PUBLIC REASON

Journal of Political and Moral Philosophy

Volume 8, Number 1-2, 2016

PUBLIC REASON

Journal of Political and Moral Philosophy

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

Public Reason is available online at **http://publicreason.ro**
ISSN 2065-7285
EISSN 2065-8958

PUBLIC REASON

Journal of Political and Moral Philosophy

Vol. 8, No. 1-2, 2016

ARTICLES

BOOK REVIEWS

Public Reason 8 (1-2): 3-13

Climate Change-The Hardest Moral Challenge?

Ingmar Persson
University of Gothenburg & Oxford Uehiro Centre for Practical Ethics

Abstract: This paper explores why it is so hard for us to do what we morally ought to do to mitigate anthropogenic climate change by reducing our carbon dioxide, CO_2, emissions. It distinguishes between two sources of this difficulty: (i) factors which make us underrate the harm that we *individually* cause when we perform our everyday CO_2 emitting acts and, thus, the wrongness of these acts, and (ii) factors which make it difficult for us to *cooperate* to the extent necessary to mitigate effectively harmful climate change by reducing our everyday CO_2 emitting acts. Under (i) are listed such factors as the temporal remoteness of climate harm, the fact that the causal connections between our acts and this harm are elusive, that countless agents together cause harm which is diffused widely over countless, anonymous victims, by acts routinely done. As regards (ii), a comparison with the problems of cooperation in the well-known tragedy of the commons is natural, but it is here argued that the problem of reducing our CO_2 emissions is disanalogous in several respects which make it harder: the world's nations differ enormously in respect of level of welfare, their record of past emissions, and the degree of exposure to climate harm; additionally, it is harder to survey compliance and apply sanctions to those who defect from agreements, in particular as future generations who have not consented to these agreements are involved. Together these factors make up a good case for saying that the problem of ameliorating anthropogenic climate change by reduction of our CO_2 emissions is the hardest moral problem humanity is facing.

Key words: climate change, CO_2 emissions, cooperation, the commons.

Tony Leiserowitz, of the Yale Project on Climate Change Communication, has said of the problem of counteracting anthropogenic climate change: "You almost couldn't design a problem that is a worse fit with our underlying psychology", and Daniel Gilbert, professor of psychology at Harvard, joins in: "A psychologist could barely dream up a better scenario for paralysis" (Marshall 2014, 91). In this paper, I will try detail factors that buttress the pessimistic diagnosis that the problem of mitigating anthropogenic climate is the hardest moral problem that humanity faces at present.[1] I don't mean that it's hard to come to a reasonable agreement about what we morally ought to do, as it can be when we have to make trade-offs between different kinds of value which is the case when, for instance, one act produces more well-being overall, but another distributes a smaller amount of well-being more justly, or one act harms people less, but involves using some of them as means. In such cases, it can be impossible in practice to reach agreement about what is the right act. No, I have in mind situations in which we can reasonably agree about what we ought to do-in the case at hand, broadly speaking, such things as significantly reducing our emissions of carbon dioxide, CO_2-but it's hard to get *a sufficient number of us to act on* what we agree that we ought to do. The fact that a moral problem is the hardest in

1] Most, if not all, of these factors are also discussed in Ingmar Persson and Julian Savulescu. 2012. *Unfit for the Future: The Need for Moral Enhancement*, Oxford: Oxford University Press. There is also a discussion of similar factors at much greater length by Stephen Gardiner. 2011. *The Perfect Moral Storm*, Oxford: Oxford University Press.

this sense doesn't mean that it's the *most serious* moral problem that humanity is up against, although it is reasonable to take it that moral problems that are candidates for being the hardest moral problems must be serious moral problems. I think that, for instance, the problem of preventing that the birth-rate in Africa will be so high that its population will increase from 1.1 billion today to 4.1 billion in 2100, unless there is an accelerating infant mortality, is likely to be a more serious problem in the sense that it will cause more human suffering and more damage to the astounding African wildlife in the present century.

We can distinguish between two sources of the difficulty of alleviating harmful climate change by cutting back on our CO_2 emissions[2]: (i) factors which make us underrate the harm that we *individually* cause when we perform our everyday CO_2 emitting acts, such as driving our cars or flying and, thus, the wrongness of these acts, and (ii) factors which make it difficult for us to *cooperate* to the extent necessary to prevent the climate harm that we cause by our daily CO_2 emitting acts by cutting down on these acts. It goes without saying that the factors listed under (i) are bound to reappear under (ii), since if it's hard for us to feel that some of our acts are wrong, we aren't much motivated to cooperate to reduce their number. Therefore, I'll start with an inventory of the factors (i).

As a point of departure, consider a situation in which it's *flagrant*, or more or less as obvious as it can be, what harm an agent causes and, thus, what reason there is to think an act is wrong if there is nothing to justify the harm: I punch you hard in the face, without having any good reason for doing so, such as your posing a serious threat to me. What are the factors that make this such a flagrant case of causing harm and, thus, of acting wrongly that most of us are shocked if we witness it, and wouldn't dream of executing the act ourselves? By sorting out these factors, I believe we could get a grip on what characterizes the acts whose harmfulness and, consequently, wrongfulness are more elusive and, so, more prone to be overlooked or underestimated, namely the ones that exemplify factors that are at the opposite end or maximally distant from the first factors. So, what are the factors that contribute to making the harmfulness and wrongfulness of our acts flagrant or evident?[3]

(1) *Temporal proximity between the act and the harm*: the pain and damage to the victim's face occur immediately after the punch. This enables us automatically to associate the harm with the punch. If it instead takes a long time for the harm to occur after an act is done, such an association won't be set up automatically, and we'll feel less uncomfortable about performing the harm-causing act. For instance, if we were forced to kill someone with a poison, we would tempted to give the victim a poison which took a very long time to kill rather than one which kills instantaneously.

2] There are other ways of alleviating harmful climate change, e.g. preventing deforestation, but I'm here focussing on the reduction of CO_2 emissions.

3] I'll assume that the victims of harm are humans. Arguably, there are factors that make us tend to underrate the harm done to non-human animals relative to humans, but these won't be discussed here.

This can be explained in part by the fact that *we are biased towards the near future*: we are more concerned about good and bad events that occur in the near future than in the more distant future. That's why we're relieved when an unpleasant event is postponed, and disappointed when a pleasant event is. This relief or disappointment is out of proportion to a reduction of probability that the postponement usually brings along. To the extent that there is this lack of proportion, there is reason to think that this temporal bias is irrational.

Now the harm caused by our CO_2 emissions is temporally very remote. CO_2 can accumulate in the atmosphere for hundreds of years, blocking radiation of heat from the Earth's surface, but letting through sunlight, thereby eventually leading to a harmful increase of the global temperature. But this is a very slow-working process which may take centuries to produce its worst effects.

(2) *The victim(s) is (are) identifiable and concrete*, that is, identifiable not in the sense that the names of the victims are known to the agents, but known in the sense that the agents are able to point them out and see what they look like. It's a familiar fact that we feel most sympathy or compassion with individuals who suffer before our very eyes. This is much harder for us to bear than suffering that is merely verbally recounted to us, even if it be the suffering of many more individuals. There is a correlation between this factor and temporal proximity: if the harmful effect of an act we perform is temporally proximate to the act, its victim is often within eyesight of us, whereas if the harmful effect is temporally distant, this is often not the case. When the harm is temporally very remote as in the case of climate change, the victims harmed will normally be anonymous, that is, we won't be able to pick them out.

(3) *Concentration of causation of harm to a single agent*: the agent who is dealing the harmful punch is just me, no other agent is involved. Contrast this with situations in which there is a diffusion or division of the causing of harm over several agents. Such a plurality of agents may either act simultaneously-like oarsmen rowing a boat-or some might act subsequently to others, as when one agent sets fire to the victim's house and another locks the doors to prevent people to escape from it. Common sense conceives moral responsibility *as being heavily based on causation*, so when causation of harm is spread over several agents, the feeling is that each agent involved is morally responsible for less harm. Indeed, even if you disperse the causation of harm over several of *your own* acts rather than concentrate it to a single act of yours-e.g. destroy a lawn by crossing it daily over the period of a year rather than by one act-you'll feel less responsible for the harm you cause. Yet on reflection it seems absurd that we could evade responsibility by such a dispersal of causation, as will become clearer when we consider the next factor.

(4) *Concentration of harmful effects to a single victim* rather than diffusion of the same quantity of harm over several victims, with the result that each suffers merely a fraction of the harm caused by the agent. Derek Parfit's 'harmless torturers' illustrate such a diffusion of harm caused (1984, § 29): instead of causing a single victim excruciating pain by increasing a painful stimulus a thousand times, each torturer in a group of 1000 torturers increases this stimulus by one unit for a thousand victims, thereby causing only

an *imperceptible* difference for each of the victims. Such a diffusion makes each torturer feel that he's acting less wrongly than he would be had he increased the stimulation a thousand times for one victim, even though the 1000 torturers together cause as much harm as they would do had they each increased the painful stimulus a thousand times for a single victim.

The reason for this apparent reduction of wrongfulness and guilt is that, while we are capable of feeling adequate sympathy or compassion for a single victim, we aren't capable of feeling adequate sympathy or compassion for several victims in proportion to their number. So, the fact that each victim is feeling less suffering diminishes our sympathy for each of them, but the fact that their number increases doesn't augment it, or at least not by far stretch in proportion. Yet, if there are several agents acting in concert, diffusion of effects doesn't exclude that the total upshot is the same as it would be if each agent had individually caused serious harm, e.g. if each of the thousand torturers had increased the painful stimulation a thousand times for a single victim.

However, large-scale diffusion of both agency and effect is precisely what happens with respect to climate change: the innumerable CO_2 emitting acts of each of us have only an imperceptible effect on the climate, but because there is such a huge number of us the total effect is harmful to a lot of the global environment, as harmful as it could be if each of us had noticeably destroyed a certain minor part of the environment. For instance, your driving your car won't make any measurable difference to the global temperature, so, you may feel that you may drive your car without being guilty of any harm. Yet, if the world's 700-800 million cars are driven by drivers who feel the same, great harm will eventually be done to the global climate.

(5) *Perspicuity of the causal process*: the causal connection between a punch in the face and pain and facial injury is so perspicuous that even a young child can understand it (though a scientific account of it may be a complicated matter). Needless to say, how CO_2 emissions cause harmful climate changes is a much more complicated matter. It takes so much of science to understand how they cause global warming that this has only been understood rather recently, and most of humanity still lacks this understanding. Moreover, a more precise knowledge of what temperature increases it takes to cause certain harmful effects, such as a certain amount of progressive melting of the vast ice caps on Greenland and Antarctica and a consequent rise of sea levels is something that even expert climate scientists disagree about. When there is some unclarity about how an act causes harm, some doubt might seep in about whether it really does. Also, uncertainty invites wishful thinking to the effect that perhaps we won't cause any climatic harm even if we don't change our extravagant life-style. People in the fossil fuel industry will not be late to exploit these sentiments, and they have the economic means to exercise a strong influence on the mass media and politicians.

(6) *The harmful act is an act out of the ordinary*: it isn't an act that we perform regularly or routinely. Most of us don't go around punching people in the face regularly, and those of us who do probably don't feel bad about it! By contrast, many of us have driven our

cars daily for years and years, and got accustomed to the idea that there isn't anything wrong about that. The fact that we and others around us have got into the habit of doing something routinely and regarding it as permissible makes it hard for us to take to heart an intellectual realization that these acts involve so much harm that they are in fact wrong, and as a result abstain from them.

This intransigence is shown also, for instance, by the fact that many people who become convinced that meat-eating is wrong find it hard to quit because they've got so used to eating meat and regarding it as permissible, and most people around them do the same. Habit and conformism make us blind to the wrongness of status quo.

Along these dimensions, then, our CO_2 emitting acts are at the opposite end to acts like punches in the face: their harmfulness is discreet or unobtrusive rather than flagrant or evident and, thus, we're spontaneously inclined to ignore or underrate their harmfulness and, so, their wrongness. It's plausible to hypothesize that evolution has programmed us to adopt moral aversion towards such flagrantly harmful acts as punching people in the face because they are actions that have been elements of our behavioural repertoire throughout our history, and their consequences have been invariably the same. But the causation of harm by CO_2 emitting acts is a recent addition to this repertoire, since they presuppose advanced technology and a huge number of agents performing them together. Consequently, it isn't surprising that we have a hard time convincing us that they could be harmful to an extent that could make them wrong.

On reflection, however, it seems clear that all of the six factors are irrelevant to the harmfulness of an action. The only exception is the non-perspicuity or elusiveness of a causal link when this factor makes it rational to doubt that there *is* such a link, and this is no longer true with respect to the causal link between our CO_2 emissions and harmful climate change. Nevertheless, all six factors contribute to making us spontaneously inclined to disregard the harmfulness of our CO_2 emissions. Therefore, if these acts benefit us, even slightly-which they certainly do-we'll be reluctant to abstain from them.

Let's now turn to the second source of difficulties (ii). To prevent the harm that we are causing by our CO_2 emitting acts, it isn't enough that *some* of us abstain from these acts, we need a majority of us to agree to do so in order to ensure effectiveness. This is due to the diffusion of agency and effect over several agents and victims, that is, factors (3) and (4). It goes without saying that the six factors that make us individually disinclined to cut down on our CO_2 emitting behaviour also make it difficult to animate a sufficient number of us to cooperate effectively to cut down on this behaviour, but cooperation introduces additional complications, which I'll now explore.

There is a well-known cooperation problem called *the tragedy of the commons*. It's natural to take it as a point of departure for a discussion of the problem of cooperation to mitigate anthropogenic climate change by reducing our CO_2 emissions. The tragedy of the commons consists of the herdsmen of a village trying to agree on restrictions on the grazing of their cattle in order to avoid overgrazing of the commons, and subsequent starvation for the herdsmen and their families. There's a problem of establishing cooperation here since,

although each of the herders has a self-interested reason to cut down on the grazing of their own cattle as a means to preventing overgrazing-which will ultimately inflict starvation on them and their families-they're likely to have a stronger self-interested reason not to do so. They might hope that a sufficient number of the other herdsmen reduce the grazing of their cattle, and free-ride on this reduction without making any reduction themselves. This strategy has the additional advantage that in the event that others by and large decide not cut down, they haven't made any useless sacrifice of their own welfare. But, obviously, if all or most of them reason and behave in this way, the collective grazing won't be reduced sufficiently to avoid overgrazing and eventual starvation, which is bad for all of them. There are however significant disanalogies between this situation and the problem of reducing global CO_2 emissions which make the latter a more pernicious cooperation problem. I'll now survey these disanalogies.

(A) *Cooperation to reduce effectively CO_2 needs to be more or less world-wide, involving at least bigger nations which are significantly different from each other.* A global agreement is clearly harder to establish than an agreement in a village in which everyone knows everyone, and share the same ethnicity and culture. This sharing is something that facilitates the growth of some measure of altruistic concern and trust among the herders. By contrast, there are deep ethnic, cultural, and political differences between many of the biggest countries of the world, countries like the USA, China, India and Russia. Some of them also have long histories of war and conflict. As a result, there will be minimal fellow-feeling between them, and trust that any costly agreements will be kept.

