenhauer, Bernard P. 2000. A Good Word for a Modus Vivendi. In *The Idea of a Political Liberalism*, eds. Victoria Davion and Clark Wolf, 204-20. Lanham, MD: Rowman & Littlefield.
- Dees, Richard H. 1999. Establishing Toleration. *Political Theory* 27: 667–93.
- Eberle, Christopher J. 2002. *Religious Conviction in Liberal Politics*. Cambridge: Cambridge University Press.
- Galston, W. 2010. Realism in Political Theory. *European Journal of Political Theory* 9: 385-411.
- Gaus, Gerald F. 1996. *Justificatory Liberalism*. Oxford: Oxford University Press.
- . 2003. *Contemporary Theories of Liberalism*. London: Sage.
- Gaus, Gerald F., and Kevin Vallier. 2009. The Roles of Religious Conviction in a Publicly Justified Polity: The Implications of Convergence, Asymmetry and Political Institutions. *Philosophy & Social Criticism* 35: 51-76.
- Gauthier, David. 1995. Public Reason. *Social Philosophy and Policy* 12: 19-42.
- Gray, John. 2000. *Two Faces of Liberalism*. Cambridge: Polity Press.

34] Some of the obvious names here would be Hume, Mill and Joseph Raz.

35] I wish to thank Sarah Broadie, Rowan Cruft, John Haldane, John Horton, Tim Mulgan, John Skorupski, Leif Wenar and audiences at the Universities of Manchester Metropolitan, Oxford, Swansea and St Andrews for helpful comments on various versions of this paper. My research has been supported by an AHRC Doctoral Award.

- Haldane, John. 1996. The Individual, the State, and the Common Good. *Social Philosophy and Policy* 13: 59-79.
- Hershovitz, Scott. 2000. A Mere Modus Vivendi? In *The Idea of a Political Liberalism: : Essays on Rawls*, eds. Victoria Davion and Clark Wolf, 221-30. Lanham, MD: Rowman & Littlefield.
- Hollenbach, David. 1993. Contexts of the Political Role of Religion: Civil Society and Culture. *San Diego Law Review* 30: 877-901.
- Horton, John. 2003. Rawls, Public Reason and the Limits of Liberal Justification. *Contemporary Political Theory* 2: 5-23.
- . 2006. John Gray and the Political Theory of Modus Vivendi. *Critical Review of International Social and Political Philosophy* 9(2): 155-69.
- . 2009. Towards a Political Theory of Modus Vivendi. Unpublished manuscript, Keele University.
- Iverson, Duncan. 1997. The Secret History of Public Reason: Hobbes to Rawls. *History of Political Thought* 18: 126-47.
- Larmore, Charles. 1996. *The Morals of Modernity*. Cambridge: Cambridge University Press.
- Lynch, Sterling. 2009. The Fact of Diversity and Reasonable Pluralism. *Journal of Moral Philosophy* 6: 70-93.
- Macedo, Stephen. 1991. *Liberal Virtues*. Oxford: Clarendon Press.
- Mills, Claudia. 2000. "Not A Mere Modus Vivendi": The Bases for Allegiance to the Just State. In *The Idea of a Political Liberalism: : Essays on Rawl*, ed. Victoria Davion and Clark Wolf, 190-203. Lanham, MD: Rowman & Littlefield.
- Mouffe, Chantal. 2000. Deliberative Democracy or Agonistic Pluralism? *Political Science Series* 72, Institute for Advanced Studies, Vienna.
- Nagel, Thomas. 1987. Moral Conflict and Political Legitimacy. *Philosophy and Public Affairs* 16: 215-40.
- Neal, Patrick. 1993. Vulgar Liberalism. *Political Theory* 21: 623-42.
- Newey, Glen. 2001. *After Politics*. London: Palgrave Macmillan.
- O'Neill, Onora. 1986 Kantian Politics: The Public Use of Reason. *Political Theory* 14: 523-51.
- Perry, Michael. 1993. Religious Morality and Political Choice. *San Diego Law Review* 30: 703-27.
- Quinn, Philip L. 1995. Political Liberalisms and Their Exclusions of the Religious. *Proceedings and Addresses of the American Philosophical Association* 69: 35-56.
- Quong, Jonathan. 2004. The Rights of Unreasonable Citizens. *Journal of Political Philosophy* 12: 314-35.
- Rawls, John. 1993. *Political Liberalism*. New York: Columbia University Press.
- Rossi, Enzo. 2008. *Liberal Legitimacy: A Study of the Normative Foundations of Liberalism*. PhD Thesis, Department of Philosophy, University of St Andrews.
- Ridge, Michael. 1998. Hobbesian Public Reason. *Ethics* 108: 538-68.
- Solum, Lawrence B. 1993. Constructing an ideal of public reason. *San Diego Law Review* 30: 729-957.
- Waldron, Jeremy. 1987. Theoretical Foundations of Liberalism. *The Philosophical Quarterly* : 127-50.
- Williams, Bernard. 2005. Realism and Moralism in Political Theory. In *In the Beginning Was the Deed*, 1-17. Princeton: Princeton University Press.

”Scales of Justice” and the Challenges of Global Governmentality

Ina Kerner
Humboldt University of Berlin

Abstract: This essay is a critical assessment of Nancy Fraser’s recent account of the “scales of justice” in a globalizing world. In particular, I examine the third dimension of justice introduced by Fraser, that of representation. In light of how civil society in many countries of the global South is affected by a form of power that we can call “global governmentality”, I argue that Fraser should not restrict her concern with problems of representation to issues of access to civil society, but also address problems arising from power mechanisms that currently shape and reshape it.

Key words: global justice theory, governmentality, development aid, NGOs, transnational civil society.

Methodological nationalism remains fairly widespread in the field of political theory. At first sight, there are even some good reasons for this. One might hold, for instance, that despite new forms of governance the nation state is still quite alive and that we should therefore continue theorizing it. One could also say that power often works locally, for example within the boundaries of a nation, and must therefore be assessed at this level as well. Upon closer examination, however, the problems of this ongoing trend become clear. For when employing methodological nationalism, transnational phenomena as well as their effects on the national sphere are hardly addressed. Furthermore – and this might relate to where most political theorists were trained and are based – the content of theoretical reasoning is often at least implicitly related to problems and conditions of OECD-countries. Given this, authors who methodologically transcend the frame of the nation state deserve positive attention, for their work promises to forgo the problems of methodological nationalism.

Among these authors is Nancy Fraser, who within the last years has undertaken the task of systematically globalizing her thought to transcend what she calls the “Keynesian-Westphalian frame” and “passé Westphalianism” (Fraser 2008b, 12, 71) not only in general, but also with regard to her own former work. Accordingly, *Reimagining Political Space in a Globalizing World* serves as the subtitle of her most recent collection of essays, *Scales of Justice*. And it is noteworthy that in this collection of essays Fraser does not only bring globalization into political theory. She also introduces the domain of the political as a crucial aspect of justice, thus expanding the two-dimensional redistribution-cum-recognition model of social justice that she had formerly proposed (Fraser 1995) by adding a third dimension, that of representation. So it seems that once we methodologically transcend the frame of the nation state, not only additional localities come into view but new issues appear on the theoretical agenda as well. But which kinds of global phenomena and which aspects of the political is Fraser addressing now? Is her new account sufficiently

broadened to address the old as well as the new problems that we are facing in today's globalized world?

In this essay, I will focus on one problem which I hold Fraser should and – considering the concerns that motivate her theorizing – should *want to* address, but omits: the power effects of relatively new forms of north-south-politics on civil society actors in the global south. Fraser introduces the third justice dimension, the dimension of representation that corresponds to the domain of the political, precisely because she is concerned with forms of political misrepresentation, which she considers as detrimental to justice. But as I will argue, when she addresses such forms of misrepresentation, she does not go far enough. For she is only concerned with matters of access to the sphere of representation – and does not also address power effects upon the form and the content of what is dealt with in this sphere, effects that, as I suggest, can also cause misrepresentation, even though of a different kind.

But there is also a second reason for discussing the aforementioned problem in conjunction with Fraser's globalized account of justice. If this problem is as severe as is sometimes argued (e.g. Edwards 2008), it should be addressed, in the realm of the political as well as in the realm of those strands of scholarship that attempt at helping to solve problems of injustice. Concerning the particular problem this essay is concerned with, Fraser's multi-faceted approach to justice, which explicitly deals with issues of representation, suggests itself as a suitable starting point for this endeavor. So in this sense, the following considerations are not exclusively considerations for theory's sake. They are also an effort to integrate an under-theorized but pressing problem into the debate on global justice.

In what follows, I will proceed in four parts: First, I will briefly point out how, in her most recent publications, Fraser has globalized her account of justice and thereby introduced the domain of the political into her theorizing. In the second part I will address what Fraser omits, which is the proliferation of conditioned development aid offered to NGOs in the global south. I will suggest interpreting this as a form of global governmentality with productive power effects on the political agendas of civil society actors. In the third part, I will briefly consider how Fraser herself addresses what she calls "globalized governmentality", how her conception relates to what I call global governmentality, why her account of justice remains insufficiently Foucaultian, and how her position should be pushed further in this direction. Fourth and finally, I will suggest how the problem that I have identified here, namely global governmentality, might be addressed in the light of Fraser's theory.

I. GLOBALIZATION, JUSTICE AND THE POLITICAL

The "westphalian" version of Fraser's framework of social justice was based on problems concerning the class structure and the status order of societies. Against these,

she proposed the combination of redistribution and recognition.¹ She has enlarged this framework in two essays.

First, in *Reframing Justice in a Globalizing World* (Fraser 2008b, 12-29), Fraser stresses that in “postwestphalian” times, which are characterized by an emerging sense of the globality of many issues and events as well as by new forms of governance, the question of who is and should be implied when reasoning about justice and means of redressing injustice, a question which was formerly answered vis-à-vis the nation state, was now an open one; that this question had become a new subject for reasoning about justice itself. To give this new area of justice concerns a name, Fraser introduces the notions of framing and of representation – and it is exactly here where the political comes in. For representation is to the political what redistribution is to the economic and recognition is to culture: a means of redressing injustice concerning this sphere. So in times in which the proper frame for discussions of justice claims is not taken for granted any more, the political question of representation emerges as an additional matter of justice on the meta-level.

Forms of injustice that concern the political are not uniform, though. Fraser distinguishes three levels of such types of injustice. The first level – already known with regard to the nation state and thus from westphalian times – relates to issues of “ordinary-political misrepresentation” which refer to political decision rules that deny full political participation to some individuals or groups within a given frame (18-19). The second level of injustice, which became widely visible only with globalization, is “misframing” and refers to the way in which a political community’s boundaries are set; the basic diagnosis here is that in a globalizing world, the nation state does not always serve as the appropriate frame for addressing issues of justice anymore (19ff.). The third level of political injustice, finally, concerns what Fraser calls the “grammar of frame-setting” (25) and consists in “meta-political misrepresentation”, the failure to institutionalize “parity of participation” in deliberations and decisions concerning the “who” of justice, thus concerning the appropriate framing and internal rules of the units within which justice claims are to be taken up (26).

In her essay *Abnormal Justice* (Fraser 2008b, 48-75), Fraser comes back to her distinction of the three levels of representational issues and gives them a different twist. Here, she distinguishes among “what”, “who”, and “how” questions concerning justice-claims and asks for appropriate forms of redress when confronted with situations of dissent in the attempt to answer them. It is precisely these situations of dissent that she calls abnormal, even though she does concede that historically they have rather been the rule than the exception (50). But let’s briefly go through the three questions one by one.

The “what” question refers to the substance of justice and has been of central concern to Fraser’s thinking about social justice since the 1990s. Against one-sided approaches that focus on the economy or the cultural sphere alone, she calls for a “multi-

1] For the debate that this framework has elicited, see Fraser 2008a.

dimensional social ontology" that integrates concern for socioeconomic redistribution, for legal and cultural recognition, as well as – now explicitly integrating the political – for representation. To evaluate justice claims with regard to these three areas, she suggests her principle of "parity of participation" for all three of them. This justice principle calls for the dismantling of "institutional obstacles that prevent some people to participate on a par with others" in social interaction. Fraser argues that such obstacles can relate to all three spheres of justice and injustice: they can consist in the impediment by economic structures that deny some people the resources to interact as peers (maldistribution), in the prevention of full participation as partners by institutionalized hierarchies of cultural value that deny the requisite standing to some people (misrecognition), as well as in decision rules that deny equal voice in public deliberation and democratic decision-making to some people (misrepresentation) (60).

The "who" question, by contrast, refers to the scope or frame of justice and has to do with misframing, the second level of representational injustice that Fraser had distinguished earlier. According to her, unanimity regarding this question stems from the challenging of the hegemony of the westphalian frame by three distinct groups: by localists or communalists who seek solutions in subnational units – an example are independence-movements within nation-states; by regionalists or transnationalists, like strong proponents of the European Union, who go for larger, yet not fully universal units; and, finally, by globalists and cosmopolitans who transcend all boundaries by giving equal consideration to all human beings (56). Not entirely in accord with any of these three groups, Fraser herself suggests "reflexive and determinate" theorizing (61) to work through these conflicts – and proposes the "all-subjected principle" to solve them. "What turns a collection of people into fellow subjects of justice," she writes, "is neither shared citizenship or nationality, nor common possession of abstract personhood, nor the sheer fact of causal interdependence, but rather their joint subjection to a structure of governance that sets the ground rules that govern their interaction." The examples for such governance structures that she gives are the World Trade Organization (WTO) and the International Monetary Fund (IMF) as well as other organizations that "regulate the interaction of large transnational populations." (65)

The "how" question, finally, is connected to the third level of representational injustice and has to do with the modes in which "who" questions are assessed. Here, the hegemony of states and elites as agents of such assessments is challenged. Fraser suggests the application of the all-subjected principle to disputes over the "who", as well, and calls for "dialogical and institutional" theorizing (67) to tackle the problems arising here. According to her, that theorizing would have to integrate both the civil society as agent for democratic dialogue and formal institutions that have the capacity to warrant claims and make binding decisions (69).

II. GLOBAL GOVERNMENTALITY

In order to think about what is missing from Fraser's account of globalized justice, I would like to start off on a personal note. When I last went to Guatemala, in the spring of 2009, several of my friends there – most of them men who once studied agronomy and who have been interested in questions of land distribution for a long time – were working as gender officers. None of them showed particular excitement about this assignment, which was one among other tasks and responsibilities on a job with one of the country's peasants organizations, or rather, given the ongoing trend of NGO-ization of Guatemalan civil society, with one of the country's NGOs working on peasant issues.² The lack of excitement didn't stem from my friends' weak affinity with feminist claims, however. Rather, it was due to their perception of the organizations' political priorities and to knowing that their gender agenda comes from complying with the conditions and demands of the foreign donors that provide for large parts of the organizations' budgets. And such influences of foreign donor agencies on the agendas of NGOs are not an issue that is particular to Guatemala. Neither are they an issue that is restricted to demands about the integration of gender into the agenda of peasants organizations – in fact, gender-based organizations themselves have been severely affected by such influences. According to Sonia Alvarez, for instance, shifts in donors' priorities have led women's NGOs in many countries of South America to turn away from movement-oriented activities to more technical oriented ones (Alvarez 1999, 196-97). Islah Jad, for her part, has observed the transformation of Arab Women's Movements into a set of NGOs. Following Jad, this process has changed these movements in several respects. First, concerning their aims, she identifies a shift away from cultural, political, and charity concerns to social aims alone. Second, concerning numbers, she points to a decline of involved women and, along with this decline, a decreasing reach of the organizations in question. Third, she highlights an increasing hierarchization of the internal organizational structure of women's groups turning into NGOs (Jad 2004). And there are many more examples from these and from other parts of the world, as well (Hudock 1999; Townsend u.a. 2002; Bebbington u.a. 2008). Even governance studies have taken up this issue and discuss it – sometimes critically, but mostly stressing its potentials – under the headings of "soft power" or "soft forms of governance" (Brunnengräber/Randeria 2008; Göhler u.a. 2009).

I myself suggest calling this type of subscription of NGOs to the conditions, the thematic and organizational guidelines set by international donors "global governmentality". I hold that this form of power, or rather, its effects, is a serious problem that should be addressed when talking about justice in a globalizing world. So let me briefly explain how the conditioning of aid can be interpreted as a form of global governmentality.

2] It is particularly interesting in the context of Fraser's work to note that peasant issues in Guatemala are issues of the distribution of land, thus of maldistribution, as well as of racism and disrespect concerning indigenous norms and culture, and thus of misrecognition.

Michel Foucault introduced the term governmentality in his lecture series *Security, Territory, and Population*, to describe a new and complex form of governance, of decentralized state power, which is organized through a diverse set of institutions, procedures, analyses and tactics and that addresses the population. Its historical precedent is pastoral power, the benevolent power of the shepherd who looks for his flock. The new form of governmentality Foucault talks about integrates central elements of pastoral power – its explicit aim, regarding the population, is the management and prosperity of the entire unit and the well-being of its single members. It is a form of power that does not so much work with compulsion or discipline but rather by establishing norms of the sound and the rational that are to affect its subjects' thinking and self-conduct (Foucault 2007).

In *Governing through the Social*, Christina Rojas has argued that aid to poor countries was a mechanism of global government of the sort described by Foucault, one of its basic means being the establishment of "a relation between donor and recipient regulated by the promise of transforming the recipient country" (Rojas 2004: 98).³ Concerning the mechanisms of this mode of governance by aid, Rojas convincingly stresses the role of conditionality. But while in her analysis she mainly looks at how IFIs (international financial institutions like IMF and World Bank) as well as big bilateral donors like the US Aid Agency (USAID) have influenced the *states* that have been receiving their funds and programs, I hold that the governmentality paradigm aptly describes what often happens to political and social movements as well. This process has at least two steps. The first step is their transformation into "proper" NGOs with an appropriate internal structure for receiving foreign funds as well as aid workers/consultants. The second step, then, is that they adjust their rhetoric and, almost inevitably, their actions and agendas to the conditions and ideas of their donors. These adaptations are not an entirely voluntary act, since those kinds of institutions are in need of funds. But neither are they forced acts, since, in principle, the organizations could refuse the funds. Additionally, the donors who formulate the conditions for receiving the funds usually act with best intentions regarding the recipients' well-being, or even the well-being of the addressees of their recipients,

3] In this light it is striking how, for instance, the German development corporation GIZ presents itself. If you go to the organization's English language website and click on "About GIZ", you can read: "Working efficiently, effectively and in a spirit of partnership, we support people and societies in developing, transition and industrialised countries in shaping their own futures and improving living conditions. This is what the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH is all about. [...] As a federally owned enterprise, we support the German Government in achieving its objectives in the field of international cooperation for sustainable development. We are also engaged in international education work around the globe. [...] We advise our commissioning parties and partners on drawing up plans and strategies, place integrated experts and returning experts in partner countries, and promote networking and dialogue among international cooperation actors. Capacity building for partner-country experts is a key component of our services, and we offer our programme participants diverse opportunities to benefit from the contacts they have made." (<http://www.giz.de/en/profile.html>; accessed March 3rd, 2011) So here, too, the implementation of development politics is presented as a mode of global partnership that improves people's well-being.

namely, the populations of the countries they are active in. Nevertheless, I hold that these acts are induced by power, and that what is happening here is a reshaping and thus the distortion of political agendas.⁴

III. CHALLENGING THE “SCALES OF JUSTICE” FRAMEWORK

It is not as if Nancy Fraser hasn't addressed possible forms and implications of governmentality in a globalizing world. In fact, her essay *From Discipline to Flexibilization? Reading Foucault in the Shadow of Globalization* (Fraser 2008b, 116-30) is dedicated precisely to this task. Here, Fraser diagnoses the emergence of “a new type of regulatory structure, a multi-layered system of globalized governmentality, whose full contours have yet to be determined” (Fraser 2008b, 124). Nevertheless, what she calls “globalized governmentality” does not include what I myself have characterized as “global governmentality”, namely forms of transnational power relations between international institutions and OECD countries with their organizations of development cooperation on the one hand, and the states and civil society organizations that receive these kinds of “cooperation”, often conditioned financial aid and consultancy, on the other hand. That Fraser doesn't herself address these forms of power when thinking about “globalized governmentality” does not mean that they were not compatible with it, however. According to her, globalized governmentality is characterized by first, new multi-levelled governance structures that transcend the nation state; second, the dispersion of regulating entities and the formation of networks of such regulating entities like states, supranational organizations, international nongovernmental organizations (INGOs) and quasi-NGOs (QUANGOs); and third, new forms of subjectification addressing and affecting actively responsible as well as flexible agents (125ff.). Given these three elements, I hold that the first two accurately describe current issues of development aid related forms of global governmentality, especially with regard to new forms of donor cooperation as well as the outsourcing of tasks formerly undertaken by state controlled development institutions to private sector firms. Concerning Fraser's third element, however, the new forms of subjectification, I do not think that it so far encompasses all of what happens in the course of global governmentality. For what Fraser does not address are power effects on collective actors like movements and NGOs in the global south, effects that refer to the identity, self-understanding and political agenda of the various institutions constituting civil society;

4] These effects of power on the political agenda of specific groups might be reminiscent of the third dimension of power that Steven Lukes has put forward in *Power: A Radical View*, namely, “that people's wants may themselves be a product of a system which works against their interests” (Lukes 2005, 38). In some sense I do think that there are some similarities. But besides the fact that other than in Lukes' case, the focus here is on global phenomena, the governmentality framework provides for a way of arguing that does not imply any commitment to determine what people's real interests are. Nevertheless, I do hold that in the case of global governmentality, too, something like the distortion of political wants is happening.

power effects that I think we nevertheless can conceptualize as somehow similar to the subjectification or subjection of individuals.