These differences make it difficult for some nations to cooperate *in general*, but there are also differences between the world's nations which are relevant for cooperation about the reduction of CO_2 emissions specifically. Let's review these differences.

(B) *The immense differences between the world's nations as regards their level of welfare, or GDP, and their level of CO_2 emissions per capita*. In the tragedy of the commons, the herdsmen might be thought to be roughly equally well-off, have a roughly equal number of cattle whose grazing needs to be reduced, and have equally many dependents to feed. This makes it comparatively easy for them to agree on what's required of each and everyone: they should divide equally among themselves the cut-downs of the grazing necessary to attain sustainability. The enormous differences in welfare between the world's richest and poorest nations rule out such a simple solution with respect to combatting climate change. These welfare differences make it reasonable to demand that richer nations pay more for measures to reduce the future level of CO_2 in the atmosphere because of their greater ability to pay, and this is likely to generate disagreement about how much more they should pay, and in what ways they should make extra contributions. This is something that has surfaced in international negotiations.

A related problem is that the per capita rates of emissions of the big emission countries differ greatly, and this may be so even though the total amount of emissions by the countries may be more equal because the size of their populations differs. To illustrate, consider the two countries that emit most CO_2 in the world, China and the USA; they

must surely be included in any effective cooperation. The population of China is around four times as large as the population of the US, but the per capita emission of the US is almost three times higher than they are in China. It would of course be disastrous for the climate if China were to increase its per capita emissions to the present level of the US. But it would be exceedingly difficult to get the US to accept a Draconic cut to bring them down to the current level of China's per capita emissions. So, a compromise in between which is satisfactory to both parties must be found. Clearly, it will be hard to find such a compromise which effectively reduces the global emissions of CO_2. Generally speaking, the problem is that developing countries are liable to aspire to the same standard of living as the more developed countries, a standard which the latter will be reluctant to lower markedly.

(C) *The historic record of CO_2 emissions differs between the more and the less developed nations*. Again, this can be illustrated by a comparison between China and USA: since 1850 USA has emitted roughly three times as much of the CO_2 put by human activity in the atmosphere as China. It's arguable that this estimate is largely irrelevant to current negotiations because a lot of the emissions occurred before there was any reason to suspect that they were harmful; therefore, it might be contended, there's no *moral* responsibility for the harm they've caused. But this is likely to disputed because, as noted, our commonsensical conception of responsibility bases it heavily on causation. This causal conception of responsibility might motivate the Chinese to propose that, on the basis of their more modest historical record, they have a right to a per capita rate of emissions in the future that is somewhat higher than that of the US. Personally, I don't believe that this causally based conception of moral responsibility is defensible, but it's so firmly moored in commonsensical thinking that it'll be persuasive to many. This is a complicating factor that's missing in the tragedy of the commons, since whatever the conception of responsibility, the herdsmen will be equally responsible for the overgrazing on the assumption that their cattle stocks are roughly equal.

(D) *The degree to which different countries of the world are harmfully affected by anthropogenic climate change varies widely*. Some countries are likely to sustain devastating damages, while other countries may stand to gain rather than lose by expected climate changes. Great losers are low-lying countries like Bangladesh, the Netherlands, and South Sea Islands-which run a serious risk of being inundated by rising sea levels-and regions in Sahel, Australia and the south-west of USA which will probably be exposed to severe droughts and desertification. Geographic regions which may enjoy salutary effects are Greenland, Russia and Northern Europe, though some of them might get massive waves of climate refugees from other parts of the world, e.g. Africa and the Middle-East, at their doorstep. Obviously, the losers have much more of an incentive to implement a reduction of emissions of CO_2 than the winners. The latter are asked to make substantial sacrifices of welfare largely for the benefit of other nations, and this is clearly less motivating due to the narrow limits of human altruism, which is largely confined to near and dear, like families and friends. Again, this is a feature that is missing in the tragedy of the commons in which

the herdsmen are asked to make sacrifices for the good of a collective to which they and their families belong.

Further, it should be noticed that even in countries which are expected to be comparatively severely hit by global warming, the worst effect won't be suffered by the *present* generation, who is making decisions about climate policies, or perhaps even their children, but by generations further into the future. This is because climate change is such a slow process. Thus, even these decision-makers are asked to make sacrifices for people who are to a great extent beyond the range of their limited or parochial altruism. By contrast, even the herdsmen *themselves* could be assumed to suffer from a failure to cut down on grazing. The tragedy of the commons-like *the prisoners' dilemma*-is commonly understood to show how *self-interested* agents could end up doing something that doesn't issue in the best outcome for themselves because they aren't willing to make any sacrifices for the common good. Now, due to the bias towards the near we're relatively unconcerned about effects in the more remote future even when they affect ourselves-that is why, for instance, smokers find it difficult to quit their hazardous habit. Obviously, we're even less concerned about temporally remote effects if they affect others, especially if they aren't near and dear to us, which they won't be if those affected are unknown people in the distant future or in distant countries. In those cases, the bias towards the near future and our parochial altruism join forces.

Additionally, making sacrifices of our own welfare, or the welfare of near and dear ones, for the sake of the global climate involves, as remarked, the further discouragement that the contribution we individually can make for the common good of all beings on the planet by reducing our own emissions is imperceptible or negligible because it takes a countless number of emissions like ours to produce a harmful climate effect.

(E) *Controls of compliance are lacking with respect to global treaties to reduce CO_2 emissions*. It's unlikely that there will be an effective surveillance of whether countries over decades comply fully with treaties to reduce their CO_2 emissions they have entered into. And if they are found out to have defected, there will probably be no effective sanctions to apply. Such checks and sanctions are surely necessary for there to be a reasonable guarantee of compliance, since we can't expect people all over the world to have much altruistic concern for and trust in each other, for reasons recounted above-see (A) in particular. Accordingly, these considerations have caused worry at international meetings. By contrast, in the tragedy of the commons the group of herders is so small that they can be expected know each other personally, having lived together for a considerable time. Thus, they can realistically be thought to have developed some altruistic concern for and trust in each other. Also, remember that the good of the herders themselves and their families is part of the common good, though by reducing the grazing of their cattle, the herders forgo the very best outcome for themselves-the prevention of overgrazing without making any sacrifices-and risk the worst outcome: making sacrifices while so many of the other herders don't, so that there's still overgrazing. But the risk of free-riding or defection is diminished by the fact that the group of herders is so small that they can realistically

be imagined to be able to keep an eye on each other. Since they can also realistically be thought to be joined together by bonds of fellow-feeling, they are likely to be motivated to collaborate to punish defectors and free-riders.

(F) *The effectiveness of current compliance to international agreements to reduce CO_2 emissions relies on the compliance of future agents who aren't bound by the agreements.* Cooperation about reducing CO_2 emissions has to extend far into future in order to be effective in alleviating global warming. But future generations who haven't consented to agreements about CO_2 reductions could in virtue of this fact claim that they aren't bound by them. Thus, there is a risk that when future generations realize that their standard of living is going down because of the reductions of CO_2 emissions implemented by earlier generations-reductions which may benefit primarily even later generations-they will be prone to discontinue these reductions. This is especially so, since they may fear that even if they keep them up, the following generation won't because they will be subjected to even greater hardships, and they have still greater reason to fear that the generations succeeding them won't keep in line because they'll be subjected to yet greater hardships, and so on. Such a chain of growing incentives to defect seems fatal to the possibility of reaching viable agreements.

To sum up, not only are nations at present encouraged to 'pass the bill' to future generations because these can't 'retaliate'; they are also encouraged to do so because they can't trust that future generations even of their own nations, let alone other nations, stick to necessary cut-backs agreed on. If it's hard to trust that the governments of other nations will at present stick to agreements, it's much harder to trust that their future governments will continue to do so.

Let's take stock. I've reviewed six dimensions, (1)-(6), along which our CO_2 emitting acts are at the opposite pole to acts whose harmfulness is so flagrant or evident that it's hard to deny their wrongness in the absence of justifying factors. This means not only that we'll be spontaneously disinclined to abstain from these emissions; it also means that we're unlikely to give our votes in general elections to political parties that favour reductions of CO_2 emissions. The factors (A)-(F) boost the unlikelihood of citizens voting for such 'green' parties. The result will be that liberal democracies are unlikely to have governments that give priority to efforts to mitigate global warming by cutting down on their CO_2 emissions. The parties that gain and retain power in liberal democracies are more likely to give priority to issues of employment, education, health care, restrictions on immigration, etc which directly benefit their voters. Politicians risk very little by omissions to combat climate change, since it's most unlikely that there will be any climatic catastrophe that can be definitely put down to human emissions as long as these politicians are in office, or even alive. The realism of these speculations is borne out by the fact that no sufficiently effective action against climate change has hitherto been taken, even though the problem has been on the agenda of organizations like the United Nations for more than twenty years.

Contrast with the risk for terrorist attacks. A major terrorist attack in a Western democracy might have seemed improbable before 9/11, but after it was no longer

be difficult for politicians to sell anti-terrorist policies to their voters because their harmfulness is flagrant and has been demonstrated. It's also in the politicians' interest to propose such policies, since terrorist attacks could happen while they are in office, and this would have a devastating effect on their chances of being re-elected. Furthermore, in the case of terrorism the majority of the voters aren't disinclined to accept policies which come down hard on the culprits because in this case *they* aren't the culprits as they are in the case of anthropogenic climate change. The problem in the case of terrorism is rather that, since it's as a rule easier to harm than to benefit, and the possibility of creating great harm grows with an increasing availability of more powerful technology, there will be innumerable loopholes through which great harm could creep in. It's hard to close them all without unduly curtailing the freedom of ordinary citizens. But citizens in general are probably more readily moved to approve of such measures than restrictions on their CO_2 emissions which will lower their welfare because terrorist attacks are acts whose harmfulness is flagrant like punches in the face, not the discreet kind of harmfulness that slowly and imperceptibly sneaks in on them under the cover of everyday life.

Moreover, it's in the interest of some economically very resourceful players who profit from the use of fossil fuel-in particular oil companies-to block policies to place obstacles in its path. As remarked, the causal connections between CO_2 emissions and harmful climate changes are elusive. True, there is an impressive body of scientific evidence demonstrating the influence of these emissions on the climate, but more precise knowledge about what impact various levels of CO_2 will have on the global climate and human civilization is missing. This provides agents interested in downplaying the risk of anthropogenic climate harm, like representatives for the fossil fuel industry, with room to exaggerate our lack of knowledge about the climatic impact of our CO_2 emissions. And they have the economic means to influence politicians and the media. The fact that the anthropogenic change of the climate is such a slow process and that it is masked by natural climatic variations makes us prone to overlook or dismiss it. As we have seen, evolution has wired us up to be alarmed by harm which occurs flagrantly, as is the case with punches in the face. When its occurrence is discreet or unobtrusive as in the case of anthropogenic climate change, wishful thinking has time to enter and distort the facts so that we can continue to benefit from our usual CO_2 emitting acts without any feelings of guilt.

All in all, the circumstances listed under (i) and (ii) conspire to make the moral problem of effective cooperation to mitigate harmful climate change by reduction of our CO_2 emissions maximally difficult. Its difficulty stems both from factors which make it hard for us to feel that our individual CO_2 emitting acts, which in fact contribute to climate change, are harmful, and from the fact that effective reduction of the harm they produce necessitates such an extensive cooperation of agents so different from each other. The combination of these features is what makes me think that this is the practically hardest moral problem humankind faces. The practically hardest moral problem must involve cooperation-of a kind that is hardest to establish. And it must be hard on the individual level to convince yourself that you act wrongly. The problem of achieving a requisite

reduction of CO_2 emissions scores high along both of these variables. This isn't to say that it's a problem that it is impossible to solve, but the odds are bad because the difficulties are rooted both in our psychology and in general facts about the state of the world.

Suppose that someone, say at the time of the Kyoto Protocol in 1997, had accurately predicted how things would develop up to this day with respect to measures to reduce CO_2 emissions. Then I think that at the time this prediction would have been described as pessimistic: after all, it would predict that CO_2 emissions would continue to increase quite steeply. But it would be moderately rather than extremely pessimistic, since it wouldn't have predicted any climatic *catastrophe*. Accordingly, I believe it's reasonable to be moderately pessimistic about the course of anthropogenic climate change during the next twenty years or so because it must be judged probable that we'll continue to act in the future as we've done in the past, unless some significant change of attitude occurs. But it's hard to see what could bring about such a change of attitude, at least before we have reached a tipping-point at which further deterioration is inevitable.[4]

ingmar.persson@filosofi.gu.se

REFERENCES

Gardiner, Stephen. 2011. *The Perfect Moral Storm*, Oxford: Oxford University Press.
Marshall, George. 2014. *Don't Even Think about It*. London: Bloomsbury.
Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Clarendon Press.
Persson, Ingmar and Julian Savulescu. 2012. *Unfit for the Future: The Need for Moral Enhancement*, Oxford: Oxford University Press.

4] This paper was presented at a workshop in Bucharest in June, 2017. Thanks to all participants for helpful comments.

Public Reason 8 (1-2): 15-30

Coercion and the Subject Matter of Public Justification

James Boettcher
Saint Joseph's University

Abstract: Some public reason liberals identify coercive law as the subject matter of public justification, while others claim that the justification of coercion plays no role in motivating public justification requirements. Both of these views are mistaken. I argue that the subject matter of public justification is not coercion or coercive law but political decision-making about the basic institutional structure. At the same time, part of what makes a public justification principle necessary in the first place is the inherent coerciveness of a legally organized basic institutional structure. While most public reason liberals seem to presuppose that the meaning of "coercion" is sufficiently obvious so as not to warrant further analysis, my defense of the essay's main thesis explicitly draws on an account of coercion as a powerful agent's employment of enforceable constraints to determine the will of another agent.

Key words: coercion, liberalism, public justification, public reason, rule of law.

All public justification principles specify the conditions according to which something is justified to some public. I refer to this something – e.g., laws, fundamental laws, coercive laws, constitutional principles, or other institutional arrangements – as the *subject matter of public justification*. Public reason liberals disagree about both the subject matter of public justification and the significance of coercion for a theory of public justification. Some public reason liberals identify coercive law as the subject matter of public justification, while others claim that the justification of coercion plays no role in motivating public justification requirements. Both of these views are mistaken.

In what follows I argue that the subject matter of public justification is not coercion or coercive law but political decision-making about the basic institutional structure. At the same time, part of what makes a public justification principle necessary in the first place is the inherent coerciveness of a legally organized basic institutional structure. In short, justifying coercion is essential to public reason liberalism's project of public justification even though coercive law is not the subject matter of public justification. While most public reason liberals seem to presuppose that the meaning of "coercion" is sufficiently obvious so as not to warrant further analysis, my defense of the essay's main thesis explicitly draws on an account of coercion as a powerful agent's employment of enforceable constraints to determine the will of another agent.

Sections I and II explain how the subject matter of public justification is understood according to different conceptions of the public justification principle, namely, the more classically liberal asymmetric convergence model (which I criticize) and a Rawlsian inspired alternative to it (which I endorse). These sections also introduce an argument by Colin Bird that aims to separate public justification entirely from concerns about coercion. I challenge Bird's conclusion (section III. 1) and argue that the inherent coerciveness of the basic institutional structure is part of what makes political decision-making about that structure the appropriate subject matter of public justification (section III. 2). Section

IV identifies and responds to objections to this main argument, including the objection that it is the involuntariness of political society rather than the coerciveness of its basic structure that motivates a public justification principle. A subsequent investigation into the meaning of "coercion" in section V explains why the enforcement approach to coercion is a much better fit for a theory of public justification than the pressure approach familiar from the philosophical literature on coercion. Before concluding in section VII, I also discuss why the thesis of my paper does not depend on adopting a particularly narrow or broad view about the so-called scope of public reason, i.e., the set of political decisions to which requirements of public reason and public justification should apply (section VI).

I. THE ASYMMETRIC CONVERGENCE MODEL AND COERCION

Various formulations of the principle of public justification refer to coercive law as the main subject matter of public justification, that is, as what stands in need of public justification. For example, Robert Audi's principle of secular rationale applies to laws and policies that restrict human conduct. But Audi also argues that demands of public justification are higher to the extent that coercion is more immediate or restrictive (2000, 86-9). Christopher Eberle, a critic of Audi and Rawls, sees the obligation to pursue "public justifications" for "favored coercive laws" as a defining element of public reason liberalism (2002, 10). Eberle's alternative ideal of conscientious engagement incorporates this same obligation and defends it on the grounds that citizens are understandably "deeply averse" to being coerced by others (94-100). Finally, the more recently developed and increasingly influential convergence approach to public justification – or what I call the *asymmetric convergence model of public justification* – directly specifies "coercive law" as the subject matter of public justification. According to this model:

> (PJ_1) Coercive law *L* is justified in a public *P* if and only if each qualified member *i* of *P* has sufficient reason(s) R_i to endorse *L* (Vallier and D'Agostino, 2013).

The asymmetric convergence model allows for the public justification of law *L* provided that each qualified member (hereafter "citizen") has a *merely intelligible* justifying reason sufficient to support *L*, even if some (or all) such reasons are intelligible only from the standpoints of particular worldviews, ethical conceptions of the good, or sectarian comprehensive doctrines. This model is asymmetrical insofar as it allows for a wide range of potential defeater reasons that would prevent public justification of *L*. That is, unlike justifying reasons, defeater reasons are more likely to be individually politically decisive. For example, even if reason *R* is alone insufficient to justify *L*, reason *R**, recognized by only one qualified citizen, might be sufficient to prevent *L's* public justification. So while (PJ_1) might be understood in different ways, the convergence model combines it with asymmetry in the relative weight assigned to justifying reasons and defeater reasons in practices of justification.

Elsewhere I criticize the asymmetric convergence model by challenging its strong presumption against direct state coercion (Boettcher 2015; cf. Vallier 2016). This criticism is summarized as follows: An implicit assumption of the model is that only certain positive actions of the state – especially its laws, policies, or administrative decrees that would directly and coercively restrict or compel human conduct – stand in need of public justification. But if the Marxist theory of structural coercion is at least intelligible, and if the state's failure to enact certain laws and policies – e.g., egalitarian property rules, restrictions on inheritance, employment regulations, guaranteed health care, or other measures of distributive justice – allows for otherwise avoidable forms of structural coercion in labor markets or elsewhere, then various forms of state inaction appear to sustain or promote coercion, at least from the standpoint of some citizens. Some citizens will call for measures to prevent structural coercion, while others will object to these measures as themselves needlessly or excessively coercive. The result is that no particular property rights regime would meet with the endorsement specified by principle (PJ_1), as each proposed regime might be defeated by the intelligible reasons of some qualified citizens. Thus, the asymmetric convergence model encounters an incompleteness problem according to which no publicly justified arrangements are found for a pressing set of political issues. Pursuing this criticism is not the purpose of the present paper, though I do return to a version of its incompleteness problem in section VI.

Here I begin instead with a different objection to (PJ_1), developed in a recent article by Colin Bird (2014). For a public justification principle may seem to apply even in cases that do not involve coercion in any immediate way. If so, then (PJ_1) fails to capture fully the domain of public justification. In Bird's imagined example, the proceeds of an entirely voluntary state-run lottery are used by government *noncoercively* to promote a perfectionist conception of the good. Funds are devoted to voluntary educational programs, research grants, public memorials and iconography, and the dissemination of relevant information. This proposal would directly promote an ethical conception of the good that some citizens reasonably reject. The problem, according to Bird, is that some qualified citizens would be unable to *condone* such political actions, where *condoning* suggests a willingness to associate oneself with another agent's action.

Citizens are already involuntarily complicit in democratic political actions that are carried out in their name. So even when there are no proposed coercive restrictions on or prohibitions of human conduct, as in the lottery example, citizens rightfully object to political actions they cannot at least condone. Bird maintains that laws and policies should not be based on justifications that would alienate citizens from "their civic standing as equal co-authors of democratic legislation" (2014, 203). So, according to this view, a public justification requirement is warranted, but not because of concerns about coercion. Bird concludes that a principle of public justification does not aim to justify coercion at all. I shall argue that this conclusion is mistaken.

II. RAWLSIAN PUBLIC JUSTIFICATION AND COERCION

One solution to the problem posed by the lottery example is to reconceive the subject matter of public justification. Andrew Lister distinguishes two ways in which a qualified acceptability criterion, such as the one presented in (PJ_1), might function as a *constraint* in public justification (2011; 2013). It might function as a constraint on state coercion, where the default position would be inaction, i.e., government not exercising coercive power at all with respect to some political problem or question. Alternatively, the criterion might function as a constraint on reasons for political decisions about social arrangements, where the presumption is that nonpublic reasons should not determine matters of law and policy. Where the criterion serves as a constraint on state coercion, as in the asymmetric convergence model, the public justification principle has a classical liberal tilt. As we have seen, if according to principle (PJ_1) coercive law L remains unjustified as long as one qualified individual citizen i_1 has sufficient reason to reject L, even in cases in which the balance of shareable public reasons overwhelmingly supports *L's* adoption, then many otherwise desirable laws and policies will be quite difficult to justify publicly.

The alternative suggested above is to apply the public justification principle to political decisions, or, more precisely, political decision-making concerning the basic institutional structure. The basic institutional structure is quite obviously a Rawlsian notion, denoting society's main constitutional, political, and economic institutions that together provide for a unified scheme of social cooperation over time (Rawls 2005, 11, 258). This focus on decisions about basic institutional arrangements, rather than coercion or coercive law, appears in a Rawlsian alternative to (PJ_1) and the asymmetric convergence model of public justification. According to the Rawlsian model:

> (PJ_2): Decision D about basic institutional-structural matter L is publicly justified in a public P if and only if each *reasonable* member i of P sincerely has sufficient, accessible, credible, and reasonably acceptable reason(s) R_i to endorse L instead of its alternatives.

Elsewhere I explain the many details of this principle, such as what it means for citizens to be reasonable or how to define the discursive qualifiers *sufficiency, accessibility, credibility,* and *reasonable acceptability* (Boettcher 2015). The main points to emphasize here are, first, that (PJ_2) represents an alternative to asymmetric convergence with its default of state inaction, and, second, that (PJ_2) is consistent with the intuition informing Bird's lottery example, namely, that important government decisions should be publicly justified.

Bird takes an even stronger position, however, claiming that the goal of justifying coercion plays *no role* in motivating a public justification requirement. To be sure, Bird recognizes that acts of coercion should be justified. He proposes a Coercion Principle according to which "political action is legitimate only to the extent that any coercion of private individuals required by it receives proper justification," where proper justifications

are not equivalent to public justifications (Bird 2014, 190). Bird distinguishes correctness-based complaints about why a coercive law is bad or faulty from standpoint-dissonance complaints, whereby citizens cannot reconcile their reasonable doctrinal or ethical commitments with the standpoint from which laws and policies are justified. Coercion as such is justified when all relevant correctness-based complaints have been answered, whereas public justification is said to aim at something different, namely, minimizing standpoint-dissonance and thereby respecting each citizen as a free and equal co-author of the laws that govern them. According to Bird, this is why we demand suitable public justifications in the non-coercive lottery example or even in those cases in which the citizens registering their standpoint-dissonance complaints are not the same individuals who are personally facing the prospect of greater coercion from proposed laws. Citizens must be able to *condone* the acts of the public of which they are members.

III. COERCION AND THE BASIC STRUCTURE

The subject matter of public justification should be political decision-making about the basic institutional structure of society, including political decisions made within that structure. This means that the qualified acceptability criterion in (PJ_2) should be a constraint on reasons rather than a constraint on state coercion as such. However, *pace* Bird, I also argue that the justification of state coercion must play *some role* in motivating the public justification principle. To establish this conclusion, I begin with a variation on his lottery example.

1. The Game Night Analogy

Suppose that members of a neighborhood association begin organizing a weekly fundraising event in private and public spaces around their neighborhood. While participation is entirely voluntary, "Game Night," featuring games of chance, trivia games, and other activities, is entertaining and attracts crowds from surrounding neighborhoods. Game Night is also quite profitable and the proceeds are used for weekend educational and social youth programs in the neighborhood, again entirely voluntary, that work against rigid gender roles according to which the virtues of athletic skill, risk taking, competitiveness, and self-reliance are construed as paramount for boys and men. And the reasons in support of these programs are in turn based on a progressive and ecumenical religious creed that the majority of residents share and which informs youth programming through their association. Programs are sponsored by the neighborhood and open to children from surrounding areas. The neighborhood becomes well known for both its weekly entertainment and youth programming.

Suppose that dissenting residents in my example – e.g., theologically conservative traditionalists or maybe some football-obsessed parents – experience standpoint-dissonance much like citizens in Bird's example who encounter laws and policies adopted in their name but justified primarily or solely by a doctrinal or ethical perspective

they reasonably reject. Dissenting residents not only disagree with the content of the youth programming but they resent the fact that their neighborhood is now known for sponsoring it. If the goal of avoiding such standpoint-dissonance is what motivates a public justification requirement, as Bird suggests, and the cases are relevantly similar, then by analogy the requirement should apply in my example. If so, then unless they are supported by sufficient public justifications, the neighborhood's programs would be illegitimate social practices, wrongful impositions on residents who object to the background ethical commitments and religious doctrines that motivate and sustain them.

This conclusion is plainly counter-intuitive from any liberal point of view that celebrates freedom of association. Residents organizing voluntary neighborhood programs are not running afoul of any public justification requirements, nor would they be doing so even if roles were reversed and majority-supported programming were aimed at solidifying rather than resisting rigidly traditional gender roles. And yet Bird's example is persuasive in many respects – indeed members of the democratic public arranging the basic institutional structure by allocating public funds to social and educational programs should expect decisions to be publicly justifiable and not based solely on a comprehensive religious or ethical doctrine that some citizens reasonably reject. This suggests that Bird's lottery and my Game Night example are not relevantly similar cases.

A key difference between the two cases is that political decisions concerning the basic institutional structure of society are necessarily linked to the exercise and authorization of coercive state power. To be sure, no persons directly experience coercive pressure in Bird's state lottery example. But to concentrate on that point is to risk misrepresenting the relationship between the democratic public, the state, and coercion.

2. The Basic Structure and the Law

Rawls identifies the basic institutional structure as the subject of justice for reasons based ultimately on respect for persons as free and equal citizens (2005, 257-71). First, such a focus is necessary if background conditions – e.g., the preconditions for social relations and economic transactions – are to remain fair over time, consistent with moral equality. Second, the basic structure has a profound impact on the abilities and aspirations that shape our choices, projects, and commitments as individuals and members of groups and associations. I submit that these same reasons also point to the basic structure as the subject matter of public justification. But additional support for that judgment is provided by reflection on the liberal-democratic rule of law as the inherently coercive organizing principle of an unavoidable form of social cooperation, viz., the modern state.

A basic institutional structure is organized through a legal system. While law is not reducible to acts of state coercion, law's action-guiding qualities are supported both by raising claims to validity and by stabilizing expectations, partly through issuing coercive threats (Habermas 1996). Lawful authority includes the right to authorize coercive enforcement of legal norms and duties, including the regulation of "when, and under

what conditions, coercion may be employed" (Lamond 2001, 55). The coerciveness of law is obvious in forms of punishment associated with criminal law. But private law is "rife with coercion as well" (Blake 2002, 277). Contract law and property law involve rules of ownership that are enforced through coercive measures. The same is true of the law of taxation, which ultimately assigns authority over various resources and compels citizens to undertake actions, such as filing documents or making quarterly or annual payments to the state.

Furthermore, the democratic rule of law is not simply a series of isolated rules but a unified system (Blake 2002, 54-7). This system claims an indeterminate authority over all members of the political society. Even when particular legal rules are not immediately coercive in the sense of employing substantial pressure to determine an agent's will, they are part of a legal system that is both generally coercive and recognized as authorizing acts of coercion. Moreover, the political and legal systems of the basic institutional structure maintain *final* authority over how coercion may be exercised. Relevant legal agents within this structure are authorized to address disputes about coercion between individuals, as in cases of liability and assumed risk, contractual duress, blackmail, or sexual harassment. However, typically no court of appeals functions outside of public reason for addressing complaints about constitutional principles or other fundamental political decisions and arrangements. The legal authority of political society – its state and government – is subordinate to no other authority and superordinate over the all other individuals and associations within the territory that it governs (Raphael 1970, 55).

Finally, a political society with a basic structure organized through the democratic rule of law is not like a voluntary association that one joins or abandons, depending on one's interests and values. Here my argument rests on two assumptions. First, a complex legal-political system is necessary for fruitful social cooperation under socio-historical forces of modernization and liberalization. Second, the democratic rule of law is necessary for the legitimacy of such a system. So, insofar as these assumptions are sound, then unlike in the case of voluntary associations or religious and cultural communities, the role of citizen is a generally unavoidable social role. To be sure, some persons deliberately reject this role, while others are excluded from it. But fruitful and legitimate social cooperation cannot be sustained unless a critical mass of a society's denizens identify as citizens and recognize their civic obligations. Citizens are indeed responsible for political decisions as equal members of their political society.

To summarize, there are several features of a liberal-democratic political society's basic institutional structure that together make it the appropriate site of public justification. First, membership in such a society is involuntary and participation in its political life is normally unavoidable. Second, its principal mode of organization – the rule of law – is inherently coercive. Third, its legal and political institutions are the final court of appeals for disagreements, including disagreements about coercion itself. None of these features are present in the Game Night example, but all are implicated in decisions about how to allocate public resources from a state-run lottery.