So what is missing from Fraser's take on governmentality in a globalizing world is identifying new entanglements and power relations between governing institutions and their old and new subjects. She does look at institutional change beyond westphalianism when she addresses the first two of the elements of globalized governmentality that she identifies. She also mentions changing modes of subjection when she addresses the third element that she has distinguished. Nevertheless, the subjects of governmentality she looks at are the same as in pre-globalized forms of governmentality: individual subjects only.

This lack of attention to new entanglements and power relations between governing institutions and their possible subjects, especially those between states and related institutions in the global north and states and civil societies in the global south, is also reflected in the way in which Fraser conceptualizes transnational civil society under the all-subjected principle when talking about redressing forms of injustice. I would like to argue that this conceptualization is too Habermasian, or, in other words, not Foucaultian enough. For if we take the effects of global governmentality, the possible modification or distortion of actor's agendas, seriously, the dialogue in those transnational arenas of civil society might look less democratic than it appears at first sight. Fraser is very aware of all sorts of impediments to participatory parity – but she doesn't address the problem of distorted participation. With regard to the basic principle of her justice theory, power is conceptualized as something that remains external to civic dialogue. It is in play when some people are *prevented* from participation in terms of parity (16) or when they are *excluded* from participation altogether (26). What Fraser does not address is situations in which they do participate, are taken seriously and loudly voice concerns, but in which the content of these voiced concerns might be affected by productive forms of power that are connected to global governmentality. Taking up on considerations of Dana Villa's, we might say that Fraser doesn't sufficiently address the self-surveillance of the civically virtuous world citizens and NGOs (who have internalized the hegemonic conceptions of the common good or at least their main donor's conceptions of what their basic goals should be, which usually go together with a conception of the common good) or communicatively rational agents (who have internalized the hegemonic conception of what constitutes "the better argument" and proper organizational conduct) (Villa 1992, 715). In Fraser's globalized theory, power functions as a barrier to civic participation rather than as something that might run right through such participation. Power, that is to say, is conceptualized as repressive rather than productive.⁵

5] Interestingly, Fraser has addressed the problem of global governmentality in one of her most recent essays on feminism. Looking at the ways in which second wave feminism has, even if unwillingly so, played into the hands of neoliberalism, she writes: "In the postcolonies [...] the critique of the developmental state's androcentrism morphed into enthusiasm for NGOs, which emerged everywhere to fill the space vacated by shrinking states. Certainly, the best of these organizations provided urgently needed

IV. CONCLUSION

Thinking within the triad of redistribution, recognition and representation, one could hold that the effects of global governmentality that I have talked about are unintended effects of efforts to global redistribution – for at least ideally, this is what development cooperation is all about – in the realm of representation and politics. So these effects could be seen as somewhat similar to the unintended effects of affirmative forms of redistribution in the realm of recognition that Fraser herself has addressed in her well-known essay about affirmative vs. transformative measures of redistribution and recognition within the westphalian frame – the images of the lazy “welfare queen” or, to complicate matters a bit, of the recipient of German “unemployment insurance no. 2” who is unwilling to work, are only two examples out of many. Within the frame of the nation state, these unintended effects have led Fraser to the rejection of affirmative measures of redistribution, in other words of measures which are aimed at “correcting inequitable outcomes of social arrangements without disturbing the underlying framework that generates them,” and to favoring transformative modes of redistribution, namely modes aimed at “correcting inequitable outcomes precisely by restructuring the underlying generative framework.” (Fraser 1995, 82) But is this a solution that could be globalized? And if so, how would that look like?

If we really wanted to interpret the mentioned effects of global governmentality as unintended as well as problematic consequences of development cooperation, understood as an attempt at global redistribution of wealth and knowledge, one globalized “Fraserian” solution could be the critique of such measures and the attempt to think of more transformative forms of redressing global maldistribution. In fact, much of such work has already been done in the transdisciplinary field of post-development studies, be it within its strand connected to normatively infused postcolonial theories (e.g. Rahnema/Bawtree 1997; Ziai 2007) or to its rather economist and sometimes neoliberal versions (e.g. Moyo 2009).

But it could also be that the mentioned effects of global governmentality are consequences of new forms of global governance that we either cannot easily or do not want to eliminate. In that case they had to be addressed in a different way when thinking about global justice.⁶ One version of doing this could be to integrate the distortion of

material aid to populations bereft of public services. Yet the effect was often to depoliticize local groups and to skew their agendas in directions favoured by First-World funders” (Fraser 2009, 111). And concerning transnational feminist activism and the ways in which it was able to build “a presence in ‘global civil society’ from which to engage new regimes of global governance” Fraser notes that it became entangled in similar problems – as an example she states that campaigns for women’s human rights have focused overwhelmingly on issues of violence and reproduction, as opposed to, for instance, poverty (112-13.). So far, however, Fraser has not integrated these insights into the conceptual frame of her globalized justice theory.

6] At first sight, there might be a difference between a political and a theoretical take on the power effects of global governmentality. Politically, all a theorist might be able to do is to address them, to put the problems they create on the agenda of reasoning about global justice, hoping that this enhances a general

political voice into the account of forms of injustice, in other words to count with deeper forms of abnormalities with regard to justice claims and disputes than Fraser already does.

In her essay on *Two Dogmas of Egalitarianism* (Fraser 2008b, 30-47), where she speaks about different ways of dealing with the "how" question of justice, Fraser distinguishes what she calls the "normal-social-scientific" approach, which she rejects, from the "critical-democratic" approach, which she endorses. The "normal-social-scientific" approach is characterized by the assumption that the "who" question of justice can be answered by scientifically determining who is affected by a particular issue (Fraser 2008b, 41). The "critical-democratic" approach, by contrast, combines a "critical-theoretical conception of the relation between social knowledge and normative reflection" with "a democratic political interest in fair public contestation" (42). To date it is unclear whether this latter approach is conceptualized critically and democratically enough to be able to deal with the justice deficits that arise from global governmentality, or whether for that end we must add a "new-entanglements-and-power-awareness" dimension to it. Neither do we know how that would translate to the arenas that discuss the "what" question. But globalization doesn't only create political problems; it creates theoretical ones as well. Solving the latter might still be easier than solving the former. Yet, it remains a complicated task.

ina.kerner@sowi.hu-berlin.de

REFERENCES

- Alvarez, Sonia E. 1999. The Latin American Feminist NGO "Boom". *International Feminist Journal of Politics* 1(2): 181-209.
- Bebbington, Anthony J., Samuel Hickey and Diana C. Mitlin, eds. 2008. *Can NGOs Make a Difference? The Challenge of Development Alternatives*. London - New York: Zed Books.
- Brunnengräber, Achim and Shalini Randeria. 2008. Multi-Level-Governance in der transnationalen Politik - neue Perspektiven auf Nord-Süd-Verhältnisse? In *Mit mehr Ebenen zu mehr Gestaltung? Multi-Level-Governance in der transnationalen Sozial- und Umweltpolitik*, ed. Achim Brunnengräber, Hans-Jürgen Burchardt and Christoph Görg, 19-39. Baden-Baden: Nomos.
- Edwards, Michael. 2008. Have NGOs "Made a Difference?" From Manchester to Birmingham with an Elephant in the Room. In *Can NGOs Make a Difference? The Challenge of Development Alternatives*, ed. Anthony J. Bebbington, Samuel Hickey and Diana C. Mitlin, 38-52. London - New York: Zed Books.

awareness of them. The integration of these problems into a theoretical frame, by comparison, is another task, one that appears to be more complicated than simply stating that the problems in question do exist. But there seems to be an intermediate level, as well. Activists and practitioners who are affected by the problems in question, as well as scholars who are working on the level of policies, often think about solutions, too. These might, sometimes, be local in scope, or rather focusing on the symptoms than on the reasons of a justice deficit. Nevertheless, they can be good starting points for theoretical considerations of a more general nature. Concerning the problems created by global governmentality, the ideas put forward by Hudock (1999) can serve as a very good example in this regard.

- Foucault, Michel. 2007. *Security, Territory, Population. Lectures at the Collège de France 1977-1978*. New York: Palgrave Macmillan.
- Fraser, Nancy. 1995. From Redistribution to Recognition? Dilemmas of Justice in a "Post-Socialist" Age. *New Left Review* 212(6/7): 68-93.
- . 2008a. *Adding Insult to Injury. Nancy Fraser Debates her Critics*. Ed. Kevin Olson. London - New York: Verso.
- . 2008b. *Scales of Justice. Reimagining Political Space in a Globalizing World*. Cambridge - Malden: Polity.
- . 2009. Feminism, Capitalism and the Cunning of History. *New Left Review* 56: 97-117.
- Göhler, Gerhard, Ulrike Höppner and Sybille De La Rosa, ed. 2009. *Weiche Steuerung. Studien zur Steuerung durch diskursive Praktiken, Argumente und Symbole*. Baden-Baden: Nomos.
- Hudock, Ann C. 1999. *NGOs and Civil Society. Democracy by Proxy?* Cambridge - Malden: Polity Press.
- Jad, Islah. 2004. The NGO-isation of Arab Women's Movements. *IDS Bulletin* 35(4): 34-42.
- Lukes, Steven. 2005. *Power. A Radical View. Second Edition*. New York: Pelgrave.
- Moyo, Dambisa. 2009. *Dead Aid: Why Aid is not Working and how there is Another Way for Africa*. Penguin.
- Rahnema, Majid, Victoria Bawtree, ed. 1997. *The Post-Development Reader*. London - New York: Zed Books.
- Rojas, Christina. 2004. Governing through the Social. Representations of Poverty and Global Governmentality. In *Global Governmentality. Governing International Spaces*, ed. Wendy Larner and William Walters. London - New York: Routledge.
- Townsend, Janet, Gina Porter and Emma Mawdsley. 2002. The Role of the Transnational Community of Non-Government Organizations: Governance or Poverty Reduction?. *Journal of International Development* 14: 829-39.
- Villa, Dana. 1992. Postmodernism and the Public Sphere. *American Political Science Review* 86(3): 712-21.
- Ziai, Aram, ed. 2007. *Exploring Post-Development: Theory and Practice, Problems and Perspectives*. Oxon - New York: Routledge.

Rawlsian Compromises in Peacebuilding?

Response to Agafonow

Endre Begby

Centre for the Study of Mind in Nature, University of Oslo

Abstract: This paper responds to recent criticism from Alejandro Agafonow. In section I, I argue that the dilemma that Agafonow points to – while real – is in no way unique to liberal peacebuilding. Rather, it arises with respect to any foreign involvement in post-conflict reconstruction. I argue further that Agafonow’s proposal for handling this dilemma suffers from several shortcomings: first, it provides no sense of the magnitude and severity of the “oppressive practices” that peacebuilders should be willing to institutionalize. Second, it provides no sense of a time frame within which we can hope that endogenous liberalization should emerge in the local political culture. Finally, it provides no suggestion for what the international community should do if the desired liberalization should fail to materialize within that time frame. In section II, I show that Agafonow’s argument resonates poorly with the concepts and ideas that he claims to adopt from Rawls’s *Political Liberalism*. Instead, his argument evokes the guiding ideas behind Rawls’s later work *The Law of Peoples*. I offer a critical perspective on these ideas, focusing specifically on Rawls’s treatment of women’s rights. Section III applies this critical perspective to Agafonow’s arguments, before closing with an example of a more constructive and empirically informed approach that critical studies of post-conflict reconstruction could take.

Key words: liberal peace, human security, peacebuilding, post-conflict reconstruction, Rawls, political liberalism, women’s rights.

In “Human Security and Liberal Peace,”¹ J. Peter Burgess and I undertook to defend the idea of liberal peacebuilding from a recent spate of criticism. These critics, we argued, draw erroneous conclusions from otherwise legitimate data. It does not follow from the failure of any number of liberal peacebuilding operations that there is something inherently misguided about the principles and ideals of the liberal peace as such. In fact, we argued, much of the criticism can be seen to implicitly confirm the very principles and ideals it purports to criticize: for instance, individual peacebuilding operations are said to fail because they seek to impose political institutions from outside in a way that neglects the importance of self-determination and local ownership of political processes.² But self-determination and local ownership are precisely among the core principles underlying the philosophy of liberal internationalism. The problem, then, is not with these principles and ideals themselves, but with our failure to implement them in practice.

Further, we argued that these criticisms typically rely on rhetorical moves which underestimate the depth and extent of conflict in the communities in question. We can see this from critics’ brazen reference to a putative opposition of interest between “us” – Western hegemony looking to impose our political values from outside – and “them” – the

1] Begby and Burgess 2009.

2] Cf. Begby and Burgess 2009, 98, 100.

natives whose legitimate interest in upholding their own way of life is jeopardized by such hegemonic imposition. We pointed out that most liberal peacebuilding operations occur in the aftermath of humanitarian interventions or, at any rate, in the aftermath of conflict scenarios grave enough to warrant such intervention. In such scenarios, we cannot simply speak of a unitary political subject – a “they” – whose interests we must seek to take into account. Instead, the communities in question are torn precisely by deep conflicts of interest. Thus, 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. Nothing in the critics’ arguments could so much as begin to suggest that liberal democratic institutions are not best suited to that aim, no matter how challenging it can be to realize such institutions in practice.³

At heart, much of the criticism is rooted in the view that liberal internationalism is founded on the presumption of the absolute universality and political priority of a certain conception of human rights. Such rights-thinking, the suspicion has it, is fundamentally individualistic (i.e., Western), and may therefore fail to find a footing in more traditional societies. Burgess and I were concerned to show that with the more recent incorporation into liberal internationalist thought of ideas about human security,⁴ these suspicions can be quelled, at least to some extent. Human security, we wrote, accommodates the idea that “the needs of human individuals to be part of larger communities is among their *basic needs*, inasmuch as it is through membership in such communities that individuals derive their basic sense of self and the value-sets around which they organize their lives” (Begby and Burgess 2009, 99).

I. AGAFONOW’S ARGUMENT

These latter remarks provide the starting point for a recent response paper by Alejandro Agafonow.⁵ Agafonow raises questions concerning the ability of liberal internationalism – even as tempered by ideas concerning human security – to provide a framework for thinking about peacebuilding in conflict-torn societies. While by and large sympathetic to our argument, as well as to the larger program of liberal peacebuilding, Agafonow wonders nevertheless whether the liberal peace, with its rights-centered agenda, might be blind to 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. Facing up to the exigencies of such peacebuilding tasks might require privileging “community security over personal security, institutionalizing what, from a liberal point of view, are oppressive practices.” By contrast, if such compromises were ruled out in principle, in the name of upholding a liberal conception of individual political

3] Cf. Begby and Burgess 2009, 93.

4] See, for instance, the 1994 Human Development Report and the 2001 report of the ICISS.

5] Agafonow 2010.

rights and their priority, “it might close the door for liberalism to thrive in the long run in more traditional societies” (Agafonow 2010, 78).

One example of what, on this view, might have to be compromised in order to reach stable political arrangements in more traditional societies is – predictably, one must say – the rights of women, as well as principles that directly and asymmetrically impinge on women’s essential interests, such as marriage law and family law.⁶ The flipside of this would be that allowing such compromises in the early stages of peacebuilding might provide for liberalization to be achieved “in the long run,” but now a form of liberalization which would emerge spontaneously from within the local political culture itself. Such endogenous liberalization will hold significantly better prospects for achieving long-term stability than liberal institutions imposed from outside. As an illustration, Agafonow points to the emergence of the All-India Muslim Women’s Personal Law Board (AIMWPLB) in 2005. He writes:

This act of self-determination, prompted from within the Muslim minority itself, was motivated by what is perceived as discriminatory decisions against Muslim women. [...] It is possible that this act of self-determination would have taken more time to occur if Muslims did not have to live together with the Hindu minority. (Agafonow 2010, 82)

Thus, Agafonow’s argument is structured around two main ideas. The first idea is that achieving any kind of workable political stability in post-conflict societies might require that certain matters of importance be compromised, at least for the time being. And, one might think, women’s political rights and domestic security are less pressing concerns, at least right away, than putting an end to ongoing large-scale atrocities. The other idea is that allowing such interim compromises might, in the long run, induce the political factions to liberalize on their own initiative, in ways that bear the imprint of the local culture, and which therefore might prove more sustainable than similar measures imposed from outside.

Agafonow’s argument certainly does point to a real challenge. It is doubtful, however, that it is a challenge unique to liberal peacebuilding. Any kind of foreign involvement in peacebuilding processes will face these sorts of compromise-dilemmas. Maybe the best one can say is that liberal internationalism at least requires one to be honest and explicit about the sorts of ideals and principles that would be compromised in a given case. It thus provides a framework in which we can at least begin to assess the magnitude of the predicament that the local political culture finds itself in.

Concerning the second idea, I am less convinced: ideally, of course, one would hope for liberalization to emerge spontaneously from within. But Agafonow’s argument can hardly claim to provide much in the way of a constructive proposal here: for instance, it provides no sense of the magnitude and severity of the “oppressive practices” that peacebuilders should be willing to institutionalize. Further, it provides no sense of a

⁶ Agafonow 2010, 81.

time frame within which we can hope that liberalization should emerge spontaneously from within. Finally, and relatedly, it provides no suggestion for what the international community should do if the desired liberalization should fail to materialize within that time frame.

Like so many of the recent criticisms of liberal peacebuilding, then, Agafonow's argument no doubt succeeds in highlighting a problem (though not one that is unique to liberal peacebuilding), but fails to provide anything that could reasonably be described as an alternative. I will return briefly to these issues toward the end of this paper, after first considering another aspect of Agafonow's argument.

II. THE RAWLSIAN BACKGROUND

Agafonow calls on certain Rawlsian concepts to make his points: sensitivity to the need for compromise is the hallmark of *political liberalism*; the product of the compromise is what we may call an *overlapping consensus*.⁷ As Rawls famously argued, an overlapping consensus may serve as the foundation of legitimate and properly stable political institutions in irreducibly pluralistic societies. Agafonow may be right that societies emerging out of civil conflict can indeed be marked by an irreducible pluralism in this sense. Moreover, we can surmise that their ability to find a way of recognizing and working around this irreducible pluralism would be a vital first step toward forging the foundations of a lasting peace.

But there are obvious problems with Agafonow's invocation of Rawlsian concepts such as political liberalism and overlapping consensus to bolster his arguments. Political liberalism applies to well-ordered societies. "Well-ordered society" is a technical term in Rawls, subject to at least three substantial constraints. A well-ordered society is one, first, in which "everyone accepts, and knows that everyone else accepts, the very same principles of justice;" second, where "[the] basic structure [...] is publicly known, or with good reason believed, to satisfy these principles;" and finally, where "citizens have a normally effective sense of justice and so they generally comply with society's basic institutions, which they regard as just." Rawls summarizes as follows: "In such a society the publicly recognized conception of justice establishes a shared point of view from which citizens' claims on society can be adjudicated" (Rawls 1996, 35).

The societies we are considering are emphatically not well-ordered in Rawls's sense. Nor should we entitle "an overlapping consensus" just anything that will support a relatively stable form of political co-operation. Instead, an overlapping consensus is a consensus on the *actual* principles of justice (as opposed to a consensus about what to designate by the term "principles of justice"). That is, a Rawlsian overlapping consensus is, in substantial part, a consensus precisely about individuals' basic rights and their

7] Cf. Rawls 1996.

political priority.⁸ These rights are emphatically not a matter for compromise in political liberalism. Rawls's point in developing the theory of political liberalism is to show that an irreducibly pluralistic society can be stable in spite of its pluralism, provided it can achieve an overlapping consensus about these very rights and their priority. An overlapping consensus would have all parties agree that these are indeed the rights that constitute the foundation of their political co-operation, even though they might disagree about the further reasons *why* these are the rights in question. An overlapping consensus is *stable for the right reasons* (a matter of great significance in Rawlsian theory) only in virtue of being precisely a consensus concerning these very rights; a society can be *well-ordered* only in virtue of being founded on a consensus concerning these rights. Neither of these conditions holds in the sorts of cases Agafonow considers. Accordingly, his argument cannot support itself on the strength and prestige of the Rawlsian concepts that he invokes.

Instead, Agafonow's thinking evokes another strand of Rawls's philosophy, namely that which comes to expression in his later work *The Law of Peoples*.⁹ But this strand of thought is much more controversial, and enjoys little of the plausibility and robustness of the ideas that form political liberalism. *The Law of Peoples* is about the limits of toleration in international affairs. The argument on offer is that the threshold of tolerability (and hence of legitimacy) of political systems in international affairs is significantly lower than what we – liberal democracies – would recognize as affording legitimacy in our own domestic setting. Here is one way to think about it: *Political Liberalism* aims to articulate the ideals and self-image of a pluralistic democratic society – *our* society. The claim made in *The Law of Peoples*, then, is that not every society need satisfy the standards of a pluralistic democratic society in order nonetheless to be a legitimate partner in international co-operation to a democratic society like ours, i.e., in order to qualify for full standing in the "Society of Peoples."

But such "decent hierarchical societies," as Rawls calls them, must nonetheless satisfy substantial political constraints. Specifically, they must honor a "special class of urgent rights," which includes "freedom from slavery and serfdom, liberty (but not equal liberty) of conscience, and security of ethnic groups from mass murder and genocide" (Rawls 1999, 79). Further, they must also afford to every population group some input on political processes through what Rawls calls a "decent consultation hierarchy."¹⁰

Maybe this vision of a hierarchically structured but nonetheless decent society can provide a more precise sense of what Agafonow has in mind. Liberal peacebuilders, then, would need to be open to the tolerability of political institutions which compromise on certain sorts of non-basic rights, including but not limited to the right of democratic

8] A different way of putting this point is that only "reasonable" views should be taken into account in the overlapping consensus. Bigoted views have no place there. By contrast, Agafonow speaks of the need to take into account "the specific kind of irrationalities that may be found in traditional societies" (Agafonow 2010, 81).