Bird focuses instead on the way that democratic citizens share responsibility for their political decisions and should be able to condone these decisions as co-legislators. That feature is also crucially important. Yet, as my example suggests, its significance arises through its connection to, and not independently of, the other three features. Citizens share responsibility for political decisions about the basic structure because of its involuntariness, coerciveness, and finality. It is in the context of these features of political life that an underlying duty of mutual respect for one another as free and equal citizens gives rise to a principle of public justification and requirements of public reason. We should be able to endorse the basic terms on which we unavoidably limit the freedom of our moral equals through a coercively enforceable institutional apparatus.

IV. REPLIES TO OBJECTIONS

1. The Presumption against Coercion

Bird refers to the kind of view I've sketched as a "mixed view," according to which requirements of public justification depend "both on the conditions for justified coercion and on the desiderata of democratic co-authorship" (2014, 205). He rejects the mixed view because it might weaken the presumption against coercion, implicit in his own Coercion Principle. Recall that the Coercion Principle states that political action is legitimate only to the extent that any coercion of private individuals required by it receives proper justification, where proper justification is not limited by a standard of public justification. If debates about coercion were limited by a public justification requirement, then some arguments *against* coercion, such as those deriving from Mill's Harm Principle, would be excluded because they are subject to reasonable disagreement. The unwelcomed result, according to Bird, is that coercion would be easier to justify, and this result is supposed to serve as a *reductio* against the mixed view.

This objection fails, for two reasons. First, there's the dubious assumption that the Harm Principle is generally a source of inaccessible nonpublic reasons. Insofar as the Harm Principle could be formulated in terms of basic political values and independently of Mill's comprehensive liberalism, arguments deriving from it are consistent with a public justification principle like (PJ_2). Such a formulation is already available (Rawls 2007, 291-3). The second reply concerns the connection between public justification and the accessibility of relevant arguments for and against coercion. For if arguments about coercion carry less weight when their premises derive solely from religious or other comprehensive doctrines, then those arguments, or at least many of them, will prove to be less important and influential according to most standard public justification theories. If so, then such arguments are not likely to present especially strong objections to various coercive proposals. The asymmetric convergence model is an exception, as it allows for merely intelligible nonpublic reasons to defeat potential coercive laws. But Bird does not endorse this model. He seems to think that more open-ended discussion based on

the Coercion Principle should be completed before narrower questions of standpoint dissonance are addressed through public reason. Yet this suggestion presupposes an artificial and ultimately faulty conception of public deliberation.

The problem lies in Bird's assumption that the Coercion Principle can be satisfied independently of a public justification requirement. That's not the case, since citizens are not always in a position to evaluate the soundness of various arguments for or against coercion unless those arguments are accessible in the way that a public justification principle such as (PJ_2) requires. An argument is *accessible* when it can be meaningfully evaluated in light of standards shared by citizens generally, such as reliable perception and observation, rules of inference, common sense, natural and social-scientific methods and results, basic moral-political reasoning such as human rights discourse, historical evidence, and shared democratic political values. Only by scrutinizing the accessible arguments for and against coercion would democratic citizens be able to determine with confidence whether the Coercion Principle is satisfied.

Citizens should not have to adopt doctrinal or ethical standpoints they otherwise reject just in order to have equal access to the reasons in support of (or against) basic institutional arrangements, including coercive laws and policies. Of course some citizens who disagree with coercive law *L* may still resent being coerced by *L* even though the reasons in support of *L* are fully consonant with their evaluative standards. But similarly they might resent being coerced by *L* even though *L* satisfies the Coercion Principle. Disagreement about political outcomes is inevitable. Still, if democratic citizens are the co-authors of their political decisions, and if my description of the basic institutional structure as organized through an inherently coercive legal system is correct, then citizens as agents of coercion should take responsibility for their actions by justifying laws and policies in terms that others might reasonably accept.

2. Involuntariness and Coercion

A second objection to my thesis turns on the distinction between coerciveness and involuntariness, both of which are associated with political society's legally organized basic institutional structure. Bird also emphasizes the fact that political society is involuntary and cites this fact in his condonation/democratic co-authorship argument in support of a public justification requirement. He cautions against conflating involuntariness and coerciveness, observing that the privileges of democratic citizenship are bestowed involuntarily, but not necessarily coercively (Bird 2014, 202-3). A potential objection to my thesis seems to follow: an appeal to coercion is unnecessary for a convincing account of why the basic institutional structure should be the subject matter of public justification; due attention to the involuntariness of political society, and perhaps the finality of its political decision-making, is sufficient in this regard.

A first reply to this objection is to observe that there is at least one institution, the family, which is partly involuntary but not necessarily coercive. While many people

voluntarily make choices about whom to marry or whether to have children, nobody chooses to be born into one particular family or another. Norms governing certain involuntary familial relationships, such as filial duties, are not typically thought to require public justification, as would be the case if involuntariness were sufficient for motivating public justification. Second, state coercion should be justified, for it limits the choices that persons might otherwise make as part of their freedom to pursue and revise a conception of the good. As suggested above, the justification of state coercion depends upon citizens having adequate access to the reasons for and against proposed laws and policies. This is why a plausible public justification principle such as (PJ_2) includes an accessibility criterion for the reasons that would justify (or defeat) law *L* or its alternatives.

It is worth noting that the Rawlsian political liberalism inspiring Bird's approach as well as my own also distinguishes involuntariness from coerciveness. Rawls cites both as "special features" of the "political relationship" in a constitutional democracy (2005, 135, 216). Political power is always coercive and indeed uniquely coercive insofar as only government is authorized to use force in upholding law. In a constitutional democracy this coercive political power is also the power of the public. After briefly expounding these special features of the political relationship, Rawls then writes that "[t]his raises" the question of how to understand liberal legitimacy and requirements of public reason. The subject term "this" in "this raises" refers not just to the involuntariness of political society but the coerciveness of the public political power that is "regularly imposed" on citizens (2005, 136).

The appeal to condonation and co-authorship does not fully explain why the decisions of a democratic public must be publicly justified. After all, and as the Game Night example suggests, we do not normally believe that all collective decision-making – i.e., the choices and actions of all the groups and associations of which we are members – must meet standards of public justification. What then is so special about *liberal-democratic political decision-making*? The answer that I've proposed links involuntariness to other characteristics of political life that significantly affect our ability to cooperate together fairly as free and equal citizens, namely, an inherently coercive legal system and the finality of political decision-making. There are, then, multiple but interconnected grounds for seeing the basic institutional structure as standing in need of public justification. It follows that coercion plays *some role* in motivating a principle of public justification even if the best formulation of the principle does not target coercion as such as the main subject matter of public justification. We should aim to have a publicly justified basic institutional structure in part because of its inherent and systematic coerciveness.

V. WHAT IS COERCION?

Public reason liberalism is a form of liberalism and, as such, includes a presumption against state coercion. This presumption is strong in the asymmetrical convergence model with its apparent default of state inaction where no law or policy satisfies its rather demanding test of public justification. Alternatively, while Bird's Coercion Principle implies a presumption against coercion, he does not assume a default position of state inaction with respect to pressing

political questions. Nor does Rawls when he avers that the content of public reason is provided by a reasonable political conception of justice that prioritizes basic rights and liberties over claims for the general good. These latter examples suggest that the presumption in favor of liberty and against coercion in public reason liberalism does not depend on identifying coercion directly as the subject matter of public justification, as is the case with asymmetric convergence and principle (PJ_1).

At the same time, the concept of coercion is often left unexplained by public reason liberals. "Coercion" is not listed as an entry in *Political Liberalism's* index, nor is it discussed in any detail by the book's lectures. Gerald Gaus addresses the question of coercion's meaning more directly, observing that the term is notoriously difficult to define. While Gaus does not offer a full theory of coercion, he notes that state coercion is "especially morally problematic" because "it employs force, or threatens to use force, against the persons of its citizens" (Gaus 2010b, 242). Gaus's suggestion about how to understand coercion for the purposes of public justification is promising, though it also represents something of a departure from much of the recent philosophical literature on the topic.

1. Coercion as Pressure

Leading philosophical analyses of coercion investigate how a coercer's proposal puts pressure on the coercee to influence his or her actions. Alan Wertheimer develops his influential theory by studying U.S. case law, and defines coercion accordingly:

> (C_p) A coerces B to do X if and only if (1) A's proposal creates a choice situation for B such that B has no reasonable alternative but to do X and (2) it is wrong for A to make such a proposal (1987).

Because (C_p) identifies only threats, and not offers, as coercive, it crucially depends on specifying a baseline according to which *B* will be worse off if he or she fails to accede to *A's* proposal. And Wertheimer is among those who favor defining the baseline in moralized terms, that is, by appealing to *B's* moral rights and interests. Indeed, according to his theory, a conception of moral rights and duties is necessary to determine both if *A's* proposal is wrongful and whether *B* is blameworthy for acting in the way that *A* demands.

Despite its considerable influence in contemporary discussions of coercion, this kind of theory is not well suited for explaining why coercive laws and policies typically stand in need of public justification. The second step in (C_p), the so-called proposal prong, identifies the wrongfulness of the coercer's proposal as a necessary condition for deeming a proposal coercive. The implication is that acts of coercion are always wrong, or perhaps always at least *prima facie* wrong. This understanding of coercion makes sense when the question is whether a person's actions under duress are legally or morally excusable. However, it is not especially helpful for explaining the intuition that coercion is sometimes not wrong *precisely because it is publicly justified.* For example, we normally think of criminal penalties designed to deter acts of wanton violence as coercive, but also publicly justifiable and entirely morally appropriate. It is of course possible to hold the following position: It is *prima facie* wrong to threaten would-be violent perpetrators with

criminal penalties, and thereby to coerce them, but still the most appropriate course of action all things considered. But such a rationale seems like a needlessly complicated way of arriving at the conclusion that some coercion is justified.

A second problem concerns Wertheimer's notion that successful threat proposals are essential to coercion. Public reason liberals and other political philosophers typically presuppose a broader domain of coercion, where coercive law and policy includes both threats of harmful force and actual deployments of force, punitive and nonpunitive. For example, coercive law is assumed to apply to citizens generally, even those citizens who ignore or are unmoved by the threats that are attached to noncompliance. Suppose that a particularly ruthless and reckless man, Anton, insists on doing whatever he wants to whomever he chooses, often thereby violating criminal law, regardless of the severity of the corresponding sanctions and penalties. Since the no-reasonable-alternative prong [(C_p) step (1)] is not met, he appears not to have been coerced by criminal law. We nevertheless would want to say that Anton is (fortunately) subject to state coercion – and that he is in fact coerced – when he is forcibly detained, relocated, imprisoned, incapacitated, or otherwise restricted from acting by the criminal justice system.

By contrast, some people would dutifully act just as the law requires regardless of any threat component. Hayek cites this kind of case in introducing the distinction between state coercion and the threat of state coercion (1978, 142). Hayek's conception of coercion differs from Wertheimer's by not including the wrongful proposal prong [(C_p) step (2)] though it still recognizes the way in which the coercee is essentially left with no genuine alternatives. For Hayek, *A* coerces *B* when *A* intentionally and through threat of harm determines *B's* will in accordance with *A's* purposes (1978, 134). With coercion so defined, the threat of state coercion is a kind of second-order threat, i.e., a threat, or perhaps a warning, that some choices and circumstances will engender the state's coercion through the threat of harm. According to this view, citizens are able to avoid actual state coercion simply by fulfilling their voluntarily incurred obligations, particularly in matters of private law. When individuals deliberately act as they ought, they avoid putting themselves in a position where they would encounter the coercive pressure deriving from an intentional threat of harm.

I submit that even if coercion is sometimes only indirect in this way, or merely threatened counterfactually, persons qua citizens are still subject to it. Recall from section III. 2 that law should be understood as a unified system, authorizing acts of coercion on an ongoing basis and claiming authority over all persons in its domain. Moreover, the distinction between state coercion and the threat of state coercion is difficult to maintain. Hayek cites the example of taxation as unavoidably coercive, presumably because the sanctions associated with it command individuals to act in specific ways. But surely we can imagine someone who not only harbors no resentment about being taxed, but would otherwise prefer to transfer some portion of her earnings, equivalent to or more than her annual tax, to an agent, like government, well positioned to promote the common good. According to the Hayekian distinction between coercion and the threat of coercion, this seems more like a case in which coercion has been avoided. A more plausible position, I think, is to say that taxation is necessarily coercive, regardless of whether the individuals who are taxed experience any threat or pressure associated with it.

2. Coercion as Enforcement

These problems with the pressure approach suggest that a theory of public justification must rely on a broader notion of coercion, in at least two respects. Coercion should not be limited to successful threats. Nor should it be understood as always *prima facie* wrong. Scott Anderson has developed just such an account, focusing less on the pressure experienced by the coercee and more on the power of the coercer to determine another's actions. He writes:

> (C_e) Coercion is "one agent's employing power suited to determine, through enforceable constraints, what another agent will or (more usually) will not do, where the sense of enforceability here is exemplified by the use of force, violence and the threats thereof to constrain, disable, harm or undermine an agent's ability to act" (Anderson 2010, 6).

Coercion, in short, is a powerful agent's employment of enforceable constraints to determine the will of another agent. It is made possible by the power that the coercer acquires or maintains over the coercee as well as the willingness to leverage or deploy that power. This "enforcement approach" does not encounter the problems identified earlier. For (C_e) is consistent with the idea that coercion may involve successful threats of force, or unsuccessful threats, or the actual use of force and other direct mechanisms of constraint. It does not link the existence of coercion necessarily to the coercee's responses or subjective experience. Finally, it is consistent with a morally neutral understanding of coercion, where the justification of coercion depends on the circumstances of who is exercising power over whom and for what reasons.

The enforcement approach recognizes the state as plainly coercive. Anderson notes that paradigmatic cases of coercion are those in which an agent intentionally accumulates and then employs or threatens to employ the power needed to constrain others, "where that power is usually generic in its potency, suited to work against almost any agent, and employable for a wide range of ends" (2010, 28). Obviously not everything done by government is directly coercive in the sense of determining the will of another by forcibly constraining his or her ability to act. Bird's imagined lottery is not directly coercive in this sense. But the enforcement approach to coercion nicely complements the earlier account of the basic institutional structure and the rule of law (section III. 2). Just to repeat: The basic structure is organized through the rule of law. And the rule of law is inherently coercive as a unified system, claiming the general and final authority for the coercion of individuals to assure compliance with social rules.

VI. THE SCOPE OF PUBLIC REASON

This essay began by inquiring into the subject matter of public justification, i.e., the question of what exactly is supposed to be publicly justified. The answer developed herein – namely, political decision-making about the basic institutional structure – may appear to be too imprecise to those readers familiar with debates among public reason liberals, especially Rawlsians, about the so-called *scope* of public reason. Public reason's scope refers the issues to which requirements of public reason apply. It seems, then, that my account of the subject matter

of public justification should be refined even further to specify the scope of public reason more precisely.

The principal distinction is between a narrow view and a broad view of public reason's scope. Jonathan Quong provides the following definitions:

> *The Narrow View*: The idea of public reason must apply to constitutional essentials and matters of basic justice, but need not apply beyond this domain.
>
> *The Broad View*: The idea of public reason ought to apply, whenever possible, to all decisions where citizens exercise political power over one another (Quong 2011, 274).

As is well known, Rawls adopts the narrow view, limiting the scope of public reason to fundamental political questions, namely, constitutional essentials and matters of basic justice. At the same time, the discussion of public reason's scope in *Political Liberalism* is brief and ambiguous, and at times even intimates that the broader view is more appropriate (Rawls, 2005 215). Quong evaluates three arguments in support of the narrow view – based on the basic structure, citizens' basic interests, and completeness in public reason – and ultimately finds each argument inadequate (2011, 275-87). His conclusion is that we should aspire to follow the broad view, even if public reasoning turns out to be incomplete with respect to some of the non-fundamental political questions addressed by that view.

The main arguments of this paper do not require a more precise answer to the question of public reason's scope. That is, their success does not depend on adopting a particular view, narrow or broad, of how much of the basic institutional structure must be publicly justified, as long as we assume that at least the fundamental aspects of that structure, such as constitutional essentials and matters of basic justice, are subject to public justification. In other words, the main arguments of this paper presuppose only the first clause in Quong's definition of the narrow view. Quong's defense of the broad view is plausible, though I shall offer two main comments in lieu of further discussion. First, even if the broad view is correct, constitutional essentials and basic justice nevertheless remain the more urgent and significant matters of public justification. Support for this conclusion is provided by the reasons for focusing on the basic structure in the first place. Insofar as law is *coercive as a unified system*, the constitutional order will determine how various acts of coercion are employed on an ongoing basis. Furthermore, constitutional and basic justice issues are more closely connected to the *finality* of political society's decision-making than ordinary acts of legislation or administrative power. Public reason is the only shareable normative mechanism governing fundamental questions of the former type.

A second main comment responds to the worry that the broad view is more likely to encounter cases of inconclusive public reasoning. Inconclusiveness is a form of incompleteness that results when there are multiple conflicting proposals based on reasonable (and even undefeated) arguments and no justifications are victorious, i.e., sufficient to rebut or undermine all of the competing arguments. While it is surely true that legislative debate falling under the broad view sometimes will be inconclusive, we need not broaden the scope of public reason just in order to encounter that problem. Rawls's own discussion of abortion politics suggests that inconclusiveness is possible at all levels of political deliberation and cannot be avoided just by

sticking to the narrow view. Public reason liberalism should explain how laws and policies might achieve at least some degree of public justification even in the context of ongoing disagreement, including disagreement about the premises and conclusions of the relevant arguments. Whether the scope of public reason is narrow or broad, it is possible for laws, policies, and other institutional arrangements to achieve at least a weak form of public justification, even when no proposal or decision is endorsed by all reasonable citizens (Boettcher 2015).

VII. CONCLUSIONS

The subject matter of public justification refers to that which stands in need of public justification. I have argued that the subject matter of public justification is not coercive law, as assumed by proponents of the asymmetric convergence model, but political decision-making about the basic institutional structure of society. This judgment is based on an analysis of several features of political society and its basic institutional structure, namely, (i) the involuntariness of political society and its state apparatus, (ii) the finality of political decision-making, and (iii) the inherent coerciveness of the rule of law. The legal system is coercive as a unified whole; it claims an indeterminate authority over all members of political society and thereby authorizes particular coercive acts on an ongoing basis. It is true that citizens who respect one another as politically free and equal should be able to condone and take responsibility for the laws and policies that are adopted in their name. Yet this view of liberal-democratic citizenship as responsible and respectful co-authorship is essentially connected to, rather than independent of, the involuntariness, finality, and inherent coerciveness of political society and its legally organized basic institutional structure. This is why we would demand public justifications for resource allocation decisions by government but not voluntary neighborhood associations, even when the resources being allocated are neither collected coercively nor intended directly to support particular coercive acts against individuals. The justification of coercion must play at least some role in motivating a public justification requirement.

For the purposes of a political philosophy like public reason liberalism, coercion occurs when a powerful agent employs enforceable constraints in order to determine the will of another agent. On this view, a legally organized democratic government is an essentially coercive agent, where its particular acts of coercion may take the form of either threats to use force or the actual use of force. This enforcement approach to coercion is consistent with a general presumption against state coercion and with the important notion that some acts of coercion are not *prima facie* wrong because they are justified.

jboettch@sju.edu

REFERENCES

Anderson, Scott. 2010. The Enforcement Approach to Coercion. *Journal of Ethics and Social Philosophy* 5 (1): 233-75.

Audi, Robert. 2000. *Religious Commitment and Secular Reason.* New York: Cambridge University Press.

Bird, Colin. 2014. Coercion and Public Justification. *Politics, Philosophy and Economics* 13 (3): 189-214.

Blake, Michael. 2002. Distributive Justice, State Coercion, and Autonomy. *Philosophy and Public Affairs* 30 (3): 257-96.

Boettcher, James. 2015. Against the Asymmetric Convergence Model of Public Justification. *Ethical Theory and Moral Practice* 18 (1): 191-208.

Eberle, Christopher. 2002. *Religious Conviction in Liberal Politics.* New York: Cambridge University Press.

Gaus, Gerald. 2010a. *The Order of Public Reason.* New York: Cambridge University Press.

———. 2010b. Coercion, Ownership and the Redistributive State: Justificatory Liberalism's Classical Tilt. *Social Philosophy and Policy* 27 (1): 233-75.

Habermas, Jürgen. 1996. *Between Facts and Norms*, translated by William Rehg. Cambridge, MA: The MIT Press.

Hayek, Friedrich. 1978. *The Constitution of Liberty.* Chicago: University of Chicago Press.

Lamond, Grant. 2000. The Coerciveness of Law. *Oxford Journal of Legal Studies* 20 (1): 39-62.

———. 2001. Coercion and the Nature of Law. *Legal Theory* 7 (1): 35-57.

Lister, Andrew. 2011. Public Justification of What? Coercion vs. Decision as Competing Frames for the Basic Principle of Justificatory Liberalism. *Public Affairs Quarterly* 25 (4): 349-67.

———. 2013. *Public Reason and Political Community.* London: Bloomsbury.

Quong, Jonathan. 2011. *Liberalism Without Perfection.* New York: Oxford University Press.

Raphael, D.D. 1970. *The Problems of Political Philosophy.* London: Pall Mall Press.

Rawls, John. 2005. *Political Liberalism*, expanded edition. New York: Columbia University Press.

———. 2007. *Lectures on the History of Political Philosophy,* edited by Samuel Freeman. Cambridge, MA: Harvard University Press.

Reiman, Jeffrey. 2012. As *Free and as Just as Possible.* Malden, MA: Wiley.

Vallier, Kevin. 2016. In Defense of the Asymmetric Convergence Model of Public Justification: A Reply to Boettcher. *Ethical Theory and Moral Practice* 19 (1): 255-66.

Vallier, Kevin, and Fred D'Agostino. 2013. Public Justification. In *Stanford Encyclopedia of Philosophy*, edited by Edward Zalta. (http://plato.stanford.edu/archives/spr2013/entries/justification-public/).

Wertheimer, Alan. 1987. *Coercion.* Princeton: Princeton University Press.

Yankah, Ekow. 2008. The Force of Law: The Role of Coercion in Legal Norms. *University of Richmond Law Review* 42 (3): 1195-256.

Public Reason 8 (1-2): 31-49

Is Rawls Really a Kantian Contractarian?

Baldwin Wong
Chinese University, Hong Kong

Abstract: In most of the introductions to Rawls and contemporary contractarianism, Rawls is seen as the representative of Kantian contractarianism. He is understood as inheriting a contractarian tradition that can be traced back to Kant and which has inspired followers such as Barry and Scanlon. This paper argues that the label does not fit Rawls. While a Kantian contractarian would presuppose a monistic conception of practical reason, Rawls is a hybrid contractarian who presupposes a dual conception. I shall first argue that the way in which a contractarian model is labeled is determined by its conception of practical reason. Then I show that Rawls and Kantian contractarians assume different conceptions of practical reason, and therefore should be seen as belonging to two strands of thought. I further argue that, although Rawls acknowledges his intellectual affiliation with Kant, he cannot be considered a Kantian contractarian in the commonly understood way. In his *Lectures on the History of Moral Philosophy*, Rawls interprets Kant as endorsing a hybrid contractarian model that is similar to his. By understanding Rawls as a hybrid contractarian and not confusing his philosophical project with that of Kantian contractarians, Rawls's contribution to the history of contractarianism can be better evaluated.

Key words: Rawls, Kant, Sidgwick, Kantian contractarianism, Scanlon, practical reason.

In contemporary discussions of contractarianism, contractarian theories are usually divided into Hobbesian and Kantian contractarianism.[1] In the literature, David Gauthier (1986) is viewed as a representative of the former, and John Rawls (1999, 2005) is seen as a prominent representative of the latter variety. Rawls is commonly seen as the father of Kantian contractarianism; as inheriting a contractarian tradition that can be traced back to Kant and which has inspired followers such as T. M. Scanlon (1998) and Brian Barry (1995). This label is widely used in the introductions to Rawls as well as in introductions

1] Here I should clarify my use of the term "contractarianism," since there is no standard definition in current academic discussions. Stephen Darwall (2003) uses "contractualism" to refer to the strand of social contract theories which are influenced by Kant and which emphasize agreements which could be reasonably accepted by everyone, whereas he uses "contractarianism" to refer to the strand of social contract theories which are influenced by Hobbes and which emphasize agreements which could advance the rational interests of everyone. However, not everyone agrees with this usage. For example, Geoffrey Sayre-McCord (2000, 247-67) and Jean Hampton (2007, 478-92) use "Kantian contractarianism" to refer to the former strand, and "Hobbesian contractarianism" to refer to the latter strand. Nicholas Southwood (2010, 25-85) uses the labels "Hobbesian contractualism" and "Kantian contractualism." Samuel Freeman (2007a, 17-44) prefers the labels "interest-based contract" and "right-based contract." Despite not using the word "contract," Brian Barry's (1989) distinction between "justice as mutual advantage" and "justice as impartiality" also makes the same kind of division.

I disagree with all of these ways of distinguishing theories since they overlook the hybrid model. Here, in this essay, "contractarianism" is employed as a broad term that can be used to refer to all theories that use social contracts to justify principles of justice, regardless of whether a contract is rational or reasonable. Hence, unlike Darwall's usage, this term does not specifically refer to the Hobbesian social contract model.

to contractarianism.[2] However, there is some lack of clarity about what it means to be a contractarian. It is also unclear whether Rawls, Scanlon and Barry can be seen as belonging to the same strand of contractarianism. The purpose of this paper is to challenge the label as applied to Rawls and to argue that the Rawlsian contractarian model is in fact a non-Kantian, hybrid model that presupposes a dual conception of practical reason.

The paper consists of five sections. First, using Kantian contractarianism as an example, I shall outline the theoretical structure of a contractarian model and argue that the question of whether a theory fits a certain contractarian model is decided by its conception of practical reason (Section I). I shall then discuss the hybrid character of the Rawlsian contractarian model and show that Rawls assumes a dual conception of practical reason (Section II). Having discussed the character of Rawlsian contractarianism, I shall compare Rawls with two Kantian contractarians, Scanlon and Barry, and show how Rawls's conception of practical reason is plainly different from theirs (Section III). I will briefly introduce Sidgwick's dual conception of practical reason and show that, ironically, the philosopher who has a conception of practical reason that is most similar to Rawls's is not any one of the Kantian contractarians, but the utilitarian Sidgwick (Section IV). Finally, I shall discuss a possible objection, namely that Rawls should be seen as a Kantian contractarian due to his self-confessed intellectual affiliation with Kant. My argument will be that Rawls's intellectual affiliation should not be understood without his distinctive, "hybrid" interpretation of Kant. His *Lectures on the History of Moral Philosophy* shows that Rawls reads Kant as constructing a hybrid contractarian model that is similar to his (Section V).

I. THE THEORETICAL STRUCTURE OF CONTEMPORARY CONTRACTARINAISM

Contractarians advocate different moral or political principles, but they share the same methodology. The methodology of contractarianism is to specify moral or political principles by an agreement that would be made by people when they reason correctly. A contractarian theory should consist of three components:

(1) A conception of practical reason

(2) A state of nature

(3) A hypothetical contract (a set of moral or political principles)

A conception of practical reason is an account of how reasoning should be when people reflect on practical issues. For Kantian contractarians, the conception of practical reason is *monistic*, i.e., the correctness of practical reasoning is measured by a single

2] For examples of introductions to Rawls who interprets him this way, see Maffetone (2010, 33-34); Audard (2006, 50-51); Mandle (2009, 84-85); Freeman (2007c, 8); Pogge (2007, 188-95). For examples of introductions to contemporary contractarianism which interpret Rawls in this way, see Darwall (2003); Sayre-McCord (2000); Hampton (2007); Kymlicka (1993).

standard, which is reasonableness.[3] Reasonableness means that the process of thinking is regulated by an absolute moral constraint that requires people to "act in ways that could be justifiable to others" (Scanlon 1998, 154). People should respect the reasons given by others and be open to revising their thought in order that it should be publicly justifiable. This revision may imply that one has to compromise and give up satisfying one's rational interests. Nevertheless, Kantian contractarians believe that reasonable people are obliged to do so; being reasonable should be accorded priority over the rational pursuit of interests.

After assuming a conception of practical reason, contractarians then define hypothetical contractors who are placed in a state of nature, which is a hypothetical, idealized choice situation. The function of this hypothetical situation is to demonstrate a collective procedure of deliberation and conversation in which people reason correctly. Kantian contractarians, such as Scanlon, assume hypothetical contractors to be well-informed, reasonable people who share an end of "living with others who are also reasonable" (Scanlon 1998, 192). Due to this end, people understand that they are standing in a relationship of mutual recognition with one another, which means that they would respect one another as beings that are capable of asking for justification.

Contractarians then try to identify what moral or political principles would be agreed by these hypothetical contractors. Kantian contractarians, such as Scanlon, claim that all reasonable people eventually would choose "principles that no one could reasonably reject." (Scanlon 1998). Reasonable people would go into conversations with one another and propose different principles (Kumar 1999, 277-78; O'Neill 2004, 29). It is a process of reason-exchanging and contractors would give generic reasons to support principles that they propose. Generic reasons refer to reasons which are based on "commonly available information about what people have reason to want," rather than being based on "the particular aims, preferences, and other characteristics of specific individuals" (Scanlon 1998, 204). The strength of a reason depends upon the burden which would be imposed on someone if this generic reason were rejected. The larger the burden is, the stronger the generic reason is. In the contract situation, contractors would compare different generic reasons and would finally reach an agreement on principles supported by the strongest generic reason, that is, principles that are "least unacceptable to the person to whom [they are] most unacceptable" (Kumar 1999, 294). These are principles that "no one would reasonably reject," for no other people could further give any generic reasons to reject these principles. These principles represent a moral ideal of equal respect because they impose burdens on all society members fairly without being affected by unequal bargaining power.