9] Rawls 1999.

10] Rawls 1999, 71-78.

representation, but which do not compromise on basic or fundamental rights, such as the right to life and freedom from enslavement.

We note that, like Agafonow, Rawls pins the tolerability of such arrangements in part on the psychological supposition that such toleration may well, in the long run, prove the best way to get these societies to liberalize.¹¹ But we note also that, like Rawls, Agafonow is *not* explicit about the fact that such internally generated liberalization is not to be relied upon. 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. We further note that in Rawls's theory, decent societies are to be tolerated (and thus to be regarded as legitimate) *as they currently are*, not for what we hope they might become if left to their own devices. As pointed out above, Agafonow commits himself to no comparable stance concerning the long-term tolerability of compromised political arrangements which fail to precipitate the desired kind of endogenous liberalization. This is a serious lacuna in Agafonow's argument.

Finally, there is one further structural feature of Rawls's Law of Peoples which is worth remarking on here. The Law of Peoples is intended to satisfy the idea that non-Western, non-democratic societies may be hierarchically organized, in ways that are at odds with our liberal ideals (and which are thus "not fully just" by liberal lights¹²), yet which may be legitimate by domestic criteria. In this way, Rawls aims to make room for an alternative to the perceived individualism of Western society. Yet the way Rawls structures this proposal is striking and in many ways peculiar. In a decent, non-liberal society, political legitimacy is founded not on the assumption that every individual has adequate political representation (*qua* individual), but rather on the assumption that every *group* has adequate political representation, and that every individual is a member of some such group.

Rawls thereby appears to assume that the communities in question divide neatly along familiar group lines, say, of ethnicity or religion. His idea is that a political arrangement can be decent if each individual person is a member of some such sub-group, and every sub-group receives adequate political representation through the consultation hierarchy. What is remarkable about this proposal is that it asks no questions about the quality of representation *internal* to any of these groups, so long as basic human rights are not violated. Accordingly, there is a double sense in which a decent, non-liberal society can be hierarchical: it can be hierarchical, first, in the sense that not every group has the same political status.¹³ But there is also a second sense in which a decent society can be hierarchical, namely that each group could itself be hierarchically organized. Thus,

11] Cf. Rawls 1999, 62.

12] Rawls 1999, 24, 62.

13] For instance, it is compatible with decency that political office is restricted to members of a privileged group.

particular individuals can be doubly disenfranchised: first, in being a member of a non-privileged group; second, in being a non-privileged member of such a group.

These peculiarities are compounded by the fact that Rawls gives no serious consideration to the political interests of groups that cut across the recognized group lines, prominently, women. In the few places that Rawls does consider women as a group in their own right, it is with an eye toward stipulating that in decent hierarchical societies, their basic human rights are not violated.¹⁴ But even when these basic human rights are secured, there is still ample room for the institutionalization of oppressive practices, with no outlook toward improvement. With this, Rawls seems to hold that our definition of political decency has no need to take into account any special interest groups apart from the familiar religious or ethnic divisions, and that women's rights are adequately dealt with in terms of gender-neutral basic human rights. Needless to say, this is extremely controversial and deeply problematic.¹⁵

III. CLOSING REMARKS

Analogous causes for concern arise from Agafonow's argument. On his view, as we saw, women's rights are among the political principles that we might have to be prepared to sacrifice in order to achieve stable political institutions. Thus, we should have to be prepared to institutionalize practices that are oppressive from the point of view of liberal thought. To my mind, such proposals merit serious consideration only when they can meet a set of further constraints. First, they should offer a clear sense of the magnitude and severity of the compromises that we should be willing to accept, or, perhaps better put, what sorts of compromises we should not be willing to accept. Second, they should offer a clear sense that these compromises are interim measures, and that the status of these oppressive practices should be reevaluated on a relatively sharply defined time frame. Third, they should offer a clear sense of how we are to comport ourselves if the desired liberalization fails to precipitate at the end of that time frame. Finally, the rationale for the compromises in question should draw on actual empirical evidence concerning what is and what is not conducive to peace, stability, and political justice under the relevant conditions. They should not, that is, rest merely on speculative psychological claims about what "traditional societies" may and may not be ready for at the present time.

As a suggestion about the form that such studies might take, it might be helpful to consider the recent work of Paul Collier and associates.¹⁶ In opposition to the widespread assumption that democratization is intrinsically conducive to peace, Collier and associates

14] Cf. Rawls 1999, 75, 110.

15] For further remarks on the status of women in the Law of Peoples, see Nussbaum 2002; Tasioulas 2002, 384; Cabrera 2001, 174-75; Martin and Reidy 2006, 14-15.

16] Cf. Collier and Rohner 2008; Collier, Hoeffler, and Söderbom 2008. A popular exposition can be found in Collier 2009.

have shown that in low-income post-conflict societies, democratization actually increases the likelihood of relapse into civil war. This is a remarkable finding. But the explanation for the finding is surely not, for instance, that these societies are not ready for democracy, or that democratization must emerge spontaneously from within the local culture itself if it is to take hold. Instead, a rather more plausible and concrete explanation is that rapid democratization drastically reduces a government's ability to repress rebellion before such time as it has been able to properly address the issues that would typically provide incentives for rebellion. Further, Collier and Rohner take care to point out that these results are only "superficially troubling for the agenda of promoting democracy in low-income societies. [...] democracy may still be highly desirable because of its intrinsic merits. An implication is that in low-income societies that democratize additional strategies may be needed to secure peace" (Collier and Rohner 2008, 533).

This sort of research provides an empirically informed perspective on the kinds of challenges that confront liberal peacebuilding. In no way does it purport to overthrow the discourse of liberal internationalism as such, so much as to point out that the order and timing of reforms is relevant to our prospects for a peaceful and just society. Such nuance is altogether missing from much of the current criticism of liberal peacebuilding. Even when these criticisms are sound and draw on empirical example, they do not offer an alternative to the liberal peace. They serve rather as reminders that building stable and just political institutions takes time, and that it would be naïve and counterproductive to seek to implement all the relevant reforms in one go.¹⁷

endre.begby@gmail.com

REFERENCES

- Agafonow, Alejandro. 2010. Human Security and Liberal Peace: Some Rawlsian Considerations. *Public Reason* 2(1): 77-84.
- Begby, Endre and J. Peter Burgess. 2009. Human Security and Liberal Peace. *Public Reason* 1(1): 91-104.
- Cabrera, Luis. 2001. Toleration and Tyranny in Rawls's "Law of Peoples." *Polity* 3(2): 163-79.
- Collier, Paul. 2009. *Wars, Guns, and Votes: Democracy in Dangerous Places*. New York: Harper Collins.
- Collier, Paul and Dominic Rohner. 2008. Democracy, Development, and Conflict. *Journal of the European Economic Association* 6(2-3): 531-40.
- Collier, Paul, Anke Hoeffler and Måns Söderbom. 2008. Post-Conflict Risks. *Journal of Peace Research* 45(4): 461-78.
- International Commission on Intervention and State Sovereignty. 2001. *The Responsibility to Protect: Report of the International Commission on Intervention and State Sovereignty*. Ottawa: The International Development Research Centre.

¹⁷] This paper was written with financial support from the Research Council of Norway.

- Martin, Rex and David A. Reidy. 2006. Introduction: Reading Rawls's *The Law of Peoples*. In R. Martin and D.A. Reidy, *Rawls's Law of Peoples: A Realistic Utopia*. Malden, MA: Blackwell Publishing.
- Nussbaum, Martha. 2002. Women and the Law of Peoples. *Politics, Philosophy, and Economics* 1(3): 283-306.
- Rawls, John. 1996. *Political Liberalism*. New York: Columbia University Press.
- . 1999. *The Law of Peoples, with "The Idea of Public Reason Revisited."* Cambridge, MA: Harvard University Press.
- Tasioulas, John. 2002. From Utopia to Kazanistan: John Rawls and the Law of Peoples. *Oxford Journal of Legal Studies* 22(2): 367-96.
- United Nations Development Programme. 1994. *The 1994 Human Development Report: New Dimensions in Human Security*. New York: Oxford University Press.

The Basis of Universal Liberal Principles in Nussbaum's Political Philosophy

Matthias Katzer
University of Siegen

Abstract: In her political philosophy, Martha C. Nussbaum defends liberal political principles on the basis of an objective conception of the good of human beings. This paper examines whether her argument succeeds. It identifies three methods to which Nussbaum refers in order to select the central human capabilities, whose exercise is seen as constituting the human good. It asks whether these methods – the interpretation of actual ways of human self-understanding, the search for necessary anthropological features, and the idea of an overlapping consensus – can yield liberal political principles. The paper concludes that it is doubtful that the first two methods will lead to this result. As to the third method, it may yield liberal principles only insofar as it relies on the notion of human dignity which is interpreted in a way that contains a strong view of equality of all human beings. In this way, universal liberal principles are not primarily based on considerations of the human good, but on a genuine moral standpoint.

Key words: capability approach, dignity, equality, liberalism, Nussbaum, overlapping consensus.

In her work on political philosophy, Martha C. Nussbaum combines views that are often believed to be in tension with each other. She expresses support for universal liberal political principles, claiming extended liberties for each individual (such as religious freedom, free speech, and free choice of one's profession and spouse), the right to non-discrimination, and the right to participate in the democratic governance of the state. Thus, each individual should be recognized by the state as a being with an equal status, an equal power of reasoning and the equal capacity to make public and private choices. But Nussbaum bases these principles of liberal equality on a thick and objective conception of the human good, so that the good of an individual is independent of his/her personal desires, beliefs, or values. This objective conception of the good forms an integral part of her "capability approach", which identifies a set of central human capabilities that should be made the target of all political institutions because they permit human beings to achieve well-being.

A conflict between these views is not strictly necessary. It is not incoherent to think that it is part of the good of every human being to be recognized by others as a being with an equal status. But many liberals have insisted that any argument for liberal political principles requires, in one way or another, the idea of a fair adjudication between competing individual interests or claims. The equal moral status of persons, as it is expressed in principles of equal liberties and provisions of non-discrimination, is often not regarded as something that is good or valuable for all individuals as such, but only follows from a genuine moral standpoint. From this perspective, the argument presented by Nussbaum can be seen as a short cut from considerations on what is valuable to considerations on what social institutions should be like, circumventing the moral standpoint of fair

adjudication between competing interests or claims. It may be questioned whether this argument succeeds: Is Nussbaum's approach able to establish an argument for liberal political principles?

This is the question this paper wants to pursue. I will proceed in the following steps: First, I will have a closer look at liberal political principles and sketch Nussbaum's position with respect to different meanings of "liberalism". I will then give an outline of Nussbaum's capability approach that is supposed to present an argument for liberal principles. Following this, the paper will distinguish three methods Nussbaum uses in order to identify the central human capabilities and examine in turn whether they provide us with good reasons to adopt liberal political principles. The conclusion will be that the argument does not succeed if it is based solely on considerations on what is valuable for individual human beings: Liberal principles follow only if substantive moral premises are presupposed.

The details of Nussbaum's capability approach have somewhat developed over her various works in which she defends it. Nevertheless, the central claims of her view have remained sufficiently constant to present it as a coherent project and to draw on different texts in order to explain it. When the paper comes to examine the methods used to identify the central human capabilities, it will pay more attention to different stages of the development of Nussbaum's views than is necessary at the beginning of the paper.

I. LIBERAL PRINCIPLES

What makes a normative conception of the state a "liberal" one? Like many philosophical notions, the notion of "liberalism" is notoriously unclear and is employed in different senses by different philosophers. However, we may broadly distinguish two basic meanings of the term.

In the first meaning, "liberalism" is defined by the *content* of the moral principles for the legitimate use of state power: In this sense, any normative view of the state is only liberal if it requires states to respect extended liberties for all human beings and does not discriminate against specific groups of the population. The exact delimitation of the sphere of individual liberties may differ between different liberal theories, but any liberal view will have to include freedom of religion, free speech, the right to freely choose one's spouse, freedom of association and assembly, free choice of one's occupation, and more.¹ In addition, liberal principles require the state not to discriminate against individuals according to characteristics such as sex, skin colour, political and religious views etc. Liberalism is also strongly associated with the right of every citizen to equal democratic

1] Mill's harm principle, according to which a society may only limit the liberty of citizens in order "to prevent harm to others", may be regarded as the generic principle of liberalism, from which particular liberties can be derived (Mill 1991, 14). Indeed, Feinberg has defined the liberal position (with respect to the criminal law) as the position that the harm principle (complemented by an "offense principle", which he also attributes to Mill) exhausts the class of good reasons for limiting individual liberty (Feinberg 1984, 26).

participation in legislative and executive state power. In this text, by “liberal principles” we will understand primarily these three elements of liberty, non-discrimination and democracy.² They give expression to a certain conception of the equality of all human beings with respect to state power: All human beings should have an equal capacity of making public and private choices.

In the second meaning of “liberalism” this term is not defined by the content of moral principles, but rather by the *procedure of justification* through which moral principles for the use of state power are derived. This procedure is modelled in a way that gives weight to voluntary agreements between rational choosers with partially conflicting beliefs about questions such as the good life and substantive metaphysical issues, and consequently sets aside these controversial questions. Rational choosers agree on certain moral principles to be implemented by the state, but deeper metaphysical questions are avoided.

Of course, it is possible to be a liberal in both the procedural and the content-centred senses of the term. This is a position taken (in different versions) by famous liberals such as Rawls, Dworkin, Habermas and others. But it is equally possible to be a liberal only in the content-centred sense of the term, and to justify liberal principles by a wholly different method. This position has famously been taken by John Stuart Mill, who tried to defend liberal principles by using a utilitarian procedure of justification.

Martha Nussbaum has repeatedly emphasized her support not only for universal moral and political principles, but also for liberalism: “Any universalism that has a chance to be persuasive in the modern world must, it seems to me, be a form of political liberalism.” (1999, 9) She defends her approach against the objection that it is incompatible with liberal respect for individual autonomy (1992, 225). She has argued for equal treatment of women in social institutions (2000, 213ff.), for the equal rights of homosexuals (1999, 184ff.) and for religious tolerance (2008) – all clearly liberal positions in the content-centred sense. Moreover, it is important for Nussbaum that these liberal principles have universal scope, i.e. are valid for all cultures (2000, 5).

But it is less clear whether Nussbaum is a liberal in the procedural sense. To be sure, she claims to follow central elements of Rawls’ political liberalism (2000, 76). But in contrast to many procedural liberals, she makes use of what she calls a “thick vague conception of the good” (1990, 205). She concedes herself that this has not been accepted by many liberals: “liberalism has to take a stand about what is good for people, and I argue that it needs a somewhat more extensive conception of the basic human functions and capacities than many liberal thinkers have used if it is to provide sufficient remedies for entrenched injustice and hierarchy.” (1999, 11) This “more extensive conception” of the human good relies on an essentialism about human nature, i.e. the view that it is possible to achieve a determinate account of the essential properties pertaining to all human

2] In contemporary American politics, “liberal” has also come to refer to policies such as the redistribution of wealth from the rich to the poor and strong social policies in matters such as health care and education. This meaning of “liberal” is not the focus of our concern here.

beings, regardless of their culture (1992, 207). This essential account of human nature is then taken as the basis for the norms social institutions should meet. In contrast, liberals have often argued that political institutions should remain neutral with respect to competing conceptions of a good held by different individuals. Basing the task of political institutions on essential human properties has often aroused the suspicion among them that arbitrary (but allegedly objective) valuations would be smuggled into moral and political rules, unduly limiting the liberty of individuals to lead their lives according to their own conceptions of the good. Nussbaum tries to avoid illiberal conclusions by arguing that the task of social institutions should only be to provide *capabilities* for individuals, leaving each individual a free choice whether and how he or she makes use of these capabilities. Furthermore, she points out that for individuals free choice in public and private matters is itself an important component of the objective conception of the good as she sees it (1992, 225). Nevertheless, it can be questioned whether it is possible to present an objective argument for a determinate account of the human good, so that the account does not rely on arbitrary valuations by those defending it. Given the procedure of Nussbaum's approach, are we really provided with convincing reasons to adopt liberal principles in the content-centred sense?

The difficulty of finding an objective argument seems to be greatest in the case of liberal principles because they are often contested, both by different political regimes in the world and by intellectuals. Even some liberals – not least Rawls himself – have argued that liberal principles should not be regarded to have universal scope for purposes of international justice. According to Rawls, the international political order must not pressure states into adopting liberal principles (Rawls 1999, 59-62).³ Thus, these days the main challenge to Nussbaum's views does not come from a full-blown rejection of universal moral principles – although this may have been the case at the time when she developed the capability approach for the first time (Nussbaum 1992, 203ff.). Rather, it comes from a position that accepts universal rules, but holds these rules to be less expansive than a liberal account. Furthermore, the conception of equality of all human beings as expressed by the liberal principles is often thought to require an argument from a moral point of view, from some kind of fair adjudication of the competing interests or claims of individuals; thus, the argument for liberal principles could not only rely on what is good or valuable for a human individual. The question is, then, if the normative basis on which Nussbaum argues is sufficient to establish that universal moral rules need to have liberal content.⁴ In order to investigate this question, we will first have to sketch the basic

3] While Nussbaum also thinks that liberal principles should be adopted by the citizens of a state themselves, and not forced upon them by outsiders, she thinks that the list of capabilities can form the basis of policies of international institutions (Nussbaum 2006, 255ff., 316ff.). This departs significantly from Rawls' "full toleration" for decent non-liberal societies.

4] This focus on liberal principles does not mean that other elements of the capability approach are less important. Quite to the contrary, the capability approach seems to be most convincing in the case of human needs that are the object of the so-called social and economic rights (nourishment, housing, educa-

elements of Nussbaum's "capability approach" and see how it is meant to support liberal principles.

II. THE CAPABILITY APPROACH AND ITS LIBERAL ELEMENTS

Although in this text it is not possible to give a comprehensive account of Nussbaum's capability approach, we need to outline at least some of its main elements. Nussbaum starts from the Aristotelian idea that the good of an individual has to be determined in relation to the essential properties of the species of which he or she is a member. Thus, the good of all human beings has to be determined by searching for essential properties of human beings – properties that distinguish human beings from other beings (1992, 207). It is central for Nussbaum that first of all we regard ourselves as humans, and not as individuals of a certain sex, nationality, etc. Consequently, she asks us to look for features that we share with all other human beings – for features that "must be there, if we are going to acknowledge that a given life is human" (1999, 40). Closely related to this is the idea that we can identify certain areas of basic experiences of all human beings, such as mortality, physical needs, pain, mobility, cognitive capabilities, affiliation with other human beings etc., which are crucial for our well-being (1993, 263ff.).

Based on these fields of human experience, Nussbaum claims that it is possible to identify a list of central human capabilities which are judged to be important for human well-being. Thus, human well-being is understood to be the development and exercise of these capabilities, i.e. the "actual functioning". Nevertheless, it is left to the choice of each individual whether he or she will make use of any given capability. The role of social institutions is confined to aiming at the establishment of the *capabilities* to function in the ways that are judged to be valuable (2000, 87).

Nussbaum distinguishes three types of capabilities: First, there are basic capabilities, "the innate equipment of individuals that is the necessary basis for developing the more advanced capabilities, and ground for moral concern." (2000, 84) Basic capabilities are that what is given in each human being from the start, they are the constitutive features of human nature. But they are often only rudimentary and cannot be exercised without being developed (e.g., infants are able to learn a language, but this needs time and appropriate conditions). It is in virtue of the basic capabilities that human beings have a claim to support by others; they "exert a moral claim that they should be developed." (2000, 83)

Second, there are internal capabilities, i.e. "developed states of the person that are, so far as the person herself is concerned, sufficient conditions for the exercise of the requisite functions." (2000, 84) (An example would be a person that has learnt to speak and has no bodily impairments preventing him/her from speaking.) And third, there are "combined

tion, health care etc.), and indeed provides a compelling critique of some approaches to welfare economics and to social justice, although the paper does not argue for this point. The focus on liberal principles rather concentrates on the point that to me seems to be the most problematic and contentious one, given the normative basis of Nussbaum's arguments.

capabilities”, which are internal capabilities “combined with suitable external conditions for the exercise of the function.” (2000, 85) (In our example, the person has the combined capability to speak if neither his/her own condition nor threats from the outside prevent him/her from speaking.)

As Nussbaum argues, the goal of social institutions should be that every human being be secured the combined capabilities up to a certain threshold (2006, 292). In order to determine this threshold she refers to the “intuitive idea of human dignity”: if individuals fall below this level, we judge their lives to be “so impoverished that it is not worthy of the dignity of the human being” (2000, 72). Thus, the list of central capabilities is a list of the combined capabilities that give rise to claims each human being has towards society.

Nussbaum does not claim that the list expresses a metaphysical truth about human nature. Instead, she proposes that it can be based on nothing but the reflection of human beings about what is essential in their lives (1992, 207). Consequently, also she does not think that her version of the list is the final word on the issue. Instead, the list is “open-ended and subject to ongoing revision” within a cross-cultural dialogue, and the list is formulated in a way that leaves room to different specification within different societies (2006, 78). Nevertheless, she thinks that the prospects for finding a broad consensus on most items of the list are good (1992, 223).

After this brief overview over central claims of the capability approach, let us focus on some points that figure on the list of central human capabilities (in the version of 2006). Several points clearly call for liberal political principles:

“6. Practical Reason. Being able to form a conception of the good and to engage in critical reflection about the planning of one’s life. (This entails protection for the liberty of conscience and religious observance.)”