The three-component structure shown above reveals the core belief of contemporary contractarianism – moral and political principles are justified if they would be agreed upon

3] However, one should be careful about the fact that Kantian contractarians disagree over the relationship between desire and reason. Barry and Richards account for the motivation to be reasonable in terms of a moral desire (Barry 1995, 164-65; Richards 1971, 242), whereas Scanlon argues that reason, not desire, is fundamental in motivating people to action (Scanlon 1998, 17-77). Nevertheless, they all share an assumption that "being justifiable" is the absolute standard in the evaluation of practical reasoning.

in a situation where people engage in *correct practical reasoning*. While contractarians share this core belief, they assume different conceptions of practical reason. This leads to different designs of the state of nature, and thus different results of hypothetical agreement. Hence, the conception of practical reason is the most fundamental component in a contractarian model. It determines the definitions of the other two components.[4] More importantly, whether two contractarians belong to the same or different strands of contractarianism depends on what conceptions of practical reason they assume. For example, Gauthier and Scanlon belong to Hobbesian and Kantian contractarianism respectively because the former upholds the instrumental conception of rationality, which defines correct practical reasoning as the maximization of individual utility, whereas the latter endorses the conception of reasonableness, which defines correct practical reasoning as acting on the ground that could be justifiable to others. Conversely, although Barry and Scanlon derive different results from their contractarian models,[5] they are both Kantian contractarians because they use the same monistic conception of reasonableness as the starting point of their contractarian model. In short, a contractarian model is characterized by the conception of practical reason it assumes.

II. RAWLS'S DUAL CONCEPTION OF PRACTICAL REASON

With the theoretical structure of Kantian contractarianism in mind, it is worth noticing that Rawls's contractarian model is plainly different from that of Kantian contractarianism, and that the difference can be traced back to Rawls's account of practical

4] Some might argue that this definition of contractarianism is more like constructivism. It depends upon how the two labels are understood. Onora O'Neill (2014), who uses contractarianism and contractualism indifferently, has argued that contractualism presupposes an antecedent, contingent agreement on certain intuitions, while constructivism is more concerned with constructing a process of offering and rejecting reasons. She then deems Rawls a contractualist and Scanlon a contructivist. O'Neill's definitions are correct in terms of the scope of their contractarian theories; Rawls' principles of justice can only be applied in liberal democratic societies in which citizens share certain political conceptions, while Scanlon's moral principles do not have this limit in the scope of application. However, this paper is more concerned with the structure of their contractarian theories and, unlike O'Neill's paper, does not intend to differentiate the contructivist from the contractualist/contractarian. Both Rawls and Scanlon are similarly labeled, because both presuppose certain conceptions of practical reason, and both use the hypothetical contract as a heuristic device to show that certain moral and political principles would be agreed upon, or not rejected, when all people reason correctly. Although the contents of their theories are different, their structures are the same. Thus, under this definition, Rawls, Scanlon, and Gauthier are all contractarians.

5] Although both are Kantian contractarians, Scanlon and Barry differ in at least two ways. First, Scanlon upholds a distinction between reasonable non-rejection and reasonable acceptance, and argues that the hypothetical agreement would be on "principles that no one could reasonably reject" (Scanlon 1998, 213-18). Barry thinks that this distinction is untenable; "I do not think anything crucial turns on the distinction between a formulation of the criterion in terms of non-rejectability and a formulation in terms of acceptability" (Barry 1995, 70). Secondly, Scanlon is more concerned about "a narrow domain of morality having to do with our duties to other people" (Scanlon 1998, 6), whereas Barry defines his project as "a theory about the kinds of social arrangement that can be defended" (Barry 1995, 3). In simple terms, Barry is more "political" while Scanlon is more "moral."

reason. Rawls understands practical reason in terms of a *dual conception*, which means that *the correctness of practical reasoning is measured by two mutually independent standards, reasonableness and rationality*. "Rationality" means practical reasoning on how to pursue ends that can contribute to the quality of life and what means are effective to pursue these ends. "Reasonableness" means practical reasoning on how to maintain relationships with one another in a mutually justifiable way. In short, rationality is about personal interests, whereas reasonableness is about interpersonal duty. The two paths of practical reasoning are independent of each other. If an action fails to be justified according to only one of the two standards, then it is partly justified. When it is justified according to both of these standards, it is fully justified. Thus the aim of contractarians who uphold this dual conception is to show that rationality and reasonableness do not necessarily conflict with each other and the hypothetical contract could be both rationally and reasonably justified.

The idea that rationality and reasonableness are two independent paths of deliberation has long existed in Rawls's writings. Rawls denies that reasonableness should be explained in terms of rational interests. "To see justice as fairness as trying to derive the reasonable from the rational misinterprets the original position" (Rawls 2005, 53). He also does not think that rational interest could be derived from reasonableness, since determining whether something is rationally good is different from the question of whether something is reasonably acceptable (Rawls 1999, 496-97). His view that rationality and reasonableness are mutually independent of each other can be seen in the following passage:

> The reasonable and the rational are taken as two distinct and independent basic ideas. They are distinct in that there is no thought of deriving one from the other; in particular, there is no thought of deriving the reasonable from the rational. (Rawls 2005, 51)

Given this dual conception, the hypothetical contractors in Rawls's model have a distinctive hybrid character that cannot be found in the models of other contractarians. However, there is some controversy about how exactly we are to interpret Rawls's hypothetical contract. There are two readings of Rawls concerning the locus of the contractarian element in his theory. The first, orthodox reading is that the locus of the contractarian element is in the original position. Hypothetical contractors in the original position are behind the veil of ignorance and so they are not aware of much of the information that is crucial in the decision-making of actual people. This reading renders Rawls vulnerable to an objection, namely that it is unfair to bind actual people to the decisions of hypothetical contractors because of the "information gap" between actual people and hypothetical contractors.[6] This objection can explain why other Kantian contractarians reject the idea of the veil of ignorance (Scanlon 1982, 110-11; Barry 1995, 9). For the sake of bridging the "information gap" between hypothetical contractors

6] The problem concerning the original position was mentioned by a number of critics of Rawls. Cf. Brown (1988, 444).

and actual people, Kantian contractarians claim that their social contract is made in a "well-informed" manner and they believe that the veil of ignorance undermines the importance of the contractarian element in the whole theory.

Therefore, some Rawls scholars, such as Samuel Freeman, suggest an alternative reading, which conceives the locus of the contractarian element in public agreement in a well-ordered society. "[T]he way in which Rawls' justice as fairness is a social contract position has far more to do with his idea of a well-ordered society than does the original position" (Freeman 2007a, 4). In this new reading of Rawls's model, the hypothetical contractors are the free and equal citizens in a well-ordered society. They aim at reaching a public agreement as the basis of exercising political power. This new reading enables Rawls scholars to explain the relevance of the hypothetical contract in an easier way. The hypothetical contractors in the new reading are well-informed citizens who have the same information, or an even larger set of information than the actual people. No "information gap" exists between hypothetical contractors and actual people. Rather, now the difference is in terms of the level of reasoning; the hypothetical contractors act as a role model to show actual people what they are to choose if they are being rational and reasonable, given the shared set of information.

The two-fold feature of Rawls's model becomes more explicit after clarifying the locus of contractarian element. Unlike the Kantian hypothetical contractors, who are designed to consider only what is reasonable to do, Rawls's hypothetical contractors, free and equal citizens, are designed to have two tiers of practical reasoning, rationality and reasonableness. This is shown in the setting of two moral powers, the capacity for a sense of justice and the capacity for a conception of the good. Although it is widely known that Rawls defines the idea of free and equal citizen in terms of two moral powers, scholars rarely link these moral powers to practical reason. In fact, the two powers correspond respectively to the requirement of being reasonable on the one hand and the rational pursuit of interests on the other. In the following passages, I shall explain these connections.

Rawls defines a sense of justice as "an effective desire to comply with [principles of justice]" (1999, 274). Although he once said that a sense of justice exists in "a purely formal sense" (Rawls 1999, 126), this desire is indeed not an empty sense of duty that can be related to any principle of justice. Rather, it refers specifically to the motivation to embrace principles that specify fair terms of social cooperation. Rawls assumes that free and equal persons have a desire "to live with others on terms that everyone would recognize as fair from a perspective that all would accept as reasonable" (1999, 418-19). Due to this desire, they are willing to think impartially and comply with principles that could be reasonably justified to every free and equal person. Hence, a sense of justice, in fact, represents a motivation to be reasonable, which requires people to mutually respect one another.

Apart from a sense of justice, free and equal persons are also motivated to realize their conception of the good, which is no less important than their sense of justice. If the

sense of justice represents the aspect of "right" in practical reason, then the conception of the good represents the aspect of "good" (Rawls 1999, 491-92). Conceptions of the good are life plans that are made in accordance with formal principles of rational choice (1999, 358-65). Apart from these formal principles, Rawls also argues that conceptions of the good have to be compatible with the higher-order interests of human beings. Free and equal persons are interested in realizing and exercising the two moral powers (Rawls 2005, 74). These interests are basic to any conception of the good, because people without two moral powers are unable to be free and responsible agents who are capable of mastering their wants and answering for their actions. The earlier Rawls even considers these "highest-order interests" to be "supremely regulative" (1980, 312). The later Rawls takes a milder position, but still considers these interests to be fundamental, since "someone who has not developed and cannot exercise the moral powers to the minimum requisite degree cannot be a normal and fully cooperating member of society over a complete life" (2005, 74). Thus, from a rational perspective, free and equal persons understand that a good life cannot be realized without pursuing higher-order interests.

The setting of two moral powers in the conception of free and equal citizens aims to demonstrate what is rational and reasonable for actual people to do. Accordingly, the hypothetical contract of free and equal citizens, i.e., the public agreement on principles of justice, represents a political arrangement that would be justified from both rational and reasonable perspectives. The fact that Rawls's principles of justice as fairness could be agreed upon in the original position explains why they are reasonably right, "since the original position situates free and equal moral persons fairly with respect to one another, any conception of justice they adopt is likewise fair" (1980, 310). Furthermore, justice as fairness is rationally good because it can guarantee "the adequate development and full exercise of the moral powers" (2001, 104) and thereby satisfy the two higher-order interests. No matter who one is in a society, one can enjoy equal basic liberties and adequate primary goods, which are necessary for developing the two moral powers (Rawls 2001, 112-15). Therefore, "the principles of right and justice are collectively rational; and it is in the interests of each that everyone should comply with just arrangements" (Rawls 1999, 504).

To free and equal persons, rationality and reasonableness lead to the same demand of following principles of justice as fairness. Being just is not only a "right" thing, but also a "good" thing to do. The dual character of principles of justice is stated explicitly in the following passage:

> The desire to act justly and the desire to express our nature as free and equal moral persons turn out to specify what is practically speaking the same desire. When someone has true beliefs and a correct understanding of the theory of justice, these two desires move him in the same way. They are both dispositions to act from precisely the same principles: namely, those that would be chosen in the original position. (Rawls 1999, 501)

By embracing principles of justice, free and equal beings can achieve the unity of self. Rationality and reasonableness will not make incompatible requirements. Thus, internal conflict within practical reason can be avoided. People can give unity to their lives by accomplishing this "congruence of rationality and reasonableness," which is explained by the following:

> Now the unity is manifest in the coherence of his plan, this unity being founded on the higher-order desire to follow, in ways consistent with his sense of right and justice, the principles of rational choice…in ways that justice allows, he is able to formulate and to follow a plan of life and thereby to fashion his own unity. (Rawls 1999, 493)

The theoretical structure of Rawls's contractarian model is clearly different from that of Kantian contractarianism. Rawls assumes a dual conception of practical reason which evaluates human action by two standards, rationality and reasonableness. Due to this assumption, the hypothetical contractors, who are free and equal citizens, are designed to have two moral powers, which correspond to rationality and reasonableness. They want to be "right" by acting in a way that is justifiable to others, but they also want to achieve "good" by pursuing both higher-order interests. By demonstrating the hypothetical contract of these persons, Rawls shows that his principles of justice are both rationally and reasonably justified, and thus offers reasons to actual people in explaining why they should follow these principles. I shall call this model *hybrid contractarianism*.

III. DIFFERENCES BETWEEN RAWLS AND KANTIAN CONTRACTARIANISM

The differences between Rawls and Kantian contractarians should not be exaggerated. Like Scanlon and Barry, Rawls advances the view that reasonableness absolutely frames rationality. "The priority of the right over the good in justice as fairness turns out to be a central feature of the conception [of justice as fairness]" (1999, 28). When conflict occurs between rationality and reasonableness, both Rawls and Kantian contractarian believe that reasonableness should have priority. Nevertheless, my point that Rawls has a distinctive view of rational interests that is completely different from Kantian contractarians has not been undermined.

In general, Kantian contractarians use two strategies to explain why reasonableness should have priority over the rational pursuit of interests. The first strategy is a *devaluing strategy*, used by Barry and the earlier Scanlon. The importance of pursuing rational interests, they believe, is always exaggerated. Indeed, the motive to seek a reasonable agreement is a stronger desire in the hierarchy of motivation. It can normally win over other conflicting rational interests (Scanlon 1982, 117; Barry 1995). Thus, it is unimportant for a theory of justice to explain whether justice is compatible with the rational pursuit of interests.

The second strategy is a *reductive strategy*, used by the later Scanlon. The later Scanlon recognizes the tension between reasonableness and rationality, but he

believes that this tension is based on a wrong understanding of goodness, which is the object of rational pursuit. He rejects the teleological account of good and "argue[s] that [this account] has had a distorting effect on our thinking about value in general, and in particular on our view of the relation between rightness ("what we owe to each other") and other values" (1998, 79). Rather, he proposes a buck-passing account of good, which passes the normative buck of goodness to other properties. According to Scanlon, the fact that something is good is not itself a reason to care about it. Whether or not something is good is dependent on whether there are some *reasons* that make us think that it is good. In the buck-passing account, "many other values will be seen to have a structure similar to that which most obviously characterizes our ideas of right and wrong" (1998, 79). The ideas of right and good are not as distinct as many people think. The reason for doing right things is that individuals have a reason "to act in ways that could be justifiable to others" (1998, 154). The same reason may also explain the great value of the relationship of mutual recognition.[7] This great value can, in turn, explain why it is good to be moral, since individuals can enjoy standing in this valuable relationship with one another by complying with contractualist principles. In short, Scanlon reduces the two distinct questions, "what is right to do" and "what is good to do," to a more fundamental question, "what act is justified by the strongest reason."[8]

The devaluing and reductive strategy show that, to Kantian contractarians, rationality is either unimportant, or derivative. It is unimportant since the value of pursuing rational interest is overrated. It is derivative since the notion of rationality can be derived from a more primary notion of reason. Rawls, by contrast, conceives rationality as an important and distinct standard of practical reason. First, unlike Barry and the earlier Scanlon, Rawls places great emphasis on proving that justice as fairness represents something that is rationally good to people. Rawls even believes that a theory of justice that fails to prove this point "is seriously defective" (1999, 398). Thus one of his life-long concerns is to show that being reasonable can be reconciled with the rational pursuit of higher-order interests. This concern with reconciliation eventually motivated Rawls to radically revise his theory in *Political Liberalism*. The later Rawls gives up his comprehensive Kantian doctrine and spends a great deal of effort showing that the higher-order interests are by nature political and freestanding, and so they are compatible with various reasonable

7] Cf. Scanlon (1998, 95, 162). In fact Scanlon does not state clearly what reason makes the relationship of mutual recognition "appealing in itself-worth seeking for its own sake" (1998, 162). Nevertheless, if there must be a reason to explain why a positive attitude toward this relationship should be held, then the reason "to act in ways that could be justifiable to others" seems to be a very plausible candidate.