“7. Affiliation. [...] B. Having the social bases for self-respect and nonhumiliation; being able to be treated as a dignified being whose worth is equal to that of others. This entails provisions of non-discrimination on the basis of race, sex, sexual orientation, ethnicity, caste, religion, national origin.”⁵

5] We can note that these formulations of point 7B cast some doubt on Nussbaum’s claim that *all* elements of her objective conception of the good really take the form of *capabilities*, as opposed to states of affairs or “passive” properties. “Having the social bases of self-respect and non-humiliation” is clearly not a capability an individual can exercise and develop, but rather a property that depends on the social conditions under which an individual lives. While “self-respect” is indeed something that has to be actively developed by each individual, “non-humiliation” is not. Moreover, “being able to be treated as a dignified being” seems to be an awkward formulation; it is not clear in which way it differs from “being treated as a dignified being”, which is not a capability. Several authors that are in principle sympathetic to the capability approach have argued that not all human goods to be supported by social institutions are capabilities (see Buchanan 2004, 138; Ladwig 2009, 264).

“10. Control over One’s Environment. A. Political. Being able to participate effectively in political choices that govern one’s life; having the right of political participation, protections of free speech and association.”⁶

These points are meant to express “combined capabilities”: capabilities that all human beings should be able to exercise. For reasons of brevity, let us call the three central capabilities “practical reason”, “social affiliation” and “political control”, but we should keep in mind Nussbaum’s somewhat longer descriptions of these capabilities. We should also keep in mind that there are basic capabilities underlying these combined capabilities, exerting “a moral claim [to] be developed”. For example, we may assume that all human beings (from a certain age on) possess at least rudimentary forms of “practical reason”, deliberating on basic decisions about their lives, although not all of them live under conditions under which they are allowed to exercise their choices.

According to Nussbaum the recognition of the three mentioned capabilities directly leads to universal liberal principles referring to social institutions, namely liberty of conscience and religious observance, provisions of non-discrimination, the right to political participation and the protection of free speech and association. The paper will refer to these policies in a somewhat simplified way as liberty rights (covering both the “private” use of liberty, such as religious activities, and the “public” use, such as free speech in political contexts), non-discrimination and democracy.

What we will now investigate is whether Nussbaum’s approach does indeed provide good reasons to call for these liberal principles as universal rules to be implemented by political institutions in all societies. Therefore, we will not question the normative basis of her approach, but rather examine how far it will lead us.

III. THE ARGUMENTS FROM ACTUAL SELF-UNDERSTANDINGS AND FROM ANTHROPOLOGY

For the present inquiry the crucial question is, of course, how the items on the list of central human capabilities are selected. Nussbaum’s answer to this question seems to have changed over time: in different texts she explains the method used for the selection of the capabilities in somewhat different terms. It is possible to identify three different methods mentioned by her. We do not necessarily have to see these methods as excluding each other, but they point into different directions.

The method proposed in Nussbaum’s earlier writings “proceeds by examining a wide variety of self-understandings of people in many times and places.” In particular, she suggests consulting “myths and stories that situate the human being in some way in the universe” (1992, 215). Thus, fictional works from all cultures should be taken as material which tells us something about what human beings think about the elements

6] All points occur in Nussbaum 2006, 77 and with little modifications in previous versions of the list.

of well-being. We should note that this method concerns the *actual*, already *given* (if only implicitly) self-understanding of people from different cultures.

A second method consists of the intuitive reflection on necessary anthropological attributes. Essential human properties have to be shared (in their basic form) by all human beings, so that reflecting on what all human beings have in common will yield these properties. Therefore, Nussbaum asks us to ponder two questions, namely about “personal continuity” – “which changes or transitions are compatible with the continued existence of that being as a member of the human kind and which are not” – and about kind inclusion: “what do we believe must be there, if we are going to acknowledge that a given life is human?” (1999, 39-40) This method does not seem to emphasize our actual self-understanding in its entirety, but rather aims at sharpening this self-understanding through an anthropological reflection. The idea of the conditions of belonging to the species of humans is supposed to be the crucial criterion for selecting the list.

A third method consists of the idea of a cross-cultural dialogue on a common list. This dialogue does not aim so much at eliciting the self-understanding different people already have, but rather at constructing a list on which all can agree: the dialogue is “seeking a conception by which people of differing comprehensive views can agree to live together in a political community.” (2000, 102)

We will now examine whether these three methods yield the liberal principles that, according to Nussbaum, follow from her list. The first two methods will be treated in this section, the last one in the following section. We will take the methods for granted and will not question their normative significance, but rather ask if they yield the result they are supposed to yield. Throughout we should keep in mind that reflecting on the capabilities does not only apply to the identification of the *basic* capabilities. Once we have identified these, a *further* reflection will be needed in order to determine the threshold up to which the *combined* capabilities have to be a public concern, i.e. “the appropriate level beneath which it seems right to say that the relevant entitlement has not been secured.” (2006, 291)

But let us start with the basic capabilities underlying the points of the list quoted above: practical reason, social affiliation, and political control. It seems that the claim that at least the first two of these should be identified as essential human characteristics has considerable intuitive force, by either method of comparing the actual self-understanding of people from different cultures or of anthropological reflection. Let us briefly consider each of the three points in turn.

The first point is practical reason. In one sense of this term, all human beings (from a certain age on) do indeed have a capability to make reasoned choices. Consider choices such as the selection of a future spouse or the decision on how to make one’s living: It may be true that some beings are not very good at making such choices and that in some societies such choices are heavily restrained by social norms. But it seems preposterous to assume that human beings are in principle unable to reflect on these choices, and it is doubtful that they are often portrayed as such in literature. It also seems that if we

encountered intelligent beings that proved wholly unable to make reasoned choices, we would regard them as profoundly different from human beings.

In a similar way, social affiliation is, in a rudimentary form, a feature of human beings that is as universally shared as can reasonably be demanded. While certain exceptional individuals may achieve a state of detachment from and indifference towards others in the course of their lives, all human beings need the company of others to develop into mature beings.

With political control the case may be somewhat different. A labour slave in ancient Rome may be supposed to not have had any political control whatsoever, not even in a rudimentary form, but we can clearly recognise him or her as a human being. Maybe the combined capability of political control does not have a clear correlate among the basic capabilities, but rather rests on more general basic capabilities, such as that of practical reason.

Once we move to the combined capabilities, it becomes much more doubtful whether the methods proposed yield the intended results. With respect to the anthropological method, we have to notice that it is of no use here. The very idea of the combined capabilities is that they define the features “of a life that is worthy of [...] dignity”, not a life of bare survival (2006, 74). This entails that human beings can, in fact, fall below the relevant threshold and still be recognisable as human beings (although as ones leading miserable lives). The criterion of belonging to the species of humans would have to put the threshold at the level of bare survival, which is clearly not Nussbaum’s intention – and would also not yield liberal principles, since there are many dictatorships with severely restricted rights under which people nevertheless survive.

It is more difficult to ponder what would result from the method of examining the ways of self-understanding of human beings embedded in the literary heritage of different cultures. This method seems to call for a comprehensive research project which would have to confront a huge amount of material containing a great variety of the ways of self-understanding of human beings. We can certainly expect that most testimonies of human self-understanding will set the threshold for a good life considerably higher than bare survival. But it seems problematic to assume that most sources will set the threshold where it should be in order to justify liberal political principles. This is because political communities guided by liberal principles have been a relatively rare phenomenon in human history. We should not think that these principles are unique to the modern Western world. But if we only stay within Western history, the political recognition of these principles is a rather recent achievement. For most of the time, people have lived under conditions in which they had unequal civil and political rights, and in which these fundamental inequalities were not questioned by the dominant political thinking. It would be very surprising if most literary accounts of a good life from these times would indicate that a good life needs liberty, non-discrimination and democracy.

Indeed, there is evidence that some relevant sources – both from Western and from non-Western literary heritage – deny that human beings need to live in a political

community guided by liberal principles in order to lead a good life. These sources do not have to deny that a good life needs a measure of practical reason, social affiliation and political control: but they may deny that the relevant threshold for the fulfilment of these capabilities is lower than the threshold commanded by liberal principles, which entail an equal status of all human beings with respect to state power. But instead of demanding that everyone be recognized as an equal, some relevant sources might claim that equality is more than is needed for a good life. For example, pre-modern tales sometimes use organicist metaphors for explaining the relation between the individuals within a society, suggesting that the proper place for all members of society is the role to which they have been assigned by birth and custom. Social and political inequalities have often been justified this way. This is the upshot of the parable of Menenius Agrippa, as reported by Livy, which has been influential throughout many ages (Livy 1988, 325).

There is another possibility that we have to consider: Instead of holding that the necessary conditions for a good life require *less* than liberal equality, some sources might hold that a good life requires *more* than equality – that it requires to be an individual standing out from the crowd, to have more powers than the ordinary people in one's society. One could call this idea an aristocratic conception of the good life. We might find traces of this conception in the texts of Friedrich Nietzsche which have had a quasi-mythical influence on people from several generations: Consider Nietzsche's appraisal of the "sovereign individual" of whom Nietzsche says that "mastery over circumstances, over nature, and over all less reliable creatures with less enduring wills is necessarily given into his hands" (Nietzsche 1996, 41).⁷ The aristocratic conception entails, of course, that the human condition is tragic: if the conception was correct, it would, in principle, be impossible to create social conditions under which *every* human being could lead a good life. But if our reflection takes as given what relevant sources in different societies consider a good life, we cannot rule out from the start that this tragic situation might exist.

Therefore, the interpretation of human ways of self-understanding *alone* cannot tell us whether the central human capabilities – understood as being the capabilities that are necessary for leading a good life – include liberal principles. We get many contradictory answers to this question. If we want the literary heritage to give us a clear answer without these contradictions, we need to be highly selective in the choice of literary sources that we take into account.⁸ But then, we cannot claim that the method of making up the list of central capabilities consists only of interpreting the actual ways of self-understanding of people from different cultures. Rather, it consists of interpreting only those ways of self-understanding that have been chosen according to another method. This brings us to the next section.

7] Nietzsche consequently states that the idea of „a state of law“, in which „each will must recognize every other will as equal, would be a principle hostile to life, would represent the destruction and dissolution of man, an attack on the future of man“. (Nietzsche 1996, 57)

8] This point is also emphasized by Müller (2003, 324).

IV. THE ARGUMENT OF OVERLAPPING CONSENSUS

The third method for selecting the items of the list of central human capabilities is closely connected to what Nussbaum calls the “political” character of her approach. Indeed, she presents her approach as being “political, not metaphysical” and a “freestanding conception” in the way that John Rawls described his theory of political liberalism (2000, 76). This means that her theory does not aspire to be metaphysically grounded. Persons with conflicting metaphysical views could agree on the list of central human capabilities without agreeing on the reasons why they consider this list to be accurate. Thus, there could be an “overlapping consensus” on the list: The list could emerge from a process in which all persons consider the judgments of their fellow citizens on the elements of a good life, “seeking a conception by which people of differing comprehensive views can agree to live together in a political community.” (2000, 102)

This method is quite different from eliciting a common view of the elements of a good life from the *actual* ways of self-understanding of different people. What Nussbaum now claims is that a consensus on the list *could* emerge from a discursive process with a given aim, namely the search for a conception acceptable for people with different comprehensive views. Moreover, this process starts from the “abstract idea of human dignity” (2006, 75). The argument is that “by imagining a life without the capability in question,” persons from different cultures will come to believe “that such a life is not a life worthy of human dignity.” (2006, 78)

But how can we be confident that this process will lead to a consensus at all, and furthermore to a consensus on items similar to those on the list? Nussbaum’s reference to Rawls may give us two hints on possible answers to this question.

The first answer would emphasize that the overlapping consensus in Rawls’ political liberalism depends on the prior existence of common democratic institutions. As Rawls points out, his conception of justice, which he calls “justice as fairness”, constructs political principles by starting out from normative ideas that are implicitly contained in the institutions and the political traditions of a particular society: “Society’s main institutions, and their accepted forms of interpretation, are seen as a fund of implicitly shared ideas and principles. Thus, justice as fairness starts from within a certain political tradition” (Rawls 1993, 14). *Given* a liberal democratic society, we can assume that the institutions of this society gain broad support among the citizens, even if they do not share the same metaphysical views. This line of argument has the consequence that, according to Rawls, any overlapping consensus is valid only for a certain particular society with common political institutions;⁹ it cannot be taken to the international level where there exist societies with different political institutions.

9] According to Francesco Biondo, the particularity of the overlapping consensus has the consequence that it cannot be used in the way Nussbaum does (Biondo 2008, 317-18).

In contrast to Rawls, Nussbaum aims at an overlapping consensus with universal scope (2006, 304-5). The question arises, then, whether already existing political traditions indicate an overlapping consensus on liberal principles. One way to show that this is the case might be to point to the institutions that exist at the level of each society. The problem with this approach is that liberal principles are far from being universally accepted. Many states lack democratic structures and do not respect central liberties such as freedom of the press,¹⁰ and not in all of these states a strong public movement exists that demands equal civil and political rights. It might well be true that the public opinion in these countries has a tendency to accept more and more liberal ideas over time, as Nussbaum argues for the case of China, but she acknowledges herself that, for the time being, the universal acceptance of liberal institutions remains a hope to be fulfilled in the future (2006, 304). Thus, one cannot base present claims associated with universally valid political principles on institutions and public opinions prevailing in different states around the world.

Another way to identify universal institutions resting on liberal ideas would be to point to international law. The International Covenant on Civil and Political Rights from 1966, which has been ratified by more than 160 states, contains strong provisions of individual liberties (art. 18, 19, 21, 22), of non-discrimination (art. 26) and of democratic rights (art. 25). But according to international law itself, this broad support for liberal principles does not suffice for them to be universal norms. Under international law, each state is bound by treaties only if it has ratified them (Cassese 2005, 170). Even if a large majority of states has ratified a treaty, this does not mean that the remaining states are bound by it. In contrast, international customary law may forbid certain *grave* violations of human rights, such as genocide or deliberate killings of political opponents, but it does not condemn all violations of liberal principles (Cassese 2005, 59). Furthermore, persistent abuses, even by many signatories of the relevant human rights treaties, show that support for liberal principles (insofar as it is recognizable in the behaviour of states) is weak. Therefore, the existence of liberal norms within international law is not sufficient to establish that there is a widely shared political tradition of support for liberal principles. Rather, it seems that a considerable portion of the signatories of international human rights documents merely pay lip service to the norms they claim to support.

But indeed, Nussbaum has increasingly emphasized that a consensus is yet to be achieved over time (2006, 304). In order to show that such a consensus is likely to evolve, at least under certain conditions, one might point to another feature of Rawls' conception. This is the view that the overlapping consensus is only a consensus among *reasonable* comprehensive doctrines (Rawls 1993, 100). What is relevant for his political liberalism is not whether *all* persons agree on certain basic principles, but whether "reasonable and rational persons suitably specified" do (Rawls 1993, 115). In other words, the consensus

¹⁰ As indicative data, one may note that the Freedom in the World Survey 2010 by Freedom House lists 43 independent states as "not free".

can partly be seen as a hypothetical one, a consensus that we can expect to be reached given certain, specifiable conditions.

This element of the overlapping consensus resonates well with Nussbaum's view that a consensus is only likely to be achieved given certain discursive conditions. As she says, "people from a wide variety of cultures, coming together in conditions conducive to reflective criticism of tradition, and free from intimidation and hierarchy, should agree that this list is a good one, one that they would choose." (2000, 151) The conditions mentioned in this quotation, though they only seem to express necessary conditions for rational deliberation in a formal sense, depart already from many reflective processes taking place in the real world. Yet they are not sufficient to make a consensus on liberal principles likely: Even in academic discourse, where one tries to approximate the conditions of rational deliberation as far as reasonably possible, non-liberal political ideas persist. For example, the thought that individual liberties and the principle of non-discrimination can be curtailed in the name of the culture of the whole group seems to keep popularity among some authors.¹¹

So it is not surprising that Nussbaum gives further requirements for the procedure through which a consensus is to be achieved. As she says, persons have to reflect on the intuitive idea of "the dignity of the human being", i.e. they have to ask what the conditions for a human life with dignity are (2006, 74). But when the idea of human dignity is to be taken as a starting point, a substantially moral idea is presupposed within the discursive procedure. To be sure, the idea of human dignity has been associated with a variety of meanings, and it is difficult to make a definite judgment about its particular features. But it seems to be rather widely accepted that the idea, in its contemporary use, is meant to indicate a profound equality of the moral status of all human beings. Whereas in pre-modern contexts dignity was often attached to particular social roles, so that the notion was used to indicate status differences between persons, the modern use of the notion is a deliberate countermovement against differential treatment of human beings of different sex, colour, religion, and ethnic, political or other groups. Since dignity is supposed to "inhere" in every human being and gives rise to a set of fundamental rights, every human being has this set of fundamental rights and is in this respect equal to all others. Moreover, the fundamental status of every human being with respect to the power of the state is *solely* determined by the dignity which every human being equally has. This is the thought underlying the use of the notion of dignity in the Universal Declaration of Human Rights from 1948, which states in Article 1: "All human beings are born free and equal in dignity and rights." Nussbaum herself points out that the idea of dignity is closely connected to the idea of equality (2006, 292).

[11] An example might be Makau Mutua. Mutua argues that the group should be able "to determine for its individual members under what political, social, cultural, economic, and legal order they [...] live." In virtue of this demand, individual human rights, such as freedom of religion, may be restricted (Mutua 2002, 108).

It follows that if the idea of human dignity is taken as a starting point in searching for an overlapping consensus on central human capabilities, the procedure by which this consensus is found makes strong moral presuppositions. The participants in the procedure are not only asked to reflect on what is good or valuable for individual human beings and to determine universal political principles with respect to the promotion of this good. They are also asked to presuppose, in their search for universal political principles, an admittedly abstract, but strong moral point of view. The list of central human capabilities is not only determined by considerations of what is good or valuable, but also by strong moral considerations. The list expresses not only what human beings *need* in order to lead a good life, but also what they can *legitimately claim* from each other and from the state. If the moral point of view is presupposed, it cannot be ruled out that the scope of what human beings may legitimately claim is larger than the scope of what they need. This is why it might be the case that some items are on the list not because they are universal prerequisites of leading a good life, but because they express a universal moral standpoint.

Especially the liberal elements on the list might be better explained by this reason. Given the idea of human dignity and the idea of a fundamental equality of all human beings associated with dignity, it is quite convincing to adopt liberal political principles. The principle of non-discrimination seems to follow immediately from it. A principle of democracy is suggested by it, because it is only when human beings have, in one way or another, an equal right to participate in the collective use of the state power that their equal moral status with regard to state power is respected and that no one has, by virtue of the institutional design, more formal power than others. Individual liberties can also be seen as the expression of the equal right of individuals to express their views, make choices about their lives etc., because any restriction of liberty that cannot be justified by the need to respect the equal liberty of other human beings amounts to giving some human beings (to those who find a particular restriction of liberty desirable) more power than others to use coercive measures.

If the reflective procedure that Nussbaum asks us to follow in order to determine the central human capabilities does indeed presuppose a strong moral view of fundamental equality of all human beings, this does not mean, of course, that there is something wrong with the procedure. But it means that the procedure shows less than one might think. First, one cannot claim that, with respect to the justification of universal liberal political principles, Nussbaum's approach presents an alternative to other theories which employ a similar universal moral standpoint. Nussbaum's argument for liberal principles does not rest on the peculiarities of the capability approach. Rather, it depends on the universal moral standpoint she shares with other theories. Second, the procedure makes no effort to justify this universal moral standpoint, it simply presupposes it. A consensus is sought only with those who already share this standpoint. This does not mean, of course, that the universal moral standpoint is unjustified. There may well be good reasons to adopt it or to demand from other people to adopt it. But the reflection on the elements of a good human life does not help us much with finding these reasons. This reflection may flesh out what

an abstract political principle of equality concretely requires, but it does not provide this principle itself.

Therefore, the paper concludes as follows. Nussbaum claims to provide an argument for liberal political principles that is based on a reflection on the elements of a good human life. She mentions three methods of determining the central capabilities which have to be secured by social institutions for all human beings. The first two methods consist, respectively, of a comparison of the actual ways of self-understanding of people from different cultures and of a reflection on necessary anthropological attributes. These methods do not yield convincing arguments for the adoption of liberal principles. The third method, in contrast, supports liberal principles, but only because it starts from a moral point of view containing a strong principle of equality. This moral point of view is not peculiar to Nussbaum's approach. Insofar as this is true, Nussbaum's political philosophy does not provide a basis for universal liberal principles that is independent of other approaches to political theory. Her capability approach may be helpful to flesh out abstract political principles and to formulate them in a way that can be applied in different contexts, but it should be viewed as being in close alliance with other approaches, not as being in contrast to them.

katzer@philosophie.uni-siegen.de

REFERENCES

- Biondo, Francesco. 2008. Is Martha Nussbaum really political liberal? *Archiv für Rechts- und Sozialphilosophie* 94: 311-24.
- Buchanan, Allen. 2004. *Justice, Legitimacy, and Self-Determination. Moral Foundations for International Law*. Oxford: Oxford University Press.
- Cassese, Antonio. 2005. *International Law*. Second Edition. Oxford: Oxford University Press.
- Feinberg, Joel. 1984. *Harm to Others*. Oxford: Oxford University Press.
- Freedom House, ed. 2010. *Freedom in the World Survey 2010*. New York: Freedom House.
- Ladwig, Bernd. 2009. Anthropologische Argumente in der menschenrechtlichen Moral. In *Politische Anthropologie. Geschichte – Gegenwart – Möglichkeiten*, ed. Dirk Jörke and Bernd Ladwig, 245-69. Baden-Baden: Nomos.
- Livy. 1988. *From the Founding of the City. Books I and II*. Trans. B. O. Foster. Cambridge: Harvard University Press.
- Mill, John Stuart. 1991. On Liberty. In *On Liberty and Other Essays*, 1-128. Oxford: Oxford University Press.
- Müller, Jörn. 2003. Das normative Verständnis der menschlichen Natur bei Martha C. Nussbaum. *Philosophisches Jahrbuch* 110: 311-29.
- Mutua, Makau. 2002. *Human Rights. A Political and Cultural Critique*. Philadelphia: University of Pennsylvania Press.
- Nietzsche, Friedrich. 1996. *On the Genealogy of Morals*. Trans. Douglas Smith. Oxford: Oxford University Press.
- Nussbaum, Martha C. 1990. Aristotelian Social Democracy. In *Liberalism and the Good*, ed. R. Bruce Douglass, Gerald M. Mara and Henry S. Richardson, 203-52. New York and London: Routledge.

- . 1992. Human Functionings and Social Justice. In *Defense of Aristotelian Essentialism*. *Political Theory* 20: 202-46.
- . 1993. Non-Relative Virtues: An Aristotelian Approach. In *The Quality of Life*, ed. Martha C. Nussbaum and Amartya Sen, 242-69. Oxford: Clarendon Press.
- . 1995. Aristotle on Human Nature and the Foundations of Ethics. In *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*, ed. J.E.J. Altham and Ross Harrison, 86-131. Cambridge: Cambridge University Press.
- . 1999. *Sex and Social Justice*. Oxford: Oxford University Press.
- . 2000. *Women and Human Development. The Capabilities Approach*. Cambridge: Cambridge University Press.
- . 2006. *Frontiers of Justice. Disability, Nationality, Species Membership*. Cambridge: Harvard University Press.
- . 2008. *Liberty of Conscience: In Defense of America's Tradition of Religious Equality*. New York: Basic Books.
- Rawls, John. 1993. *Political Liberalism*. New York: Columbia University Press.
- . 1999. *The Law of Peoples*. Cambridge: Harvard University Press.

Religious Toleration and Public Funding for Abortions: a Problem with Christopher Eberle's Standard of "Conscientious Engagement"

Michael David Harbour
Harvard Law School

Abstract: "Justificatory liberalism" holds that citizens should refrain from advocating in favor of coercive policies for which they can only offer a religious justification. Christopher Eberle, a prominent critic of this view, calls this the "doctrine of restraint." Eberle argues that the restraint requirement unfairly burdens religious citizens by prohibiting them from acting on their religious commitments in the public sphere. As an alternative he offers what he calls the "ideal of conscientious engagement" which does not require restraint. In this paper I contend that Eberle's conscientious engagement standard fails to provide an adequate alternative to the "doctrine of restraint." Using the current controversy over public funding for abortion as an example, I argue that, under the "ideal of conscientious engagement," it is legitimate for the state to coerce religious citizens into violating their core moral commitments. Accordingly, this standard puts citizens' religious freedom at serious risk.

Key words: religious tolerance, liberalism, public reason, abortion, Rawls, Eberle.

Many liberals argue that the principle of toleration prohibits citizens from advocating in favor of coercive policies for which they can only offer a religious justification. This view is commonly referred to as "justificatory liberalism."¹ The idea is that such policies unjustly impose on citizens who do not share these same religious convictions. Tolerance, so the thought goes, dictates that the state should only coerce citizens on the basis of reasons that are, in some sense, accessible to all – in other words reasons that are sufficiently *public*.

Some critics argue, however, that this requirement itself violates the principle of toleration.² Part of what it means to hold certain religious beliefs, they maintain, is that these beliefs inform one's political advocacy. To demand that religious citizens refrain from endorsing policies on the basis of their religious commitments is to restrict their ability to practice their religion in a significant respect. If this concern is merited, then the justificatory liberal ideal of toleration appears untenable because it fails to respect the freedom of some religious citizens. In light of this apparent difficulty, one critic of justificatory liberalism, Christopher Eberle, has offered an alternative account of liberal tolerance and respect. According to Eberle (2002, 2009), the requirement that citizens should refrain from politically endorsing coercive policies for which they only have a

I would like to thank Robert Talisse for his helpful comments and encouragement on this paper.

1] This term was first introduced by Gerald Gaus (1996) to contrast his view with Rawls's (1993) "political liberalism." Here, though, I am using the term in the less narrow sense to include any view that requires that coercive policies achieve public justification, however this notion is understood. My use of the term thus coincides with Christopher Eberle's (2002) more inclusive usage.

2] In this paper I focus particularly on the work of Christopher Eberle (2009), but cf. Wolterstorff 1997.

religious justification – what he calls the “doctrine of restraint” – is too strong and lacks sufficient warrant. Instead, he offers the “ideal of conscientious engagement” which requires that citizens sincerely attempt to engage with and learn from their opponent’s point of view and that they genuinely pursue reasons their fellow citizens can accept for their preferred coercive policies. Should they ultimately fail to do so, however, the ideal of conscientious engagement does not further require that they withdraw their support for these policies.³

The virtue of the conscientious engagement standard is supposed to be that it better embodies the liberal principles of respect and religious toleration because it does not unjustly burden citizens who cannot abstain from appealing to their religious convictions when determining their political advocacy without thereby abandoning these very convictions. I intend to argue, however, that the doctrine of conscientious engagement fails in this regard. The problem is that this standard deprives religious citizens of any *in principle* objection to coercive policies that essentially *force* them to violate their fundamental religious convictions. I will illustrate this problem using the current political controversy over public funding for abortion. Under the conscientious engagement standard, religious opponents of abortion cannot, on the basis of religious toleration, object to a policy that effectively forces them to help finance an activity which they regard as morally abhorrent.

In this paper, I take no stand on the permissibility or impressibility of using tax dollars to fund abortions or on whether justificatory liberalism is the right view all-things-considered. My argument is that, irrespective of the merits of justificatory liberalism, the ideal of conscientious engagement fails to provide an adequate alternative *specifically* from the point of view of religious citizens.

I. RESTRAINT VS. CONSCIENTIOUS ENGAGEMENT

The crux of the justificatory liberal view is that state coercion must be equally justifiable to all in order to be legitimate. John Rawls articulates this commitment as follows:

Our exercise of political power is proper only when we sincerely believe that the reasons we would offer for our political actions – were we to state them as government officials – are sufficient, and we also reasonably think that other citizens might also reasonably accept those reasons. (1997, 771)

As Rawls famously argues (1993, 1997), reasons drawn exclusively from one’s particular “comprehensive doctrine” cannot serve as legitimate bases for coercive policies under this standard. Policy making, he contends, requires that we appeal only to those reasons that are shared across comprehensive doctrines or that set of reasons which forms what Rawls calls an “overlapping consensus” (1997, 776). Because not all citizens

3] For a detailed account of the principle of conscientious engagement, see Eberle 2009, 165ff.

in a pluralistic liberal democracy share the same religious beliefs, reasons anchored to a specific religious outlook fall outside the bounds of the overlapping consensus. Accordingly, citizens should refrain from endorsing coercive policies based solely on such reasons.

As was noted above, the problem according to some is that this requirement places an overly onerous burden on religious citizens, as it effectively prohibits them from acting upon some of their most fundamental commitments. As Eberle, puts it, “the liberal commitment to conscience would be quite a desiccated thing were it not to marry conscientiously formed *belief* to *action* guided by conscientiously formed belief” (2009, 158). Religious citizens cannot simply ignore their deeply held religious beliefs when voting on, and advocating for, various policy proposals without in essence abandoning their fundamental religious convictions. Demanding that they do so fails to respect them as citizens, as Nicholas Wolterstorff stresses:

It belongs to the *religious convictions* of a good many religious people in our society that *they ought to base* their decisions concerning fundamental issues of justice *on* their religious convictions. They do not view it as an option whether or not to do so [...] Accordingly, to require of them that they not base their decisions and discussions concerning political issues on their religion is to infringe, inequitably, on the free exercise of their religion. (1997, 105)

If Rawls is correct in thinking that coercive policies based exclusively on religious reasons are illegitimate, and Wolterstorff and Eberle are right in thinking that such reasons cannot be excluded if the freedom of religious citizens is to be properly respected, then we appear to be stuck in a serious bind. Eberle, however, argues that there is no reason to think that respect requires restraint, contra Rawls and others.⁴ He maintains that, out of respect, “a citizen or official should sincerely and responsibly *attempt* to articulate reasons for his or her favored coercive laws that his or her compatriots regard as sound” (Eberle 2009, 167; emphasis mine), but there is no basis for thinking that “a citizen or official [...] who discerns no such reasons, ought to restrain herself or himself from supporting those laws” (Eberle 2009, 168). As long as the religious citizen takes seriously, and sincerely engages with, the perspective of her opponents – by listening to them, trying to learn from them, and addressing their concerns – she respects their “basic worth” as citizens (Eberle 2009, 163). This is what Eberle calls the “ideal of conscientious engagement” (2009, 165-66). While her fellow citizens might reject the reasoning behind her preferred coercive policy, they will not feel dismissed or left out of the democratic process altogether as long as the conditions of conscientious engagement are met.

Take the controversy over gay marriage: if the only justification religious opponents of gay marriage can offer is that restricting the benefits of marriage to heterosexuals accords with the basic teachings of the Bible, then, as long as these citizens do their best to consider and engage with the arguments of gay marriage proponents and sincerely

4] For other proponents of justificatory liberalism see Gaus (1996), Audi (1997), and Macedo (2000).

try to formulate non-parochial justifications for prohibiting it, they have met their civic burden. If they genuinely believe that marriage between two members of the same sex is a grave sin, and that a country that permits it is doing great harm to its citizens, then they should be free to advocate for the passage of this policy. On Eberle's view, this is all that we can demand of religious citizens in the name of respect. It is unreasonable, he insists, to further contend that it is somehow disrespectful for religious citizens to act on the basis of their most fundamental beliefs and moral convictions. After all, this is how we would expect anyone to act (Eberle 2009, 158).

In a sense, Eberle's alternative of "conscientious engagement" tries to stake out a middle ground between the view that religious reasons are strictly impermissible as justifications and the view that religious citizens are permitted to simply ignore or backhandedly dismiss the concerns and objections of those who do not share their religious views. Under the ideal of conscientious engagement, citizens are free to practice their religion as they see fit – even when this includes voting for coercive policies on the basis of their religious convictions – provided that they do so respectfully, taking care to address and listen to the views of others.

II. PUBLIC FUNDING FOR ABORTION

I will not try to evaluate Eberle's arguments against the doctrine of restraint here. My goal is instead to assess the viability of the conscientious engagement alternative by drawing out some of its implications. With regards to gay marriage at least, it does seem like religious citizens would find the conscientious engagement standard more palatable than the restraint standard. Now consider, however, the controversy over public funding for abortion – a controversy which has recently reemerged in American politics with the current debate over healthcare reform. The sticking point for many reform opponents is that the use of government funds to expand health care coverage would mean that some of these funds will be used to perform abortions. The now infamous *Stupak Amendment*, introduced by Michigan Representative Bart Stupak, was designed to alleviate this concern by prohibiting any insurance policy paid for even in part by government funds from covering abortion procedures.⁵ While this amendment was not part of the final health care bill, President Obama signed an executive order shortly after the passage of healthcare reform prohibiting the use of federal funds for abortion procedures under the *Affordable Healthcare for America Act*.⁶ This is only the most recent manifestation of this controversy. Three years after *Roe v. Wade*, the Supreme Court decision that recognized the constitutional status of elective abortion rights, Congress passed the *Hyde Amendment* which prohibited the use of public funds to cover abortions under Medicaid.⁷

5] Affordable Healthcare for America Act, H.R. 3962, 111th Cong. (2009).

6] Executive Order 13535 (2010).

7] The current text of the *Hyde Amendment* can be found in Consolidated Appropriations Act of

I do not intend to address all the various controversial issues these types of restrictions raise. Even if one is a staunch proponent of abortion rights, however, it is not difficult to appreciate why such restrictions might be required in the name of religious toleration. Some religious citizens strongly believe that human life is sacred no matter what its stage of development, and, as such, abortion constitutes the murder of the most innocent and vulnerable of persons. Given their basic convictions, such citizens will not be able to regard the requirement that they financially contribute to a practice which they view as morally abhorrent as in any way reasonable. George Sher summarizes this point as follows, “any policy of government funding for abortions must draw upon tax monies collected from conservatives as well as liberals; and this must place conservatives in a position of actively supporting abortions rather than reluctantly tolerating their performance by others [...] A compromise which includes government funding of elective abortions may not be one which conservatives can reasonably be asked to accept” (1981, 371). Likewise, David Wong suggests that, in order to accommodate these citizens, “liberals could refrain from pressing for public funding for abortions, since this would require conservatives to actively contribute to the violation of a deeply held moral belief” (1984, 197). Should they refuse to do so, abortion rights advocates would fail to respect the liberty of their fellow religious citizens, as tolerance seems to require that we not force people to act against their own conscience.

In the face of this kind of coercive policy, it seems reasonable for the religious opponent of abortion to object, “it is not right for the state to coerce me on the basis of reasons that I could not possibly recognize as genuine moral reasons because they conflict so fundamentally with my core commitments.” Notice though that this is just the standard of restraint. The conviction that everyone has a positive right to an elective abortion will be completely inaccessible to religious citizens who believe that abortion is equivalent to murder. Such a reason is only accessible from within a particular comprehensive doctrine. As such, coercive policies based on this line of justification will be prohibited under the doctrine of restraint. The doctrine of restraint accordingly gives religious citizens recourse against coercive policies that would effectively force them to violate their own religious commitments.

Under the conscientious engagement standard, however, religious citizens forfeit this recourse. Suppose a group of abortion rights advocates, who believe that everyone should have access to elective abortions and that it is everyone’s duty to help finance this access, make a sincere effort to listen to, attempt to learn from, and engage with the arguments of their anti-abortion counterparts. Suppose further that they do all they can to offer reasons in support of public funding for abortions that their opponents could accept. Not surprisingly, however, they ultimately come up short. According to the ideal of conscientious engagement, these citizens have met their civic burden, and it is perfectly permissible for them to coerce their fellow citizens on these grounds *in spite of the fact that*

this will put some of their fellow citizens in the position of having to violate one of their deeply held moral and religious convictions. If conscientious engagement is the ruling principle, religious opponents of abortion cannot object to this imposition in the name of tolerance and respect.

Couldn't the opponents of abortion still argue, however, that this policy is illegitimate on the basis that it violates their *religious* liberty? Not if the ideal of conscientious engagement applies to all reasons equally and not just those reasons that some religious citizens happen to favor. If it is permissible, under the terms of conscientious engagement, for religious citizens to impose coercive policies even when the justifications for such policies fundamentally conflict with the moral beliefs of other citizens, then there is no reason why these religious citizens should be immune from similar impositions. It would clearly be unjust for a group of religious citizens to claim that they should be able coerce others on the basis of reasons that fall outside of the overlapping consensus but should not themselves be coerced on the basis of such reasons. Freedom of religion under this interpretation would amount to special coercive privileges for citizens of religious faith.

Another problem is that the religious convictions of citizens will sometimes conflict. Just imagine a religious sect of radical feminist who believe that elective abortions are a God given right. Under the conscientious engagement standard, they should be free to impose policies in favor of publicly funded abortions even if this policy violates the religious commitments of others. After all, according to Eberle, we would fail to respect the religious freedom of this sect if we were to tell them that they are not permitted to act on their core religious beliefs.

Notice that the conscientious engagement standard not only rules out the justification behind the *Hyde* and *Stupak* amendments but weaker accommodations as well. One could staunchly oppose strict across-the-board prohibitions like those required by *Hyde* and *Stupak* and still recognize the need to exempt religious citizens who have grave moral concerns about the practice of abortion in general. For example, those who believe that access to abortion is a fundamental right can still, out of respect, allow their fellow religious citizens to adopt a kind of *conscientious objector* status by granting them the ability to opt-out of taxes which are used to cover abortion procedures (Tribe 1985, 339). On the conscientious engagement standard, however, this plea for exemption lacks any warrant. If abortion rights advocates sincerely believe it is everyone's duty to equally share the burden for providing equal access to abortions, then they are under no obligation to make this accommodation, and their own moral convictions on the matter demand that they not do so in the absence of some overriding norm (such as the principle of restraint).

Recall that the virtue of the conscientious engagement standard is supposed to be that it does not produce a conflict between religious freedom and the burdens of citizenship. The above example shows that the doctrine of conscientious engagement fails in this regard. Under this principle, there could be cases in which religious citizens would be obligated to comply with a coercive mandate that conflicts quite explicitly with their fundamental religious convictions. It is hard to imagine that religious citizens would

find this consequence acceptable and consequently hard to imagine that they would regard the principle of conscientious engagement as consistent with their religious beliefs. Accordingly, I do not think that the doctrine of conscientious engagement represents a viable alternative to the doctrine of restraint.

III. CONCLUSION

Some critics are wary of justificatory liberalism because they believe it favors secularism and is unfairly hostile towards religion. For example, in a rather controversial footnote in the first edition of *Political Liberalism*, Rawls argues that any moral doctrine that does not recognize a woman's right to an elective abortion in the first trimester is "to that extent unreasonable" and therefore falls outside the overlapping consensus (Rawls 1993, 243 n. 3). This has led critics like Eberle to object that justificatory liberalism only excludes *religious* reasons while permitting similarly controversial secular justifications (like the belief that women have the right to an elective abortion in the first trimester).⁸

Perhaps this concern is merited; I have not commented on this issue in this paper. What I think Eberle fails to appreciate, however, is that alternative conceptions of tolerance and respect that do not require restraint, such as his own, also threaten the religious freedom of citizens in significant respects. Even if the restraint standard prevents religious citizens from seeking to impose coercive policies solely on the basis of their religious beliefs, it at least protects them from similar efforts by others. While the doctrine of restraint might be unacceptable to religious citizens, the ideal of conscientious engagement is hardly much better.

mharbour@jd13.law.harvard.edu

REFERENCES

- Audi, Robert. 1997. Liberal Democracy and the Place of Religion in Politics. In *Religion in the Public Square: The Place of Religious Convictions in Political Debate*, eds. Robert Audi and Nicholas Wolterstorff, 1-66. Lanham: Rowman and Littlefield.
- Eberle, Christopher J. 2002. *Religious Conviction in Liberal Politics*. New York: Cambridge University Press.
- Eberle, Christopher J. 2009. Basic Human Worth and Religious Restraint. *Philosophy and Social Criticism* 35: 151-77.
- Gaus, Gerald F. 1996. *Justificatory Liberalism*. New York: Oxford University Press.
- Rawls, John. 1993. *Political Liberalism*. New York: Columbia University Press.
- Rawls, John. 1997. The Idea of Public Reason Revisited. *University of Chicago Law Review* 64(3): 765-807.

8] In fact, one defender of the justificatory liberal view, Robert Audi (1997), equates "public" reasons and "secular" reasons.

- Macedo, Stephen. 2000. In Defense of Liberal Public Reason. Are Slavery and Abortion Hard Cases? In *Natural Law and Public Reason*, eds. Robert P. George and Christopher Wolfe, 11-50. Washington DC: Georgetown University Press.
- Sher, George. 1981. Subsidized Abortion: Moral Rights and Moral Compromise. *Philosophy and Public Affairs* 10(4): 361-72.
- Tribe, Lawrence H. 1985. The Abortion Funding Conundrum: Inalienable Rights, Affirmative Duties, and the Dilemma of Dependence. *Harvard Law Review* 99(1): 330-43.
- Wolterstorff, Nicholas. 1997. The Role of Religion in Decision and Discussion of Political Issues. In *Religion in the Public Square*, eds. Robert Audi and Nicholas Wolterstorff, 67-120. Lanham: Rowman and Littlefield.
- Wong, David. 1984. *Moral Relativity*. Berkeley: University of California Press.

Dussel's Critique of Discourse Ethics as Critique of Ideology

Asger Sørensen
University of Aarhus, School of Education

Abstract: Political philosophy should have the ambition to meet the conceptual demands of both government and governed. Critique of ideology is a classical modern way to see that such demands are met. In this perspective a marginal position is beneficial, namely when it comes to experiencing the particularity of a statement proposed as universally valid. The Argentinian-Mexican philosopher Enrique Dussel has exploited his marginality to point out shortcoming in modern critical theory, which he considers ideologically flawed. In this critique he employs Marx, Levinas, and the founding fathers of the Frankfurt school. The critique is mainly directed towards discursive ethics in terms of materiality vs. formality, where Dussel points to the material importance of economy, the body and teleological content for ethics. Apart from the epistemological benefits, being marginal has material importance, since it is in the peripheries of the world that the suffering is realized and thus experienced in the most extreme way, namely as exploitation, starvation, slavery and torture. As practical philosophy both ethics and political philosophy must be able to back up normative stands on such material matters as well as principles and procedures, and this is what Dussel reminds us.

Key words: ideology, discourse ethics, matter, victim, U-sentence.

For more than three decades the Argentinian-Mexican philosopher Enrique Dussel has been engaged in developing the Philosophy of Liberation in a critical dialogue with various philosophers, living as well as dead. Inspired by Levinas, Dussel's main concern was from the beginning to formulate an Ethics of Liberation, which was first conceived of as specifically Latin American (Dussel 1973-80), but latter simply as ethics, and the result is an impressive work, the *Ética de la liberación* from 1998, which is now being published in its fourth edition in Spanish and is in the process of being published in English.¹ Apart from Levinas, Dussel main philosophical inspiration in this project comes from a new reading of Marx, but nevertheless his favourite interlocutors during the whole period were philosophers representing discourse ethics, namely Apel, Habermas, and Wellmer. In this paper I will discuss what I consider one of the main issues at stake in the interchanges between the ethics of liberation and discourse ethics, namely Dussel's critique of discourse ethics to be merely formal and thus reductionist in its conception of ethics, that this formalistic reductionism is ideological in the classical Marxist sense, and that a better understanding of matter will lead to a better understanding of ethics.