8] This belief is also mentioned in the most recent writings of Scanlon (2014, 2): "Reasons might be fundamental in the further sense of being the only fundamental elements of the normative domain, other normative notions such as good and ought being analyzable in terms of reasons."

comprehensive doctrines (2005, 74). In short, proving that justice as fairness is in the interests of citizens is a necessary component of Rawls's project.

Secondly, while both Rawls and the later Scanlon recognize the tension between reasonableness and rationality, their strategies for answering this problem are significantly different. Unlike Scanlon, who uses a reductive strategy to blur the distinction between the ideas of right and good, Rawls believes that the ideas of right and good belong to two distinct aspect of practical reason. A reason that explains why a principle is reasonably right does not explain that this principle is rationally good as well. Another different reason must be offered in order to explain the goodness of this principle. As the following passage explains,

> [t]he concepts of justice and goodness are linked with distinct principles… More precisely, each concept with its associated principles defines a point of view from which institutions, actions, and plans of life can be assessed. (Rawls 1999, 496-97)

The distinction between the ideas of right and good often appears in Rawls's writings and is even described as being "absolutely crucial" by Rawls's closest colleague, Burton Dreben (Dreben 2003, 321-22). Accordingly, Rawls uses neither the devaluing strategy nor the reductive strategy to resolve the conflict between rationality and reasonableness. Rather, he answers the two questions, "what is rational to do" and "what is reasonable to do," separately and shows that these two questions lead to the same answer in his contractarian model.

The contrast between Rawls and Kantian contractarians on how they resolve the conflict between reasonableness and rational interests shows that they begin to construct their contractarian models with fundamentally different conceptions of practical reason. Kantian contractarians assume a monistic conception of reasonableness. Therefore, what matters is merely whether the contract would be reasonably justified or not. The question of whether the contract would be in the rational interest of contractors plays a negligible role in their theory. But Rawls assumes a dual conception of practical reason. The question of whether the contract is rationally justified should be taken seriously. If, as mentioned in the last section, the conception of practical reason is fundamental in defining the character of a contractarian model, then obviously Rawls's model should be separated from Kantian contractarianism and, consequently, it should be seen as a distinctive model, a hybrid contractarian model.[9] The differences between these two models are summarized in the following table:

9] Nevertheless, as an anonymous reviewer pointed out to me, the hybrid character of the contractarian model is not as strong throughout Rawls's life. Although Rawls has never abandoned the belief that justice should be in the rational interests of citizens, the belief becomes less important in his later works. Rather, he emphasizes more about the idea that reasonableness must have priority in justifying principles of justice. In brief, the hybrid character is stronger in *A Theory of Justice* as well as "Kantian Constructivism," but becomes less obvious in *Political Liberalism* and *The Law of Peoples*.

	Kantian contractarianism	*Hybrid contractarianism*
The conception of practical reason	The monistic conception of reasonableness: reasonableness is the only standard of evaluation for reasons and actions	The dual conception: rationality and reasonableness are two mutually independent standards of evaluation for reasons and actions
The characteristics of hypothetical contractors	Well-informed people who are wholly concerned with behaving in a reasonable way	Free and equal citizens who are concerned with both pursuing rational interests and behaving in a reasonable way
The hypothetical agreement	Agreement that could not be reasonably rejected by anyone	Agreement that could be justified from both rational and reasonable perspectives
Representatives	Barry, Scanlon	Rawls

Interestingly, Rawls himself also seems to recognize the uniqueness of his model. Although he does not deny the similarity between himself and Scanlon,[10] he consciously keeps his distance from Scanlon throughout his life. In a conversation with Samuel Freeman, Rawls mentioned that he and Scanlon were apparently undertaking two different projects:

> Rawls always referred to justice as fairness as a "contractarian" position. He was opposed to others' use of Scanlon's term "contractualism" as a generic term used to refer to justice as fairness…He said contractualism was Scanlon's own position, and that it was original, distinctive, and, in important respects, quite different from what he was trying to do. (Freeman 2007b, 36)[11]

Freeman believed that the reason Rawls differentiated his position from Scanlon's position was that Rawls was concerned with the problem of congruence, that is, whether "right" is compatible with "good," while Scanlon did not intend to answer this problem. Freeman's interpretation was incorrect here. Scanlon did indeed recognize the problem of congruence, but he answered this problem with his reductive strategy, in contrast with Rawls's approach. Nevertheless, Freeman's story reflects Rawls's awareness that he and Scanlon actually belong to two different strands of contractarianism.[12]

10] Cf. Rawls (2005, 85 n. 33).

11] A similar story can also be found in Martha Nussbaum (2006, 418 n. 9). Rawls describes himself as a social contract theorist but not a Kantian.

12] I owe thanks to an anonymous reviewer for suggesting that Scanlon, though usually named as the representative of Kantian contractarianism, does not accurately mirror Kant's thought. For Scanlon is interested mostly in the foundation of morality and uses the methodology of social contract to define the ideas of rightness and wrongness. Yet Kant reflects on the idea of social contract only at the political, but not moral, level. Therefore one can also doubt how "Kantian" Scanlon is. In fact, in recent years, Scanlon also

IV. RAWLS AND SIDGWICK: SAME QUESTION, DIFFERENT ANSWERS

Ironically, the theorist who has the same view of practical reason as Rawls does is not any one of the Kantian contractarians, but rather Henry Sidgwick, a utilitarian. In this section, I shall introduce Sidgwick's dual conception of practical reason and show that, although Rawls rejects Sidgwick's utilitarianism, he shares his conception of practical reason and understands his works as continuing Sidgwick's unfinished project.

Some scholars have argued that Rawls and Sidgwick have a similar methodology.[13] A distinctive feature of *The Method of Ethics* is that Sidgwick does not manifestly advocate utilitarianism, but rather stands in a neutral position, as impartially as possible, in evaluating different moral theories. Sidgwick (1981, 338-43) defines several neutral standards, such as precision and consistency, as the standards that a moral theory (in Sidgwick's terms, a "method of ethics") must fulfil. He then uses these standards to examine three theories respectively, i.e., utilitarianism, intuitionism and egoism. Sidgwick's methodology is reminiscent of how Rawls argues for justice as fairness. Rawls also presents the original position as an impartial procedure and shows how parties would choose among utilitarianism, intuitionism, perfectionism, as well as justice as fairness. However, I shall further argue that their similarity is not only limited to the formal methodology of comparing theories, but also in more substantial understandings of practical reason.

In comparing intuitionism and utilitarianism, Sidgwick criticizes intuitionism for failing to fulfil the several neutral standards mentioned above. Intuitionism judges the rightness or wrongness of actions by various intuitions. Yet the implications of these intuitions are always imprecise. Also, when intuitions conflict with each other, intuitionism fails to offer a consistent guide for determining what one ought to do. Unlike intuitionism, utilitarianism is precise and consistent -it requires only a universal principle that "one ought to aim at the Ultimate Good on the whole," i.e., maximizing total utility. Nevertheless, the comparison between utilitarianism and egoism does not have a clear result. Egoism is a theory centred upon a universal principle according to which one ought to aim at one's good on the whole, i.e., maximizing one's own utility. Sidgwick takes great pains to admit that the principle of egoism is able to offer implications that are as precise and consistent as the principle of utilitarianism is. Both egoism and utilitarianism can fulfil the neutral standards defined by Sidgwick. Being an egoist is as rational as being a utilitarian is. The requirement of practical reason is thereby two-fold, implying the requirements of "good" and "right":

> There is something that it is reasonable for him to desire, when he considers himself as an independent unit, and something again which he must recognize as reasonably to be desired, when he takes the point of view of a larger whole; the former of

---

begins to acknowledge the different between Kant and him. See his reply to Parfit in Parfit (2011b, 116-39).

13] A recent example is Mandle and Reidy (2017, 773)

> these objects I call his own Ultimate "Good," and the latter Ultimate Good taken universally…I apply the different term "right," to avoid confusion. (Sidgwick, 1981, quoted in Philips 2011, 132)

Sidgwick's conception of practical reason is obviously a dual conception, formed by two principles which aim at self-interest and total utility respectively. Given this conception, one has a good reason to pursue one's own interest, while one has a good reason as well to promote the Ultimate Good impartially. Sidgwick famously claims that, in "a recognized conflict between self-interest and duty, practical reason, being divided against itself, would cease to be a motive on either side" (1981, 508).

Sidgwick is pessimistic about the reconciliation between these two aspects of practical reason. He believes that *The Method of Ethics* successfully shows that, as a moral theory, intuitionism is inferior to utilitarianism. Nevertheless, he unwillingly admits that utilitarianism and egoism are equally valid and there is no good reason to explain why any one of them should be favoured over another. As a work of moral theory, this is a disappointing result. Sidgwick's original purpose was to prove the existence of the "Cosmos of Duty" by systematizing moral reasoning and showing that moral reasoning has a unique rational answer. However, the result of philosophical reflection is that moral reasoning cannot get rid of the internal conflict within practical reason. "[T]he Cosmos of Duty is thus really reduced to a Chaos: and the prolonged effort of the human intellect to frame a perfect ideal of rational conduct is seen to have been foredoomed to inevitable failure" (Sidgwick, 1981, quoted in Crisp 2015, 233). This result of philosophizing severely disappointed Sidgwick such that he once thought about writing a book where the first word would be "ethics" and the last word "failure."[14]

Rawls's philosophical ambition should be understood against the background of Sidgwick's failure. Rawls inherits Sidgwick's dual conception of practical reason, but believes that Sidgwick gives up too early in resolving the conflict between self-interest and duty. *The Method* is an unfinished project. If self-interest and duty are understood properly, then Sidgwick's failure can be avoided. Instead of the duty to maximize total utility, the duty should be understood as one of reasonableness to act in accordance with principles that would be fairly chosen in the original position. Instead of taking whatever personal good as self-interest, self-interest should be defined as rational higher-order interests in developing the two moral powers. Given these two new understandings, duty and self-interest, the two aspects of practical reason can be reconciled. Rawls does not mean that self-interest and duty can always be reconciled, but they can be reconciled in at least one case, namely when free and equal citizens conceive of self-interest and duty in these two ways. Put succinctly, the starting point of Rawls is the end point of Sidgwick; Rawls understands practical reason as Sidgwick does and begins his work attempting to resolve the internal conflict of practical reason that bothered Sidgwick for his whole life.

14] Cf. Crisp (2015, viii).

With this admittedly brief sketch of the relationship between Rawls and Sidgwick, two points are worth noting. First, the dual conception of practical reason held by Rawls is not a weird view of practical reason. Rather, it is a philosophically plausible view that is shared by notable philosophers such as Sidgwick.[15] Secondly, in terms of the conception of practical reason, the similarity between Rawls and Sidgwick is significantly greater than the similarity between Rawls and other Kantian contractarians. In fact, the attempt to reconcile rationality and reasonableness shown in Rawls's and Sidgwick's writings can hardly be found in the writings of any Kantian contractarians. If we distinguish Kantian contractarians from Hobbesian contractarians because they uphold different conceptions of practical reason, then we have little reason not to distinguish Rawls from Kantian contractarians as well.

V. THE "HYBRID" KANT IN RAWLS'S INTERPRETATION

The claim that Rawls is a hybrid contractarian rather than a Kantian contractarian seems shocking, given that Kant is no doubt the philosopher whom Rawls is mostly intellectually indebted to. Rawls mentions his connections with Kant's ideas in Section 40 of *A Theory of Justice* (1999, 221-27). He also names his theory "Kantian constructivism" in his Dewey Lecture in 1980. Furthermore, he once said that "[w]ith Kant I hardly made any criticism at all. My efforts were centred on trying to understand him so as to be able to describe his ideas to students" (Rawls 2000, xvii). However, by reading Rawls's interpretation of Kant in the *Lectures on the History of Moral Philosophy*, it is clear that, in the mind of Rawls, Kant's theory has a strong "hybrid contractarian" character. Like Sidgwick, Kant assumes a dual conception of practical reason. He aims at resolving the internal tension within this conception and even offers a better solution than Sidgwick does.

Throughout the lectures, Rawls emphasizes that Kant's conception of practical reason is two-fold. He understands that Kant's full conception of reason, *vernunfig*, "usually covers being both reasonable and rational" (Rawls 2000, 164).[16] The structure of practical reason that consists of these two powers is regarded by Kant "as implicit in our everyday moral consciousness, the fact of reason" (Rawls 2000, 237). A person with a good will is a person who can be both rational and reasonable. Reasonableness means that one is "judicious, ready to listen to reason, has the sense of being willing to listen to and consider the reasons offered by others" (Rawls 2000, 164). A reasonable person will not force others to accept a proposal that is unfair to others, but rather will use reasons to persuade others to accept this proposal. To Kant, being reasonable implies acting from the moral law, which is the maxims that pass the test of the CI (categorical imperative)-procedure.

15] Examples of contemporary philosophers who advocate a similar form of dual conception are Crisp (2006) and Parfit (2011a, 130-49).

16] See also Rawls (2000, 188, 194, 202).

If a maxim passes the test of the CI-procedure, i.e., we are able to will an "adjusted social world" in which the maxim becomes a universal norm, then it implies that everyone has a reason to act from this maxim. Treating others in accordance with this maxim implies treating others in a way that could be justifiable to others.

Unlike reasonableness, which is about how to treat others, rationality is about how to pursue one's interests. Rawls understands that there are two kinds of rational interests in Kant's writings. First, it is the interest to act from a maxim. Everyone has different situations and desires. This leads to the formulation of different maxims that rationally guide people's actions. Some can pass the CI-procedure, whereas some cannot. A maxim that fails to pass the CI-procedure does not qualify as a moral law and fails to be a conclusive reason for us, despite the fact that it is rational from our perspective. Secondly, there is the interest in the moral law itself. One has a pure practical interest in complying with moral law and achieving good will. This interest motivates people to check maxims by the CI-procedure. Without this interest, we are not bothered with the question of whether a maxim is a moral law.

Kant finds that it is common for the pursuit of the first rational interests to conflict with the requirement of reasonableness. Acting in accordance with a maxim that passes the CI-procedure could be harmful to the advancement of one's rational interests. Rational interests should in that case be abandoned or revised, since reasonableness draws a boundary within which the rational interest is permitted to be pursued. However, this does not imply that reasonableness has to be achieved at the cost of rationality. Unlike the first rational interests, the second rational interest, the pure practical interest in acting from moral law, is compatible with the requirement of reasonableness. Thus, to avoid the conflict between rationality and reasonableness, the second rational interest should have the priority over the first in practical reasoning. "[T]he interests moving our actions are of different kinds and arranged in a certain structure, with the practical interest we take in the moral law itself…always having an *effective regulative priority*" (Rawls 2000, 178, emphasis added). To interpret this distinction using Rawls's terminology, the first rational interest is a normal interest that varies according to one's situation, whereas the second is the higher-order interest that is shared among free and equal citizens.