First it must be determined in which sense practical philosophy includes critique of ideology (section I), second I will then sketch in what sense Dussel critique of discourse ethics is a critique of ideology (section II). One of the main points for Dussel is to point out that Apel's approach to discourse ethics is formalistic and thus reductionist (section III). In the case of Habermas Dussel recognizes that he allows matter to play a more important part in his idea of discourse ethics (section IV). But this is not enough for Dussel, who

1] It is being translated to be published soon at Duke University Press.

thinks of matter as intrinsically linked to universality and critique (section V). Dussel's concept of matter and materiality is thus both ambitious and comprehensive, comprising at least three distinguishable senses (section VI), and the negligence of discourse ethics concerning matter in this respect must be considered ideological (section VII). It is the intrusion of the Other that makes me side with the victim and thus make me capable of doing critical science (section VIII), but even though this shows Dussel's ethics to be of fundamental importance, it still does not make mainstream philosophical ethics and politics superfluous, since Dussel mainly focuses on what is unacceptable, whereas we would also like be able to chose the best ethical and moral alternative among those found acceptable (section IX).

I. A PRACTICAL PHILOSOPHY WITH UNIVERSALIST AMBITIONS MUST INCLUDE A CRITIQUE OF IDEOLOGY

Practical philosophy is about determining concepts in such a way that it can contribute to the realization of the good life. It is in this sense philosophers since Aristotle have considered ethics practical. How this kind of thinking is dealt with, however, varies a lot. Many philosophers naturally side with those in power and take the point of view of an agent to discuss how we – that is, those of us, who are sufficiently free – can organize our life and act in the best possible way. This point of view is perfectly sensible and legitimate, if “we” are actually a substantial amount of the population in question, and if “we” actually constitute a community, who rules itself democratically. If on the other hand we are dealing with a feudal society, an absolutist monarchy, or even a military dictatorship, where the very few rule the vast majority, then it is reasonable to ask, if the “we” of ordinary ethics is likely to express the viewpoint of people in general. Or whether the “we” is not more likely to function as a pseudo-including cover-up for serious conflicts of interests between various groups, classes, races or genders (see Adelson 1994, 4-5).

This way of questioning opens up practical philosophy to include the point of view of those members of a society, who are not as free as those in power. As Marx noticed (1969[1845-46], 46), the ruling ideas are always just ideas of the rulers, and this is relevant to point out no matter what kind of government is in power. When it is forgotten that the ruling normative ideas might have an origin that shows them precisely to be just ideas of the rulers; when it is forgotten that it is in the interest of every potential ruler to present his ideas as in the societal interest of all citizens (47), and that such ideas therefore might not be valid and beneficial for everybody; then one can consider such ideas ideological, and a critique of a set of such ideas can thus be called “critique of ideology.”²

2] Marx opposed ideology to science, whereas, for instance, Lenin thought of communism and bourgeois ways of thinking as both kinds of ideology (see Nogueira 1992, 185-88). Habermas (1971, 266-67) argues that the fault is to be found in Engels' naturalized conception of ideology. To me, it is sufficient to recognize the opposition between ideology as only of particular validity and thus opposed to what must be considered reasonable, that is, of universal validity.

Practical philosophy tries to determine normatively relevant concepts in the most universal sense, but of course philosophers are each by themselves ordinary people, influenced by time and location, by local culture and politics, and by social and economical conditions. The problem is that we often ourselves cannot see, what these conditions do to our thinking. As the Bible puts it, it is easy to see the speck in the eye of your brother, it is quite another thing to discover the log in your own eye (cf. Matthew, 7.3-5). In such processes the formation of consciousness literately happens behind our backs, as Hegel reminds us (1952[1807], 74 (A20)). Thoughts that we assume simply to be true as a matter of course, or that we sincerely believe to be the best possible expression of something universally valid, can under the right circumstances reveal themselves to appear valid to us only because of our particular living conditions. We are thus ideological in our thoughts without knowing about it; if we really knew about the matter in question, we would probably experience great difficulties believing in the ideological distortion of it.

Such a lack of consciousness of ones own dependency can be called hypocrisy, but is probably better named "false consciousness." Ideology and false consciousness consist of prejudices that we develop growing up and living in a human society. Some of them are fundamental for our orientation in an otherwise very complex world, and they are therefore very difficult both to discover and to change. This means, however, as Gadamer (1986[1960], 301-2, 457-58) has convincingly argued, that distance can be considered as a condition that contributes positively to the acquisition of knowledge, and the marginalized members of a society can therefore be said to occupy a privileged position when it comes to certain types of practical knowledge.

It is this privileged position that Dussel has taken upon him to exploit as much as possible as a philosopher. His doctoral work, which comprises both history, theology and philosophy, was done in the centre of the world system,³ in Spain, France, and Germany in the nineteen sixties, but as professor in philosophy in Mexico, he has since the seventies been back in the periphery, that is, in the position, wherefrom the ideological repression and false consciousness that rules in the centre is most easily revealed. The result is the so-called "philosophy of liberation," which is a practical philosophy in the above mentioned classical sense, that is, an ethical and political thinking, which has ambitions to be both universally valid and practically relevant. As practical philosophy it must contribute to the realization of justice in the world, and for the suppressed classes such a realization implies liberation. Therefore the expression "philosophy of liberation."

II. THE CRITIQUE OF DISCOURSE ETHICS IS A CRITIQUE OF IDEOLOGY IN A SENSE THAT HABERMAS HAS USED

For more than a decade Enrique Dussel has been conducting a dialogue with discourse ethics, mainly as it has been embodied by Apel, but also in relation to Habermas'

3] The term "world system" is employed with a conscious reference to the work of Immanuel Wallerstein, which is also of major importance for Dussel (1998a, 51-54). In this context, however, I will not discuss further questions concerning this aspect of Dussel's work.

version. During these encounters Dussel's general approach has been a critique of what he perceives as the ideological bias of discourse ethics, which he assumes to be resulting from having to view the world from within the peculiarities of the world centre. His general point is simple, namely that the position from which we experience affects our ability to experience and thus the resulting experience as such.

To back up this claim, Dussel can here draw on the force of an *ad hominem* argument, namely by referring to a quotation from one of Habermas' earlier works, *Theorie und Praxis*:

[I]n the developed countries the standard of living has been raised so much, also in the population at large, that the interest for liberation of society cannot any more immediately be expressed in economic terms. Alienation has lost its evident economic form as misery. [T]he proletariat as such is dissolved. (Habermas 1971, 228-29; see Dussel 1998a, 188; 1998b, 142)

As Habermas expresses it, it is obvious that the material living conditions can make people think of freedom in less economical and, one can add, material or even less corporal terms. In the most central centres of the world system, material necessities are almost unnoticeable. As Habermas puts it, alienation has lost its form as misery, meaning that in the rich centres of the world alienation does no longer by necessity imply starvation, pain and death in their literary senses. But, as Dussel (2005b, 341) is never tired to underline: In the periphery life is to a much greater extent confronted with death, and whether it is in the cities or on the countryside, in the periphery the daily life is marked by matter, namely by poverty and lack of social rights.

The basic thought is that apart from the spatial separation between centre and periphery, the more benign living conditions in the centre makes one ignorant to the sufferings of those excluded or living at the periphery, and that this also has an impact on the way we, the philosophers think of ethics and morality, no matter how much we discuss universality. Actually, it influences the very way we discuss universality. Dussel simply claims that it is the living conditions that conditions western philosophers such as Apel and Habermas to be less aware of those aspects of ethics, which are important for the victims of exclusion and suppression, the widows, the refugees, the orphans, the harassed women etc.

Dussel's critique of discourse ethics thus follows the classical scheme of a critique of ideology, claiming that what poses itself as universally valid – i.e. discourse ethics – can in reality be recognized as of a much more limited validity, since it has not escaped the conditions of its own origin. Hinting in this way, however, at the distinction between genesis and validity, also makes obvious that Dussel has to say more to deny the validity of discourse ethics. According to Marx (1969[1845-46], 18) an ideological conception is actually wrong, but it still remains to be shown that discourse ethics is ideological in this strong sense. Neither has discourse ethics yet been shown ideological in the even stronger sense, namely as a set of ideas, which are both false and necessary to uphold existing inequalities, since they blur the perception of those inequalities and thus function as

support of continuous exploitation and suppression. To get within reach of such strong conclusions, we have to look more closely into the basic logic of discourse ethics.

III. DISCOURSE ETHICS IS REDUCTIONIST; IT FOCUSES EXCLUSIVELY ON FORMALITY AND IGNORES MATTER

The basic idea supporting discourse ethics is that rational discourse is the original mode of language use. When one is communicating in the most basic sense, e.g. telling an interlocutor that he or she is right about something, one presupposes truth claims, and this is also the case when one is criticising, even when the critique is directed towards the very importance of communication, argument or reason as such. To criticise something means that you are arguing against something and that is already arguing, i.e. communicating is the sense used here – and, so the argument goes, therefore the original mode of language is communicative and orientated towards mutual understanding. So the basic relation between people using language is one aiming at mutual understanding, and all other uses of language are parasitic on this original mode. As Habermas puts it: “Mutual understanding is inherent as *telos* in the human language.” (Habermas 1988[1981], vol. 1, 387) To use language instrumentally or strategically one presupposes that the interlocutors still think we are communicating; it is only on this precondition that they can be influenced in a way so they are manipulated, fooled, or even deceived.

On the basis of this conception of language Habermas argues that moral norms can only be considered universally valid, if they can be submitted to a rational discourse and be accepted by all of the interlocutors possibly participating in such a discourse. This is formulated as a principle, the so-called “discourse ethical ground sentence”, or just “D” (Habermas 1983, 76). The idea is very similar to the way the truth of a proposition is supposed to be tested according to the critical rationalism of the late Popper (see Albert 1968, 29-31). If what we strive for is universal validity of norms, then we must let our values be tested by the most thorough critique in a rational discourse. Only if our ethical values can survive such a test can we consider them candidates for universal moral norms.

Such a conception of ethics, however, Dussel finds all too simple. Inspired by Levinas and Marx Dussel criticises Apel and Habermas to have put too much stress on the demonstration of the universal validity of moral norms and too little on the matter of ethics. Dussel accuses Apel to ignore “the sense of the ethical materiality of the life of the human subject” (Dussel 2005b, 341) and together with that, all empirical, historical, and material aspects. Instead, in his discourse ethics Apel allegedly only considers universal conditions of possibility for moral validity. The question of validity is given absolute priority in relation to questions of content, that is, of what is good. In the discourse ethics of Apel matter is not the core issue, neither negatively nor positively. According to Dussel, matter is simply relegated and ignored, and discourse ethics has no intention at all of grounding a material ethics. As far as I can see, this is his basic critique. But what does that mean? What is the matter? First, what can be said of matter and form in relation to

the rather traditional perspective of discourse ethics, and, then, what can be said about this matter, if we employ the perspective developed by Dussel in the ethics of liberation?

First of all, one still has to make a note concerning the use of the terms “ethics” and “morality”, and especially when discussing these matters in English. In mainstream Anglophone practical philosophy “moral” and “morality” signify something intuitive and often even collective or unconscious. “Ethics” signify the personal and systematic reflection about moral matters, as it for instance happens in moral philosophy, and this discipline is therefore called ethics. In that sense ethics is part of morality, namely the most systematic and consciously laid out part. The problem is that when we discuss these matters within philosophy in the tradition after Kant and Hegel, then the situation is almost turned upside down. For Habermas ethics thus has to do with substantial values and the good life in a particular community, whereas morality has to do with norms aiming at universal moral validity. According to Habermas discourse ethics should therefore rightly have been called a “Diskurstheorie der Moral” (Habermas 1991, 7). This means that when one makes the distinction between the content of and formal validity in German, then ethics traditionally deals with content, whereas morality is concerned with norms and their validity in terms of universality, that is, form.

This is reflected in the way Dussel argues for the importance of matter for ethics. He employs a kind of transcendental argument, stating that Apel cannot get away with his formalist strategy, since discursivity in the sense Apel uses it presupposes mutual recognition in a material sense. According to Dussel we cannot evaluate the propositional content of what is claimed in a discourse formally, if we do not listen to the content of what is said, and this presupposes respect for the interlocutor as a rational and reasonable person. So Dussel actually recognizes the relevance of Apel’s transcendentalist strategy, but claims that it has not been thorough enough. The transcendental condition of possibility for formal validity is that we in a discourse take each other seriously, i.e. that we recognize the personal dignity of the Other. This Dussel considers a “material moment” of ethics (Dussel 2005b, 341), where “moment” is understood in its Hegelian and not in its existential sense (see Hegel, 1952[1807], 73 (A18)). This, however, does not suffice for Dussel. The question of materiality is much larger in scope than just to function as a condition for the possibility of formal validity.

To appreciate Dussel’s point, we must look a little more into the logic of Apel’s transcendental pragmatics. The main idea is that it would be a contradiction to deny the propositional content of what one does in action. According to Habermas (Habermas 1983, 90) this implies a kind of grounding that breaks with the semantic idea of grounding as exemplified in deduction. Grounding in the transcendental pragmatic way relies on the recognition of a fact, namely that argumentation can only continue to be meaningful as a language game, if some pragmatic conditions are met in action. Such conditions are of the following kind:

Something, which I cannot deny without being confronted with an actual self-contradiction, but which I cannot ground deductively without being involved in a formal logical *petitio principii*. (Apel 1976, 72-73; see Habermas 1983, 92)

Deductively speaking such grounding would lead to a vicious circle. But Apel's point is pragmatic, concerning the action involved in speech. The claim is that if one's speech action takes form of an argument, then one cannot without contradiction deny that logic and reason matters (Habermas 1983, 93). It will be a performative self-contradiction to argue that one can simply ignore arguments. Quite the contrary, letting oneself be engaged in such speech actions actually implies that one recognizes in action, i.e. in practice, the forceless force of the better argument. Involved in this idea of inescapable conditions of possibility is thus as mentioned some kind of an empirical fact, or something ontological, just as the question of meaningfulness here is pragmatic as related to action, not linguistic or semantic, that is, not related to propositional truth. The crucial point here is the general logic of recognizing something as a transcendental condition, which cannot be ignored. For Apel such a condition is an action performed, but for Dussel such a condition is material in a more general sense than just an action.

IV. HABERMAS' VERSION IS DIFFERENT SINCE HE RECOGNIZES THE IMPORTANCE OF MATTER FOR ETHICS

Part of the idea we can get from Habermas' version of grounding in discourse ethics. His way of formulating discourse ethics draws explicitly on Apel's idea of transcendental pragmatism, and he accepts the basic idea of criticising by exposing performative contradictions and thus basing the validity of ideas on something not ideational. Habermas illustrates the logic of such a way of grounding with Descartes' *cogito*. By using such an illustration, however, Habermas actually displaces the idea as it thought of by Apel. And this displacement is not just a coincidence; it is symptomatic of the differences in perspective between Apel and Habermas, since for Habermas the conditions not to be ignored are precisely not only pragmatic, but also material. What I am thinking can be illustrated by reference to one of the paradoxes of classical antiquity, namely the one where a Cretian put forward a proposition stating that all Cretians are liars. In this case the propositional content of the material condition is in contradiction with the proposition expressed. This is not just a point concerning the speech action, but something material concerning values, namely the culture on Crete.

That Habermas finds the focus on matter much more relevant for discourse ethics than Apel is already indicated in the idea of another principle, namely the 'universalisation ground sentence' U. According to U a norm is only valid, if the consequences and side effects of its being followed can be accepted by everybody affected by that activity (Habermas 1983, 75-76). According to Habermas the U-sentence states a condition, which is basic to discourse ethics. U is called a "bridge principle", which is considered on par with the principle of induction, and that is because it brings us from our private

material considerations into ethics proper. It is by following the universalization rule that private interests can be transformed to values, and one can pass from unethical behaviour to ethical actions in Habermas' sense.

U is clearly concerned with material matters, and actually I think that *Faktizität und Geltung* is best considered as a full development of the principle stated in U. *Faktizität und Geltung* formulates the rules to be followed to realize a society, where discourse ethics proper, the abovementioned principle of D can rule our normativity. I will leave the attempt to form a detailed argument for this to some other time, but here it is still worth remembering that U stand for 'universalization', not 'universality'. It refers to a process that aims at bringing individual interest to the levels of ethical values, not a criterion for validity. Universality as criteria is relevant, when we discuss the discourse ethical principle, D, which is formal in that sense that it demands acceptability by all those possibly affected as participants in a practical discourse. And it is that principle, D, not U, which according to Habermas (1983, 103-4) express the moral core of discourse ethics.

U specifies norms that one must accept living in a society of people pursuing different interests. The specification of the condition for acceptance of such norms, however, it is not deontological, but consequential and material. And I think this must mean that we are simply talking of the political-economical conditions for realizing the discourse ethics. U concerns the universal acceptance of the material consequences of each individual in pursuit of his or her happiness, that is, of economical and political freedom. Habermas wants the material condition (U) realized universally as a condition of ethics, but that is ethics in terms of the good life, and thus a matter, which can be dealt with in practical politics, whereas ethics as morality is to be secured by the criterion expressed in D, which is formal and procedural, i.e. concerned with moral norms.

What is at stake is the level of political-economical freedom acceptable in a society, and in this aspect Habermas is much more demanding than for instance Rawls. Rawls (1999, 65, §13) allows for a political-economical inequality, if it can be argued to be to the benefit of the least advantaged. Habermas simply demands that there should be universal acceptance of norms and that leaves the question of inequality up to the verdict found universally acceptable in discourse. And we should remind ourselves that universality in the classical Kantian sense is expressed by D, not U, in spite of the latter's name. U states that we have to live with the fact of each other as following different life plans. What is expressed is thus a social liberal precondition accepting each individual life as ideally a realization of an individual plan, just as we can find it by John Stuart Mill (1961[1859], 304-5, § 3), although Habermas' version (1995, 114-15) is more deontological than what is expressed by Rawls' difference principle. To Habermas matter thus concerns the political condition for ethics. As he mentions in passing, discourse ethics presupposes highly rationalized life forms (Habermas 1988 [1981], 119), and for Habermas (1981, vol. 1, 205-8) rationalisation must always be understood in continuation of Max Webers theory of western capitalist modernity. Discourse ethics thus presupposes the life forms, which we are in the process of developing in the most affluent parts of the world.

I have paused a little on this subject because the understanding of the U-sentence has been a matter of some controversy among readers of Habermas (cf. e.g. Finlayson 2000, Langlois 2001, Abizadeh 2005, and Lumer 1997). Apel (1998, 733-35) argues that D is just a variation of U, and in Scandinavia some of Habermas' most dedicated readers seem to have passed rather quickly over the difference and relation between U and D (e.g. Larsen 2005, 143, 209 and Glebe-Møller 1996b, 15, 18). Some have even argued that in some contexts the U-sentence is identical to the D-principle (e.g. Bordum 2001, 30; Glebe-Møller 1996a, 86; Glebe-Møller 1996b, 21, and Eriksen & Weigård 1999, 214; 2002, 239).⁴ Dussel, however, does not fall into that trap. In his analysis of Habermas' contribution to discourse ethics Dussel (1998a, 185) recognizes that Habermas characterises the fundamental proposition of universalization, the U-sentence, in terms of matter, and that Habermas by formulating U is arguing for the importance of material conditions for ethics and morality.

V. ETHICS MUST BE DEVELOPED IN TERMS OF MATTER, BUT STILL AS UNIVERSAL AND CRITICAL

When Dussel criticises discourse ethics for being only focussed on formality and ignoring matter, then it is primarily Apel, who is the target. Dussel recognizes that Habermas includes matter in his version of discourse ethics, but still he is not satisfied. And the reason I think is found in Dussel's understanding of matter and materiality in relation to ethics.

In order to get a proper understanding of what Dussel is up to with his ethics of liberation, it is important to underline that Dussel adheres to the traditional philosophical ambition of universal validity. Dussel (1999, 116-18) thus recognizes the formal aspect of ethics as necessary, but the point is that it is not sufficient. He also explicitly recognizes the claim that philosophical thinking should be rational and scientific, adding just that science as such should be critical in a sense, he attributes to the original Critical Theory of Max Horkheimer. To be 'critical' as a scientist means for Dussel something very specific, namely that in the scientific research one should place oneself beside the victim (1998c, 191-93), that is, beside the poor, the hungry, the orphan, the woman, the Indian etc. To be critical thus means to take side for those, who are marginalized and excluded from the centres of the world system, i.e. from the modern capitalist society, no matter whether they are found in first, the second or the third world.

To stand beside the excluded, however, does not only mean that you take an ethical and political stand. It also means that one gets in the position to experience the limitations that practical philosophy has to overcome. It is the victims' point of view that reveals governing thoughts as ideological in the classical sense, namely as the thoughts

4] Some of the references and my understanding of these matters owe a lot to discussions with Natascha Schlottmann (2009).

of those in government. Dussel's focus on the victims does not mean that he is trying to construct a practical philosophy designed specially for the periphery, the exploited and the suppressed. His constant reference to those, who are excluded, is part of his critique of current schemes of thoughts as ideological; when practical philosophy wants to claim universal validity, letting those, who are excluded, be the centre of attention constitutes a perfect test case.

This way of referring to the victims of modern western capitalism, however, keeps within the formal approach of discourse ethics, and Apel has also recognized that the idea of the appeal of the victims does pose a challenge to discourse ethics, although he does not consider it a fundamental problem (see Pinero 2005, 32). Those excluded, however, are also interesting in relation to ethics in another sense, namely because they in a very material way – as victims – feel the consequences of the order of the world system. A focus at those excluded reminds us that practical philosophy must never forget the body as the material foundation of the consciousness, and that brings forth a sense of matter as precisely the material basis in the form of a body, a sense of matter, which Dussel (1998a, 130-32) demonstrates, was also recognized by Marx.

It is with the body that we enjoy, but it is also with the body we feel the pain. It is with the body that I starve, get tired, am worn out, suffer and eventually go down because of economical and political inequality. It is bodies that every day must give up, when thousands of people in the third world die because of abuse, starvation, or illness. Even in our first world middle class centres it is with the body we feel stressed as a consequence of the anxiety produced by local managerial, ideological and economical pressure. The focus on the corporal materiality of the victims thus reminds us of something, which the proponents of discourse ethics seems to have forgotten, namely, as Dussel (1998a, 252) repeatedly emphasizes, the basic material concern of ethics, which is the preservation and development of the life of every single human subject. That, however, is far beyond the sense in which Habermas thinks of materiality in formulating U, and in order to get Dussel's line of thought right, we must therefore pause to consider the concept of matter a little more closely.

VI. MATTER MUST BE UNDERSTOOD IN AT LEAST THREE SENSES

As indicated matter can be understood in more than one sense, and Dussel wants to employ a least three of them. In mainstream practical philosophy "matter" first of all can mean *subject matter* or *content*, and in this sense matter is opposed to form in formalistic conceptions of morals philosophy such as discourse ethics. In this tendency discourse ethics is simply continuing the liberal enlightenment ambition of Kant. As content matter can signify specific happiness, human rights, conceptions of human dignity or ideals of human interaction as for instance friendship, care etc. In the vocabulary of modern philosophical ethics matter concerns the 'thick' conceptions of the good life, whereas form focuses on 'thin' conceptions of norms, which aims at universal validity. For both Apel

(1987, 178) and Habermas (1983, 132) it is very important to underline that discourse ethics does not side with any particular or substantial ethical conceptions of the good life, and one therefore also express the distinction by saying that matter concerns the good, whereas form concerns the right. Habermas seems to accept the understanding of matter as content, and Dussel (Dussel 1998a, 187) does it explicitly. It was this understanding of matter that was presupposed in the critique of discourse ethics mentioned above in paragraph 3, namely that discourse ethics presupposes mutual recognition of personal dignity.

Second, Dussel of course recognizes the metaphysical understanding of matter as just something material, i.e. something physical. Dussel refers to a distinction in German between “material” with an “a” and “materiel” with an “e”, where the first one signifies *content* as opposed to *formality*, whereas the second signifies something physical as opposed to something mental or spiritual. One of the reasons why Dussel (1998a, 130-32) considers Marx important for the conceptual development of the ethics of liberation is precisely that he understands Marx’s materialism as an ethics of content, as saying something substantial about the good life. The point is here that the dialectical materialism of Engels, Stalin and generations of orthodox communists refers to matter in the second sense, the metaphysical or ontological sense (see Dussel 1998a, 621-22). So when Apel claims that Marxism as such must be given up, Dussel simply answers that Apel has overtaken a “standard Marxism” (2005a, 231), that is, a simplified and reductionist understanding of Marx (see Piñero 2005, 37). Instead, Marx should be read as an ethical critic of capitalism (Dussel 1998a, 315, n. 63.), and that makes Marx important for the ethics of liberation.

It is worth noting, however, that the ontological sense of matter can be distinguished in at least two senses, namely a mechanical understanding, which takes physics as the model, and an organic sense, which refers to biology. Matter can thus be considered both dead and alive, and Dussel will of course stick to the latter sense, whereas at least some orthodox communist would be mechanical. This, however, only becomes clear, when matter is understood in a third sense. As mentioned above Dussel underlines that ethics must contribute to the preservation and development of the life of every single human subject, and this he considers the basic material core of ethics. It is ethical content, substantial in the sense just mentioned, but it is also material in that way that the aim, i.e. its *telos*, is practically to preserve and develop the life of individual human beings. It is this specification of the matter of ethics that Dussel (2005b, 344-47) considers the universal material aspect of ethics. Within the discourse ethical framework Wellmer distinguishes clearly between intersubjective validity and objective truth. By employing this distinction Dussel can state that he is not just maintaining the ambition of formal universal validity, when it comes to morality; he also wants to argue for a materialistic ethics, which is universally true (1998a, 202-5).

In his formulation of the U-sentence Habermas is acknowledging the ethical importance of matter as content as well as human practice, namely by pointing to how acceptance of norms should depend on the interests of those affected and on the effects

of human interaction. By further restricting the validity of discourse ethics to certain life forms, he also implicitly acknowledges the importance of matter in its practical aspect. Dussel, however, wants to make this more explicit, first of all by reminding of the economical sense of matter, that is, that human practice involves both production and reproduction, which both require material resources. Matter in its practical aspect therefore involves an understanding and critique of capitalist economy, which was what Marx gave us. This understanding of matter, however, presupposes valuing matter in an even more basic sense, namely in terms of preserving and developing the human life of every individual human being. It is precisely here that the third sense of matter emerges in its most complete form, namely when Dussel remind us that a human life is not just organic or biological, that is, it is not just ontological in the reductionist sense. Human life is specifically human, and that means a life, which must include the practice of politics, economy, and culture at large. This is the sense of matter that I presupposed in the interpretation of the classical paradox of the Cretians in paragraph 4. One can say that in this understanding of matter the two first senses merge, since to Dussel culture, history and economy is living matter bestowed with meaning, and in that sense matter can be referred to in a transcendental argument concerning the foundation of ethics.

VII. THE NEGLIGENCE OF MATTER IN DISCOURSE ETHICS IS IDEOLOGICAL

According to the critique of Dussel, discourse ethics then does not reflect sufficiently on the material aspect of ethics. As ethics it is lacking an understanding of matter and materiality in all of the senses mentioned, i.e. lacking understanding of the importance for ethics of the content, the living body and the human history, culture, and economy. In order to back up these claims, in addition to Levinas and Marx Dussel also refers to the teachers of Habermas, Horkheimer and Adorno, who Dussel (2008, 296-98) recognizes as having a good understanding of the conditions and realities of human suffering, that is, beyond mere bodily pain.

As indicated in the first sections, Dussel considers the omissions of Apel and Habermas in relation to ethics with respect to matter to be ideological. Overlooking such obvious aspects of ethics is only possible, because western middleclass in its daily life has put such a distance to material needs and suffering that they do not seem urgent. Dussel (1998a, 188) acknowledges to have learned a lot from the young Habermas, but as a privileged contributor to the ruling discourse the mature Habermas is no longer able to take the position as standing beside the victim. Even though the mature Habermas has a better understanding of the material aspect of ethics than Apel, he is not critical in Dussel's sense, which Dussel finds demonstrated by the lack of criticism of capitalism in discourse ethics (1998a, 200). In contrast Dussel makes clear that the ethics of liberation is basically critical by always standing beside the victims, and that it aims at a transformation of society; but he also makes clear that this transformation does not have to be revolutionary (1997, 22; 2005b, 340).

What is important for Dussel, however, is not just that Habermas' discourse ethics should be criticised morally, ethically and politically. Habermas' loss of marginality means that he has lost the privileged access to knowledge of the material conditions of the excluded, which he was still able to uphold in *Theorie und Praxis*. Habermas' negligence can thus be explained as a symptom of the material changes in his own personal life-world. To continue the Biblical metaphor mentioned above, one could say that over the years his eyes have become infected by small specks to such a degree that it has in all likelihood affected his way of experiencing and thinking about ethics, making it ideological in the sense mentioned in the first paragraphs.

Now, however, we are in a position to indicate that discourse ethics is ideological in the even stronger sense than the one mentioned above, namely by maintaining the focus of attention on non-material matters, in spite of the very material reality of the victims of global capitalism. What we see is an insistence on restricting the discussions on justice in philosophical ethics to universality and formal criteria for rightness,⁵ functions as the culture industry does for Adorno. Discourse ethics fills our heads with formal matters that appear to be the answer to our idealist aspirations, but in reality it just consumes the mental attention and corporal energy that we materially have at our disposal. Discourse ethics proclaims to be universally valid, but, as mentioned above, Habermas himself seems to admit that in reality it is only of a very limited validity, namely for those lucky enough to be participating in highly rationalized life-forms, and such life-forms are at a global scale the privilege of a very small minority. In sum, discourse ethics can make us aware of shortcomings within our own communities, and this is important for the continued development of our substantial values; but it also tends to make us in the centre negligent for the material sufferings in the periphery. And when we are the beneficiaries of the current global order, such negligence makes us even less inclined to fight for justice on behalf of the victims of this order, and then when we continue this intellectual strategy, we are not just negligent, but actually accomplices. Discourse ethics can thus be said to be ideological in the strong sense of being in the interest of those, who benefits from the existing distributional and political inequalities.

VIII. IT IS THE INTRUSION OF THE OTHER THAT MAKES ME SIDE WITH THE VICTIM IN ETHICS

Thus for Dussel the focus on matter plays an important part in the critique that reveals discourse ethics to be ideological in the very strong sense. Now we come to the last part, namely that the better understanding of matter also benefits the philosophical understanding of ethics and morality as such. What I am thinking at is Dussel's insistence on the corporality of the victims, which brings forth another aspect of ethics than thinking of matter just as condition or resource for normative assessment and action. It is one thing

5] A current example of this strategy can be seen in Forst 2009.

to be the subject, who performs an action; it is quite another thing to be object for an action performed. The ethical assessment of an action is different depending on the role you play, active or passive, actor or victim, or whether you take a first person or second person perspective; and the point is that this difference cannot be abolished just because one attempts to achieve impartiality by taking a third person perspective, i.e. by acting as just a witness to an action, taking the view from nowhere, or placing oneself behind the veil of ignorance, as it has been famously put.

As mainstream philosophical ethics discourse ethics presupposes freedom of action and is therefore on the active side of a relationship, and the perspective is first person, even though it is sometimes disguised behind third person ways of expression. Dussel too wants think as an active citizen, but his interpretation of doing critical science means that he constantly reminds us that there is always another side of the action; whenever there is a first person, there is a second person. The basic human relation is an I-thou relation, as it has just as famously been expressed by a religiously committed thinker like Martin Buber (1923), and such a relation can never be one of equality. In this aspect of ethics, however, Dussel (1998a, 255-58) takes mainstream sociology as his point of departure, namely in a critique of the classical analysis of Talcot Parsons of the relationship of double contingency between *ego* and *alter ego*, which has recently been given a renewed prominence by the system theory of Niklas Luhmann. The point is here precisely that even though Luhmann quite clearly can point to the strained relation between the individual and the social system, neither the first person perspective of strategically calculating what the other will do, nor the objectified third person perspective can account for "the intrusion of the alterity of the Other" (257), i.e. the emergence of the Other as another autonomous subject critical of the social system.

According to Dussel, who in this aspect follows Levinas, ethics means that I have the responsibility to stand on the side of the victims, and that again means that I must take the perspective of those, who are the passive recipients of the consequences of my own actions. When I do something, I must always try to imagine the consequences as experienced from your perspective. In the vocabulary of Levinas Dussel presupposes that a 'totality' always has an 'exteriority', i.e. that any kind of orderly action always will imply victims, and being ethical means that one must stand up for such victims. Such an ethical perspective, however, is very difficult to reconcile with the active perspective of ethics aiming at political institutions, which are supposed to include everybody as equals, but which, in the eyes of Levinas and Dussel, nevertheless always will imply some elements of material constraint and force.

For Dussel ethics basically must take side for the victim. The material universal implied by ethics is to preserve and develop the life of every human subject, and since those affected, i.e. the victims, are always the majority in relations to the actors, taking ethics seriously in its universality implies standing up for the victim. So because the world will never be ideal, ethics implies a responsibility, similar to what Apel says of the ethics of responsibility. The difference is that for Apel this is supposed to be something only

temporary, which can be dealt with in part B as the application of ethics (Piñero 2005, 43), whereas Dussel considers this situation permanent and therefore wants to include it in part A, the foundation of ethics. So for Dussel inequality is ontological for the human way of being and ethics is precisely brought into being to deal with this condition, whereas Apel thinks that ethics logically and thus counterfactually must presuppose equality in order to be ethics at all.

Discourse ethics is, according to Apel (1988, 9), about the possibility to establish universal validity for moral norms, and that is supposed to happen through a process, where the real community of communication approaches the ideals of the ideal community of communication. What makes the case of Dussel different is that the insistence on standing beside the victims is considered part of ethics proper, not just a psychological motive or a material precondition for ethics. In mainstream philosophical ethics, including discourse ethics, ethics as such only begins, when we have made our claims and want to give reasons for them. The model is a group of free citizens gathered at the assembly, who are able to discuss and chose between alternatives. Such citizens want to choose the right alternative in a more general sense than what is good just for an individual by her- or himself. The discussion might of course take place within ones own consciousness as conscience reminds us about our responsibilities, but that does not change the model; remember the classical illustration of moral scruples as a devil that discusses with an angel.

IX. DUSSEL'S WAY OF DOING ETHICS IS A NECESSARY SUPPLEMENT TO MAINSTREAM ETHICS

As it should be clear by now, Dussel's way of doing ethics is different. For Dussel ethics has already begun, when we strive to stand beside the victim. This much more basic ethical drive can be considered the material or ontological basis for social relations as such. It is feelings such as pity, solidarity, love, or, when it is a little bit more articulated, conscience. The point to be made here is that this is the way of doing ethics that the victims are most likely to benefit from in practice, either simply from the ethical urge of those, who have the resources to be active, or by the appeal that Levinas has become famous for, but which can also be found in, for instance, Sartre' writings on ethics (1983[1947-48], 285-88; see Rendtorff 1993, 69-71). Apel recognizes the question of appeal as ethically relevant, but it is in quite a different sense (see Piñero 2005, 32). The basic conflict between Dussel and Apel is that even though they both are engaged ethically in doing philosophical ethics, and even though such an engagement is also acknowledged as philosophically relevant by mainstream philosophical ethics, including discourse ethics, to Dussel the ordinary way of being ethically engaged in doing ethics philosophically is simply not ethical enough.

The point is simply that there is a huge difference in the conceptions of ethics between, on the one side, mainstream philosophical ethics and discourse ethics, and, on the other, the ethics of Levinas and Dussel, which acknowledges an understanding of human relations mainly emphasized by various types of theology. Dussel is aware of these

differences in perspective, which to him reveals themselves as between an ethics, which as politics takes the perspective of those who can act, and an ethics, which primarily demonstrates the solidarity with those who cannot act, but can only wait to see what happens, when the others act. He actually makes this difference constitutive for both the structure of his ethics and his history of political philosophy (1998a, 207; 2007, 71-72), and as it is most often the case, the two sides of the distinction are not on the same level.⁶ The point is that what makes something right in a more general sense than just good for me, is precisely the ability of the actor to take the perspective of the other, of the victim of ones own actions. Ethics is practical, related to things that could be different, holding actions, which are goals in themselves in high esteem, and this means that ethics must have some consideration concerning matter. So no matter whether one argues transcendently, ontologically, or materially, the ethics of discourse and most other mainstream philosophical ethics presuppose the ethics of the kind that Dussel presents us with. In that sense Dussel is right and gains the upper hand.

This upper hand, however, still is not conclusive. Realizing the material basis of ethics does not offer much help to decide, when we are in the active role. It only gives us some universal, but still minimal constraints on the alternatives, between which we can chose. It helps us rule out unacceptable alternatives; it does not seem equally convincing, when we want to find the best alternative among those acceptable. So we must also go further, changing the victim's perspective into the citizen's – and this, I think, is actually what has happened with Dussel over the years. Apel has not changed his position during the dialogue, whereas Dussel has taken the opportunity to develop and refine his position (see Piñero 2005, 44). Dussel has developed his idea of ethics from just having the critical perspective based on the sufferings of a responsible, but passive human being in accordance with a traditional Christian and Jewish perspective to include also the active political perspective, which was the commonsensical precondition of the citizens of Athens in classical antiquity. Dussel has simply changed the perceptive of pity into an ambition to change the world and minimize the reason for pity. Greek ethics allows giving reasons for normatively guiding human life within a larger community, not just to protest against injustice, which is beyond human influence.

By his ethical critique of ideology Dussel reminds us that in the centres of the centre of the world system, where we live – that is, where we live together with Apel and Habermas – we are only rarely confronted materially with the perspective of the victim, the starving, the orphan, the widow etc. But these are perspectives that Dussel is confronted with every day, when transporting himself between his private home and his university department in Mexico City. Dussel is right in reminding us that universality in ethics must never just be considered in terms of formal validity; when it comes to the truth of the matter,

6] Although this point should be a commonplace after having been made forcefully by Derrida in several writings (e.g. 1968, 128), when writing philosophy in English it is still worth mentioning both the point itself and the reference. In the English speaking world the understanding of Derrida is often polemical and based on stereotypes (see Critchley 1998, 6-9).

universality has a material aspect in another sense, namely the human life of every actual individual. Ethics must never forget this, not even in its most philosophical or political arguments.⁷

ASO@dpu.dk

REFERENCES

- Abizadeh, A. 2005. In Defence of the Universalization Principle in Discourse Ethics. *Philosophical Forum* 36(2): 193-211.
- Addelson, K. P. 1994. *Moral Passages*. London: Routledge.
- Albert, Hans. 1968. *Traktat über kritischen Vernunft*. Tübingen: J.C.B Mohr.
- Apel, K.-O. 1976. Das Problem der philosophischen Letzbegründung in Lichte einer transzendentalen Sprachpragmatik. In *Sprache und Erkenntnis*, ed. B. Kanitscheider, 55-82. Innsbruck.
- . 1987. Der postkantische Universalismus in der Ethik im Lichte seiner aktuellen Mißverständnisse. In Apel1990[1988], 154-78.
- . 1988. Vorwort. In Apel1990[1988], 7-14.
- . 1990[1988]. *Diskurs und Verantwortung*. Frankfurt a.M.: Suhrkamp.
- . 1998. *Auseinandersetzungen in Erprobung des transzendental-pragmatischen Ansatzes*. Frankfurt a.M.: Suhrkamp.
- Apel, K.-O. and E. Dussel. 2005. *Ética del discurso y ética de la liberación*. Madrid: Trotta.
- Bordum, A. 2001. *Diskursetik og den positive selvreference*. Frederikberg: Samfundslitteratur.
- Buber, M. 2003[1923]. I and Thou (excerpts). In *The Continental Ethics Reader*, eds. M. Calarco and P. Atterton, 74-87. New York & London: Routledge.
- Critchley, S. 1998. Introduction. In *A Companion to Continental Philosophy*, eds. S. Critchley and W. R. Schroeder. Oxford: Blackwell.
- Derrida, J. 1972[1968]. La pharmacie de Platon. In Derrida. *La dissemination*, 77-213. Paris: Seuil.
- Dussel, E. 1973-80. *Para una ética de la liberación latinoamericana*. T. I-V. Buenos Aires, Mexico y Bogota.
- . 1997. The architectonic of the ethics of liberation. On material ethics and formal morality. *Philosophy & Social Criticism* 23(3): 1-35 (a later version in Spanish is Dussel 2005b).
- . 1998a. *Ética de la liberación en la edad de la globalización y de la exclusión*, Madrid: Trotta.
- . 1998b. Globalization and the Victims of Exclusion: From a Liberation Ethics Perspective. *The Modern Schoolman* LXXV(2): 119-55.
- . 1998c. El programa científico de investigación de Karl Marx. In *Los retos de la Globalización*, ed. F. López Segrera, 185-217. Caracas : UNESCO/CRESALC.
- . 1999. Principios, mediaciones y el bien síntesis (de la ética del discurso a la ética de la liberación). *Signos Filosóficos* (México) 1(1): 115-38.
- . 2005a. Respuesta inicial a Karl-Otto Apel y Paul Ricœur. In Apel & Dussel 2005, 217-47.

7] Some very preliminary reflections for this article have, in Danish, been published as A. Sørensen. 2008. Enrique Dussel. Ideologikritik fra syden. *Det ny Clarté* 9: 24-30. Translated into English it has become A. Sørensen. 2009. Enrique Dussel: Critique of Ideology from the South. *Kult* 6: 148-60.

- . 2005b. Hacia una arquitectónica de la ética de la liberación. In Apel & Dussel 2005, 339-66.
- . 2007. *Política de la liberación. Historia mundial y crítica*, Madrid: Trotta.
- . 2008. Den kritiske teori og frigørelsesfilosofien. In Dussel. *Frigørelsesfilosofi*, København: Politisk Revy: 293-328.
- Eriksen, E. O. and J. Weigård. 1999. *Kommunikativ handling og deliberativt demokrati*. Bergen: Fakkbokforlaget (2003. *Kommunikativt Demokrati*. København: Hans Reitzel).
- Finlayson, G. 2000. Modernity and Morality in Habermas's Discourse Ethics. *Inquiry* 43: 319-40.
- Forst, R. 2009. Zwei Bilder der Gerechtigkeit. In *Sozialphilosophie und Kritik*, eds. R. Forst, M. Hartmann, R. Jaeggi und M. Saar. Frankfurt a.M.: Suhrkamp.
- Gadamer, H.-G. 1986[1960]. *Wahrheit und Methode*, Mohr Siebeck.
- Glebe-Møller, J. 1996a. *Jürgen Habermas*. København: Gyldendal.
- . 1996b. Indledning. In J. Habermas. *Diskursetik*, Frederikberg: Det lille forlag.
- Habermas, J. 1971. *Theorie und Praxis*. Frankfurt a.M.: Suhrkamp.
- . 1988[1981]. *Theorie des kommunikativen Handelns*. Vol. 1-2. Frankfurt a.M.: Suhrkamp.
- . 1983. *Moralbewußtsein und kommunikatives Handeln*. Frankfurt a.M.: Suhrkamp.
- . 1992[1991]. *Erläuterungen zur Diskursethik*. Frankfurt a.M.: Suhrkamp.
- . 1995. Reconciliation through the Public Use of Reason: Remarks on John Rawls's Political Liberalism. *The Journal of Philosophy* XCII(3): 109-31.
- Hegel, G. W. F. 1952[1807]. *Phänomenologie des Geistes*, ed. Hoffmeister. Hamburg: Felix Meiner Verlag.
- Langlois, L. 2001. Discours Moral, Discours Éthique Et Situation Pratique: Remarques Sur La 'Diskursethik' De J. Habermas. *Symposium: Canadian Journal of Continental Philosophy (Revue canadienne de philosophie continentale)* 5(1): 67-82.
- Larsen, Ø. 2005. *Den samfundsetiske udfordring*, København: Hans Reitzels forlag.
- Lumer, C. 1997. Habermas' Diskursethik. *Zeitschrift für philosophische Forschung* 51(1): 42-64.
- Marx, K. and F. Engels. 1969[1845-46]. *Die deutsche Ideologie*. In *Marx Engels Werke (MEW)*, Bd. 3. Berlin: Dietz Verlag.
- Mill, J. S. 1961[1859]. *On Liberty*. In *Mill. The Essential Works of John Stuart Mill*, ed. M. Lerner, 253-360. New York: Bantam.
- Nogueira, J. V. 1992. Revolución Rusa y leninismo. In *Historia de la Teoría Política, 4. Historia, progreso y emancipación*, ed. F. Vallespín, 170-211. Madrid: Alianza.
- Piñero, M. A. H. 2005. Introducción: El debate entre ética del discurso y ética de la liberación. In Apel & Dussel 2005, 29-44.
- Rawls, J. 1999. *A Theory of Justice*. Cambridge, Mass.: Harvard University Press.
- Rendtorff, J. D. 1993. *Frihedoetik i Jean-Paul Sartres filosofi*. Aalborg: Nordisk Sommeruniversitet.
- Sartre, J.-P. 1983[1947-48]. *Cahiers pour une morale*, Paris: Gallimard.
- Schlottmann, N. 2009. *Deliberativt demokrati – en undersøgelse af dannelsens normative grundlag*. København: Danmarks Pædagogiske Universitetsskole.

Book Reviews

Amartya Sen, The Idea of Justice, Belknap Press, 2009, Pp. 496, ISBN: 9780674036130

It is a mark of the greatness of John Rawls' *A Theory of Justice* that, after nearly 40 years, it still retains its position as the one work of post-war political theory which must be contended with. One might say of Amartya Sen's new book what Jürgen Habermas, in his own debate with Rawls, once said: namely, that any disagreement between the two is basically a family feud. The dispute, to be sure, is a significant one. It concerns the viability of an idealized theory of society such as Rawls has presented it in *A Theory of Justice*. I say viability since it is not really Sen's concern to argue about the philosophical foundations of Rawls' approach. Rather, Sen wishes to substitute a more practically oriented theory of justice, the capabilities approach, for Rawls' foundationalism about institutional structures and their relation to justice. Sen's main criticism is thus that Rawls' idealized theory excludes too much of what we should care about: the person's real position in the world, as well as parts of the world which are not included in Rawls' closed system of the state. In short, Sen is driven by the practical need to provide a theory which is truly universal in scope, but which is also able to deal with vexing questions of real world politics like persistent inequalities among people and relative lack of basic freedoms. This situates Sen's work at the intersection of political theory and social science.

Sen's book is in many ways an introduction to and summary of work done over the past 50 years. Readers who are already familiar with Sen will find here a compelling elaboration and systematization of the many topics that Sen has worked on: economics, philosophy, and, it seems, all of the social sciences which lie in between. New readers will be dazzled by the copious references not only to Sen's own work, but by the near encyclopedic knowledge Sen evinces of work done in the past 50 years in the social sciences and political philosophy. Sen peppers his pages with references to friends and colleagues, especially at Harvard, Oxford and Cambridge, who have influenced him and who, inevitably, have been influenced by him. His style in these references is chatty and warm, and further contributes to the impression that what we have in these pages is a disagreement only within liberalism, which, in different forms, Sen believes has carried the day. (The only prominent references to a non-liberal thinker that I could detect are to the late G. A. Cohen, whose own recent book takes issues with the central tenets of Rawls' theory.)

Sen has many philosophical heroes but chief among these is Adam Smith who, for Sen, himself an avowed child of the enlightenment, represents a sort of counter current to the Kantian enlightenment, which Sen sees Rawls embodying in his emphasis on abstraction. This counter-current includes thinkers like Mary Wollstonecraft, the Marquise de Condorcet, Jeremy Bentham, John Stuart Mill and even Karl Marx. What all of these thinkers have in common, according to Sen, is their comparative rather than foundationalist approach. Thus, Smith's concept of the *impartial spectator* is championed as a comparative version of Kant's foundational categorical imperative. The differ-

ence between the two is that the impartial spectator, while impartial, proceeds by comparison of different perspectives rather than by moral judgment concerning the correct position to take, as the Kantian categorical imperative does. Smith's view is meant to permit us to see other positions without judging them right or wrong, but rather giving the spectator an appreciation of the relative advantages and disadvantages of a particular comprehensive position.

Sen's thought, as he makes clear on several occasions, is also influenced by Indian philosophical thought. Here the distinction between *niti* and *nyaya* conceptions of justice is particularly important to him. The *niti* conception of justice is characterized by "organizational propriety and behavioral correctness", while the *nyaya* conception "stands for a comprehensive concept of realizing justice" (20). And with this, the central concern of theory has been named: Sen is concerned with realization of justice rather than the determination of its pure definition. Sen writes: "an approach to justice can be both entirely acceptable in theory and eminently useable in practice, even without its being able to identify the demands of perfectly just societies." (401) What Sen wants, in other words, is a theory that is both responsive to the needs of justice and responsive to real world problems.

For someone who has worked as extensively on real world problems, from the Bengal famine of 1943, which Sen witnessed first hand, to the problems of the 'missing' generation of women in Asia (due to the impact of inequality on life expectancy of children and newborns), Rawls' idealized theory of justice surely leaves something to be desired. The problem with a procedure like Rawls' original position is that, as Sen puts it, "there is still a large question about how the chosen institutions would work in a world in which everyone's actual behavior may or may not come fully into line with the identified reasonable behavior [stipulated by Rawls]." (68) The problem, however, is not only that people are not sufficiently rational, but also, that formulating the question of justice in the way Rawls does means that certain important features of justice will be neglected. These are, as I have noted, the agent's real position and her real possibilities.

Sen develops this theory out of the social choice model developed by Kenneth Arrow. While much social contract theory depends on an idealized conception of rational agents, the social choice model, as developed by Sen, relies on our actual and only somewhat rational ability to rank different outcomes based on our particular situation. Such ranking, for Sen, means that we can weigh outcomes that are not ideal against each other, regardless of how close to the ideal outcome they get. Social choice theory thus makes room for incompleteness.

Sen identifies several features of the social choice framework. (1) *Focus on the comparative, not just the transcendental*. For Sen this means that "a theory of justice must have something to say about the choices that are actually on offer". (2) *Recognition of the incapable plurality of competing principles*. This means that there will be different conceptions of freedom that can be employed in social choice theory and this, in turn, means that there will be different and potentially incompatible outcomes of social choice theory. (3) *Permissibility of partial resolution*. This means that the result of social choice may be tentatively incomplete (a work in progress) or that it may arrive at an impasse which is conceptually accounted for, though also open to further revision. (4) *Social choice*

theory reflects diversity of input which means that social choice can give us functional connection between individual rankings and priorities on the one hand and results relevant to social policy on the other. These too will be helpful in determining a closer approximation to justice. (5) Finally, social choice, even if it yields conclusions which conflict, tells us a great deal about current positions and thus should help us come up with better solutions than the ones we already have (106-111).

Sen argues that this framework can help us evaluate relative freedom and justice in terms of outcome as well as in terms of agency. Avowing that he shares Rawls' commitment to the priority of liberty (62), Sen argues that justice in actual human affairs is not simply a matter of cumulative outcome (what results) but also of comprehensive outcome (what results and how it is brought about), as in Rawls' proceduralism. Sen thus sides with Rawls against economists and utilitarians who tend to see human freedom merely as welfare. Sen's approach of freedom as capability holds that freedom is valuable for at least two reasons: freedom must provide *opportunity* to pursue our ends, and freedom must give us a *choice* about which opportunities to pursue (228). Sen calls this capability. The approach is pluralist in the sense that it "points to an *informational focus* in judging and comparing overall individual advantages" without specifying how information may be used (232). It does not specify any particular 'primary goods' as Rawls does. Its second decisive feature is that it "is inescapably concerned with a plurality of different features of our lives and concerns." (233) These are what Sen calls functionings. Sen writes: "The capability that we are concerned with is our ability to achieve various combinations of functionings that we can compare and judge against each other in terms of what we have reason to value." (233) Freedom as capability thus seeks to do justice to the diversity of human positions as well as to the diversity of human interests.

But how are these different interests to be coordinated? Sen's answer is: through public debate. For Sen the model of debate is essentially given by the comparative model of Smith's impartial spectator, as I have noted. The impartial spectator permits open and open-ended discussion, hence the opportunity for people to revise their views based on their changing positions and changing information. This, for Smith, is taken as a fairly unproblematic point.

It is difficult to do justice to such a synthetic work in a few pages, and perhaps more difficult to launch any substantive criticisms of it. However, philosophy being what it is, it is perhaps worth dwelling for a few moments at least on what Sen's book is and what it is not. As I said at the outset, Sen's book is not concerned with the grounding of its theory in any meta-ethical sense, that is, Sen does not seek to give a final justification of his pluralist approach to liberty nor to his claim that the impartial spectator is the best sort of model for deliberation. He does argue, however, that the impartial spectator can help us become more clear about the ways in which human agency could be more fully realized in terms of both functionings and capability. He has, however, also argued that Rawls' conception of the original position is idealized in a way that makes it unsuitable to tackle real world problems. This too I will not dispute. Sen's book is a work that seeks to form a bridge between philosophy and social policy in the sense that it argues for the methodology of the capability approach largely from a mixture of philosophical and

empirical observation but does not try to systematize these in a philosophical foundational way.

For all of Sen's criticism of Rawls there is one concept that Sen barely mentions: the concept of the *reflective equilibrium*. The reflective equilibrium, which in Rawls is meant to ground the whole theory of justice, is the idea that we can abstract from our own perspectives, check these against those of others and adjust our own views accordingly. The reflective equilibrium, in my view, is remarkably close to Sen's own view of public deliberation. When all is said and done, Sen's and Rawls' theories bare remarkable similarity to one another when it comes to their philosophical commitments about the role of rationality and human agency. What they share is a rejection of Kant's attempt to provide a metaphysical foundation for morality. This means that they have, from a philosophical point of view, more in common than Sen may think.

Stefan Bird-Pollan
University of Kentucky

Axel Gosseries and Lukas Meyer (edited by), Intergenerational Justice. Oxford: Oxford University Press, 2009. Pp. 419. ISBN: 9780199282951

Justice between generations is now a major preoccupation for many human sciences, especially for political and moral philosophy. This development is partly a result of the complexity of the issue, about which numerous debates offer new and exciting challenges, such as discussions on what we owe to people who do not yet exist. But this phenomenon is mostly due to the social consequences of the question. Intergenerational justice leads us to think about the stability and sustainability of retirement plans, about environment damage, etc. It urges us to conciliate the individual interests or needs of current, future and even past generations. With their book *Intergenerational Justice*, Axel Gosseries and Lukas Meyer provide the reader with an exhaustive and sustained overview of these questions, thanks to the insight of many specialists. In the first part of the book, the authors try to accommodate different theoretical approaches, in particular in the face of specific challenges arising in intergenerational issues. In the second part, the other contributors to the volume deal with applicative problems.

The contributions gathered in *Intergenerational Justice* present several lines of thought and the perspectives are sometimes substantially different. However, it is possible to distinguish an internal logic and to underline grounds for disagreement: each author relies implicitly on an account of why individuals from current generations should take into account future generations. At the same time, all articles deal with theoretical challenges specifically related to intergenerational issues. From these two perspectives, Rawls's influence seems to be predominant. *Intergenerational Justice* provides different interpretations of this Rawlsian approach, notably through an important debate between egalitarianism and sufficientarianism. The latter interpretation seems to prevail. We are going to try to understand whether this prevalence is justified.

Intergenerational Justice gives the reader an opportunity to identify the main debates, in particular the discussion about the reason why current generations should act for future ones. Some authors consider that individual interests are good and sufficient

reasons to act. It is the case of contractualist theories defended by Gardin (chapter 3), and reciprocity-based theories, introduced by Gosseries (chapter 4), even if the latter considers that generations must also avoid “free-riding”: a generation is unfair toward future generations if it consumes all the resources it has inherited from preceding ones. Birnbacher defends a similar conclusion in his reflection about motivational assumption, when he takes into account the absence of future generations in individuals’ behavior (chapter 10). According to these theories, an individual would act on behalf of future generations only because of his personal interest or his offspring’s. Others show however that personal motivations toward future generations are of a moral order. For example, Bertram invites us to make use of the Marxist concept of *exploitation* in order to know whether future generations could ask for compensation from the current one (chapter 5). For instance, a generation which made efforts to improve future one conditions of life without back reward, for itself or next generation, could ask for compensation. According to Thompson, in the context of a “weak communitarianism”, the motivation lies in the respect we owe to our ancestors: people must cultivate and promote their predecessors’ efforts and values (chapter 1). In a libertarian perspective, Steiner and Vallentyne defend the principle of equal opportunity between individuals in an intergenerational context (chapter 2). According to them, current generations do not have an obligation to save, but they have a duty not to limit the benefits that future generations could derive from natural resources as well as their capacity of ownership.

Finally, authors who adopt a Rawlsian perspective consider that current generations should make it possible for future generations to live in good conditions through a “just saving principle”: in a hypothetical situation, participants decide what they should save for the next generations (1971, *A Theory of Justice*, Oxford: Oxford University Press. See especially § 44, “The Problem of Justice between Generations”, 251 – 58). But the correct interpretation of Rawls is a controversial matter, in particular between egalitarianism and sufficientarianism, from the difficulty this perspective encounters when trying to cope with intergenerational justice. John Rawls applied to it his theory in order to determine the principles of justice between generations: under a “veil of ignorance”, participants do not know their “real” situation – but they do know that they belong to the same generation. Since the first principle of justice grants equal liberties for all, the participants must define a just distribution principle between generations, called the just saving principle: a trade-off between consumption and saving which allows each individual, whatever his generation, to develop and pursue his conception of the good life.

According to Rawls, two stages must be distinguished. In the first one, the “accumulation stage”, participants must save enough capital to build institutions that warrant equal liberties for members of future generations and respect the first principle. In the second one, the “steady state stage”, institutions already exist and the participants have mainly to preserve them. Therefore, the first stage benefits from a “lexical” priority: before determining a just distribution between individuals from different generations, it is required to insure equal liberty for all. The egalitarian interpretation, represented here by Attas (chapter 7), supports a fair distribution of living conditions, based on a reflection about the correct level of savings from current to future generations. On the

other hand, the sufficientarian interpretation, embodied by Meyer, Roser (chapter 8) and Wolf (chapter 13), requires the establishment of minimum standards below which living conditions are no longer considered as decent. The latter establish equivalence between basic institutions advocated by Rawls in his reflections on intergenerational justice, and his concept of *basic needs* as it appears in *Political Liberalism*. Rawls had suggested that we should add to his theory of justice a principle requiring that all individuals could dispose of sufficient or decent resources or life conditions (1993, *Political Liberalism*, New York: Columbia University Press).

Like Rawls's, all theories of justice have to adjust their principles to cope with the specific challenges raised by intergenerational justice issues. The main difficulties come from the fact that future generations do not yet exist. The peculiar statute of people who do not exist makes it difficult to impose on members of current generations obligations on behalf of members of future ones. How could we grant rights to people who do not exist? Furthermore, how could we harm people who may not exist at all, and whose existence depends on our choices? These paradoxes, pointed out in particular by Derek Parfit (1984, *Reasons and Persons*, Oxford: Clarendon Press), called respectively the "non-existence" and the "non-identity" challenges, reveal the limits of some theories about intergenerational justice issues. Attas considers that egalitarian principles of justice can be applied as Rawls did to such a context: by virtue of a Kantian universalization principle, the duties that members of current generations have towards future individuals is almost similar to the duty they have towards their contemporaries. A Rawlsian original contract allows us to determine a level of obligations since a just distribution principle chosen by participants from the same generation must be followed by individuals from future ones: the principles of social justice established between contemporaries are also valid for other generations (chapter 7).

However, Heyd argues for the exact opposite position. According to him, the intergenerational situation does not coincide with David Hume's circumstances of justice: cooperation and obligations between individuals are required only if certain conditions are fulfilled (moderate scarcity in resources, same territory, etc.). Therefore, individuals from current generations do not have any obligations towards future ones. Moreover, individuals may not be rewarded for their efforts in favor of their successors: the risk of "chronological injustice" is serious. A solution could consist in building an intergenerational solidarity or cooperation through another institution, as family. A mother, or a father, has special obligations toward her son. The obligations of members of the current generation would depend indeed on the strength of the relation. Birnbacher's article is mostly an attempt to solve the non-existence challenge by underlining the fact that the effective absence of future generations discourages the current ones to take them into account. Nevertheless, the author considers that a chain of cooperation may stem from indirect causes: a selfish behavior could have indirect consequences benefiting the next generations. A chain of cooperation and family links promote some kind of reciprocity between individuals from different generations: each individual, whatever his generation, equally gives and receives. But such a proposition does not resolve all difficulties specific to intergenerational context. Gosseries puts forward a discussion about reciprocity-based theories to cope with the "population challenge". According

to him, there are inconsistencies when we consider demographic fluctuations between generations: it seems impossible to respect the reciprocity condition.

The problems that the egalitarian approach meets, despite accommodations, are one of the reasons why sufficientarianism seems to prevail on this issue. Meyer and Roser's interpretation of the Rawlsian perspective provides consistent discussions about chronological injustice. Thanks to a minimum threshold establishment, each individual is insured to benefit from decent conditions of life. Sufficientarianism seems also to be able to solve other difficulties egalitarian perspective cannot deal with. For instance, the establishment of a decent threshold is independent of the identity of the persons, as well as of the future size of the population. Wolf argues as well for a sufficientarianist approach based on Rawlsian basic needs through a reflection about the long-term consequences of climate change. Climate degradation could become an obstacle to autonomy. In such a case, it is required to insure decent conditions of life for all, so that each person is able to develop her own conception of the good life. Bertram argues also for a sufficientarianist solution but with a different basis, namely Sen's concept of *capabilities*: all individuals need to get decent conditions of life – good health, access to drinking water, access to education, etc. – to enjoy a real liberty (1992, *Inequality reexamined*, Oxford: Clarendon Press). According to Wolf and Bertram, sufficientarianist theories can be used in intergenerational context.

Intergenerational Justice aims at better understanding general issues and specific theoretical challenges, through a wide range of contributions. By their rank and importance, the theories inspired by an interpretation of Rawls's work seem to be the most influent. The book illustrates the important debate between two such interpretations: egalitarianism and sufficientarianism. However, egalitarian perspectives are not able to deal with specific paradoxes. The latter approach appears to be the most adequate for intergenerational justice issues since it warrants all generations decent living conditions. But such an approach encounters with internal debates and difficulties that must be resolved. First of all, there are *a minima* two models, inspired by Sen and Rawls. Is it required to choose between them whereas *capabilities* concept appears so closed to basic needs principle? It is also necessary to define what a decent life means, and to consider whether such a definition will apply to all individuals in each generation. One difficulty comes indeed from the radical inequality of life conditions around the world: for instance, in developed countries, the access to drinking water is not anymore a problem, but it is a crucial issue in the developing world. Therefore, whether we assume that decent life conditions mean that human being must get capacity to enjoy autonomy, obstacles to autonomy vary from a region to the other. Should we establish the same threshold for all, or is it required, for instance, to impose a priority for the world's poorer? Whereas we argue for sufficientarian perspective in intergenerational justice issues, it seems crucial to keep in mind inequalities of life conditions in current generations around the world.

Cédric Rio
University of Poitiers

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 Independenței, 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.ub-filosofie.ro>). CSRB is located in the Faculty of Philosophy, Splaiul Independenței 204, Sector 6, postcode 060024, Bucharest. For more information see <http://www.csrc.ro/EN/home>

ISSN 2065-7285

EISSN 2065-8958

© 2010 by *Public Reason*

ARTICLES

Anti-paternalism and Invalidation of Reasons
Kalle Grill (Uppsala University)

Modus Vivendi, Consensus, and (Realist) Liberal Legitimacy
Enzo Rossi (Social Ethics Research Group, University of Wales, Newport)

“Scales of Justice” and the Challenges of Global Governmentality
Ina Kerner (Humboldt University of Berlin)

Rawlsian Compromises in Peacebuilding? Response to Agafonow
Endre Begby (Centre for the Study of Mind in Nature, University of Oslo)

The Basis of Universal Liberal Principles in Nussbaum’s Political Philosophy
Matthias Katzer (University of Siegen)

Religious Toleration and Public Funding for Abortions: a Problem with
Christopher Eberle’s Standard of “Conscientious Engagement”
Michael Harbour (Harvard Law School)

Dussel’s Critique of Discourse Ethics as Critique of Ideology
Asger Sørensen (University of Aarhus, School of Education)

BOOK REVIEWS

Amartya Sen, *The Idea of Justice*
Reviewed by Stefan Bird-Pollan

Axel Gosseries and Lukas Meyer (edited by), *Intergenerational Justice*
Reviewed by Cédric Rio

<http://publicreason.ro>

ISSN 2065-7285
EISSN 2065-8958