To Kant, one's action has moral worth only when the following two conditions are met:

> i. When a question of duty is involved, we decide the case entirely by considerations of duty, leaving aside all reasons of interest and inclination;
>
> ii. Our pure practical interest in acting from the considerations of duty is strong enough by itself to ensure that we do as we ought. It is only when our pure practical interest is not strong enough to ensure this, and those other motives are needed for us to act properly, that our will, or moral character, is less than fully good. (Rawls 2000, 204)

These two conditions imply that moral worth is meant to be achieved when rationality and reasonableness can be reconciled with each other. The first condition

represents that one acts reasonably, whereas the second condition represents that one pursues the second rational interest. An ideal moral agent should have two "grounding reasons" to act from moral law, the "justifying reason" to act in a reasonable way as well as the "reason of rationality" to pursue the pure practical interest (Rawls 2000, 166). The former motivates people to respect each person as an end in itself, whereas the latter drives people to realize the ideal of a realm of ends, which is the greatest goodness in human life.

According to Rawls's interpretation, Kant is, similar to Sidgwick and Rawls himself, a theorist who assumes a dual conception of practical reason. The distinctiveness of Rawls's interpretation can be made more explicit if it is compared with other interpretations of Kant. I shall use the interpretation of Paul Guyer as an example. In Guyer's interpretation, Kant assumes a monistic conception of practical reason, which consists of a single principle of rationality as the standard of evaluation. While each action aims at achieving a certain end, how rational an action is depends on how valuable the end is. Most of the ends only have conditional worth; they are worthwhile because they are chosen by rational agents, not because they are valuable in themselves. Therefore, "there is no other candidate for the unconditionally valuable source of conditionally valuable ends than our own capacity to choose those ends, so our capacity of choice must be the very thing that has unconditional value" (Guyer 2006, 189). A rational agent should always aim at fully realizing the capacity to choose ends, i.e., being free.

To be free and to fully realize this capacity, a rational agent should be independent of contingent inclinations. This is then related to Kant's ideas of hypothetical and categorical imperatives. A hypothetical imperative is a maxim that tells you what you must do if you are inclined to attain some end, while a categorical imperative is a maxim from which each of us should act, regardless of what inclinations we have. When an agent conforms to hypothetical imperatives he is still unfree because his choice is contingent on his inclinations. Only when an agent acts merely from duty in accordance with categorical imperatives, is he free because he is "to be governed by a law that [he] gives [himself] rather than to allow [himself] to act on whatever mere inclination happens to be alluring at the moment" (Guyer 2006, 204). In brief, a rational agent should aim to be free, and being free implies acting from duty in accordance with categorical imperatives, which constitute moral law.

When comparing Guyer's version of Kant with Rawls's version, one realizes that the former is not so much concerned with the conflict between the duty to moral law and the rational pursuit of self-interest, which is a conflict taken by Rawls's version of Kant as a fundamental philosophical problem which remains to be solved. If a person breaches the moral law for the sake of pursuing self-interest, then she/he is irrational. There is not much a philosopher could say in this case except to ask this person to be rational. Acting from moral law cannot necessarily be reconciled with the pursuit of self-interest. Practical reason has only one standard -acting from categorical imperatives for the sake of freedom. One is rational when this standard is fulfilled. The way in which Guyer's Kant

is uninterested in the conflict between duty and self-interest is shown in the following passage:

> The fundamental principle of morality, Kant has claimed, must be unconditionally valid for any rational being. If any being were perfectly rational, it would automatically act in accordance with this law, and the law would therefore not appear to be a constraint. But we human beings are not perfectly rational, and thus although we recognize the unconditional validity of the moral law, it also appears as a constraint to us, something that may be in conflict with our irrational side… ….But *Kant takes it to be obvious and not in need of any special argument that we will often experience the stringent demands of morality as a constraint*; thus, although his arguments are aimed at a derivation of the categorical imperative, all of his effort is aimed at demonstrating the content of the fundamental principle of morality and proving that it is valid or binding for us, not at reminding us that we often experience that validity as a constraint. (Guyer 2006, 179-80, emphasis added)

Here I do not intend to argue whether Guyer's interpretation or Rawls's interpretation are correct. What I want to show is only that Rawls has a distinctive interpretation of Kant that describes him as assuming a dual conception of practical reason, which is different from the interpretation of other Kant scholars. Hence, when Rawls claims that he is intellectually affiliated with Kant, we should not understand this as meaning that Rawls accepts the monistic conception of practical reason that is commonly understood to be held by Kant. Rather, Rawls has in mind that Kant thinks the same as him in endorsing the dual conception of practical reason, thus belonging to a tradition of hybrid contractarianism.[17]

VI. CONCLUSION

Labels matter in some ways and do not matter in others. On the one hand, labels do not matter in that a philosopher usually has multiple labels. Mill has been labelled as a utilitarian, a socialist, a libertarian, or a perfectionist (Edwards and Townshend 2002, 180-97). A single label can hardly tell us all the substantial beliefs of that philosopher. Nevertheless, labels matter because labels usually determine our first impression of how a philosopher is linked with other philosophers in the history of thought. A label represents a strand of thought in which philosophers share certain core beliefs. Some are recognised as forerunners who first emphasize the importance of these beliefs, while some are followers who are inspired to construct a more perfect theory in this strand of

17] I owe thanks to an anonymous reviewer for indicating how Rawls's interpretation of Kant can be made sense. To construct a Kantian political theory, one may not directly apply the theory of categorical imperative to the realm of political action. In the political domain, political principles not only have the negative dimension (you shall not! / i.e. refraining), but also have the positive dimension (you shall! / i.e. acting). The theory of categorical imperative only considers the former dimension. To take the latter dimension into account, a Kantian political theory should explain how people are rationally motivated to act in certain ways that support just institutions to work efficiently. This may explain why Rawls constructs a contractarian model that is hybrid in nature.

thought. A label affects how we understand the position and contribution of a philosopher in the history of thought. A wrong label may lead us to overestimate or underestimate the contribution of a philosopher.

In this sense, Rawls is more or less the victim of a wrong label. Rawls is usually conceived as the father of Kantian contractarianism in contemporary political philosophy, who revives a strand of thought traced back to Kant. Despite his huge influence, Rawls's contractarian model is sometimes regarded as a flawed project. An example is Barry's critique of Rawls, namely that Rawls is concerned too much with whether justice is beneficial to each party (Barry 1989, 242-54). Barry believes that the part of Rawls's theory which proves how justice is in the rational interest of people is redundant. Therefore, Rawls's theory should be further improved to a more "pure" version of Kantian contractarianism, which is concerned only with whether justice is impartially justified and has no interest in showing that it is beneficial to people. If we look from the perspective of Kantian contractarians and assume Barry's monistic conception of practical reason, then the part of proving that justice is rationally good is truly redundant. However, as I argued in this paper, Rawls in fact assumes a dual conception of practical reason that is significantly different from the monistic conception held by Barry and Scanlon. Instead of treating Rawls as a "inspiring but deficient forerunner" of Kantian contractarianism, he should be seen as a contractarian who belongs to a distinctive strand of contractarianism that should be distinguished from Kantian contractarianism. By labelling Rawls as a hybrid contractarian, his relationship with other contractarians as well as his contribution to the history of contractarianism can be better understood.[18]

bwwong@cuhk.edu.hk

REFERENCES

Audard, Catherine. 2006. *John Rawls*. Stockfields: Acumen Publishing.
Barry, Brian. 1989. *Theories of Justice*. California: University of California Press.
— — —. 1995. *Justice as Impartiality*. Oxford: Clarendon Press.
Brown, Henry Phelps. 1988. *Egalitarianism and the Generation of Inequality*. Oxford: Clarendon Press.
Crisp, Roger. 2006. *Reasons and the Good*. Oxford: Clarendon Press.
— — —. 2015. *The Cosmos of Duty: Henry Sidgwick's Method of Ethics*. Oxford: Clarendon Press.
Darwall, Stephen. 2003. Introduction. In *Contractarianism / Contractualism*, edited by Stephen Darwall, 1-8. Malden, MA: Blackwell Publishing.
Dreben, Burton. 2003. On Rawls and Political Liberalism. In *The Cambridge Companion to Rawls*, edited by Samuel Freeman, 316-46. New York: Cambridge University Press.
Edwards, Alistair, and Jules Townshend, eds. 2002. *Interpreting Modern Political Philosophy: From Machiavelli to Marx*. New York: Palgrave Macmillan.

18] I am grateful for valuable advices and comments from David Gauthier, Paul Kelly, Chandran Kukathas, Matt Matravers, Christopher Morris, Geoffrey Sayre-McCord, Jonathan Quong and Albert Weale. I also wish to thank all those who commented on the earlier versions of this paper in the Political Theory Doctoral Workshops at LSE and the MANCEPT Workshop in Political Theory.

Freeman, Samuel. 2003. Congruence and the Good of Justice. In *The Cambridge Companion to Rawls*, edited by Samuel Freeman, 277-315. New York: Cambridge University Press.
— — —. 2007a. *Justice and the Social Contract*. New York: Oxford University Press.
— — —. 2007b. The Burdens of Public Justification: Constructivism, Contractualism, and Publicity. *Politics, Philosophy & Economics* 6(1): 5-43.
— — —. 2007c. *Rawls*. New York: Routledge.
Gauthier, David. 1986. *Morals by Agreement*. Oxford: Clarendon Press.
Guyer, Paul. 2006. *Kant*. London: Routledge.
Hampton, Jean. 2007. Contract and Consent. In *A Companion to Contemporary Political Philosophy*, (2nd edition), edited by Robert E. Gordin, Philip Pettit and Thomas Pogge, 478-92. Oxford: Blackwell Publishing.
Kumar, Rahul. 1999. Defending the Moral Moderate: Contractualism and Common Sense. *Philosophy and Public Affairs* 28(4): 275-309.
Kymlicka, Will. 1993. The Social Contract Tradition. In *A Companion to Ethics*, edited by Peter Singer, 186-96. Oxford: Blackwell Publishing.
Maffetone, Sebastiano. 2010. *Rawls: An Introduction*. Cambridge: Polity Press.
Mandle, Jon. 2009. *Rawls's A Theory of Justice: An Introduction*. Cambridge: Cambridge University Press.
Mandle, Jon, and David A. Reidy, eds. 2017. *The Cambridge Rawls Lexicon*. Cambridge: Cambridge University Press.
Nussbaum, Martha. 2006. *Frontiers of Justice*. Cambridge, Mass.: Belknap Press.
O'Neill, Onora. 2004. Constructivism VS. Contractarianism. In *On What We Owe to Each Other*, edited by Philip Stratton-Lake, 19-31. Oxford: Blackwell Publishing.
Parfit, Derek. 2011a. *On What Matters, Vol. 1*. Oxford: Oxford University Press.
— — —. 2011b. *On What Matters, Vol. 2*. Oxford: Oxford University Press.
Philips, David. 2011. *Sidgwickian Ethics*. Oxford: Oxford University Press.
Pogge, Thomas. 2007. *John Rawls: His Life and Theory of Justice*. Oxford: Oxford University Press.
Rawls, John. 1980. Kantian Constructivism in Moral Theory. In John Rawls, *Collected Papers*, edited by Samuel Freeman, 303-58. Cambridge, MA: Harvard University Press.
— — —. 1999. *A Theory of Justice*, revised edition. Cambridge, MA: Harvard University Press.
— — —. 2000. *Lectures on the History of Moral Philosophy*, edited by Barbara Hermen. Cambridge, MA: Harvard University Press.
— — —. 2001. *Justice as Fairness: A Restatement*, edited by Erin Kelly Cambridge, MA: Harvard University Press.
— — —. 2005. *Political Liberalism*, expanded edition. New York: Columbia University Press.
Richards, D. A. J. 1971. *A Theory of Reason for Action*. Oxford: Clarendon Press.
Sayre-McCord, Geoffrey. 2013. Contractarianism. In *The Blackwell Guide to Ethical Theory* (2nd edition), edited by Hugh LaFollette and Ingmar Persson, 247-67. Malden, MA: Blackwell Publishing.
Sidgwick, Henry. 1981. *The Method of Ethics* (7th edition). London: Hackett.
Scanlon, T. M. 1982. Contractualism and Utilitarianism. In *Utilitarianism and Beyond*, edited by Amaryta Sen and Bernard Williams, 103-128. Cambridge: Cambridge University Press.
— — —. 1998. *What We Owe to Each Other*. Cambridge, MA: The Belknap Press of Harvard University Press.
— — —. 2014. *Being Realistic about Reasons*. Oxford: Oxford University Press.
Southwood, Nicholas. 2010. *Contractualism and the Foundation of Morality*. Oxford: Oxford University Press.

Public Reason 8 (1-2): 51-64

Why Dignity is not the Foundation of Human Rights

Stamatina Liosi
University of Kent

Abstract: This essay questions what is argued by many scholars today, namely that the moral concept of human dignity provides the basis for the establishment of human rights. More specifically, I critically discuss the two most prominent conceptions of human dignity, the 'status' and the 'value' (coming from the Catholic and the Kantian traditions) conceptions of dignity, which are suggested today as the foundations of human rights (sections I and II). Ultimately, I propose a different, 'duty-based' philosophical account for the justification of the latter (section III).

Key words: Dignity, justification, human rights, status, value, duties.

Since 1948, human rights have been widely accepted and ratified by most countries. However, a great number of human rights violations and abuses still occur worldwide. Even if there is a broad human rights' framework, the latter is considered to be ineffective. One of the main reasons of its ineffectiveness is the fact that in the Universal Declaration of Human Rights (UDHR), in the International Covenant on Civil and Political Rights (ICCPR), and in the International Covenant on Economic Social and Cultural Rights (ICESC), rights are described in abstract terms. But for human rights to be respected they have to be clear and concrete. I argue that the formulation of rights by the drafters of the major human rights documents should be supplemented by their philosophical establishment. It is argued by many scholars today that human dignity provides the basis for the establishment of human rights. In this essay, I question the claim that human dignity is the foundation of human rights. More specifically, I critically discuss the two most prominent conceptions of human dignity, namely the 'status' and the 'value' (coming from the Catholic and the Kantian traditions) conceptions of dignity, which are suggested today as the foundations of human rights (sections I and II). Ultimately, I propose a different, 'duty-based' philosophical account for the justification of the latter (section III).

I. THE STATUS CONCEPTION OF DIGNITY AS THE FOUNDATION OF HUMAN RIGHTS

1. Jeremy Waldron's Account

One of the most prominent dignitarian accounts for the justification of human rights today is status-based. Jeremy Waldron is one of its main supporters. In his Tanner lectures *Dignity and Rank* and *Law, Dignity, and Self-Control* (2012); in his essays: *Citizenship and Dignity* (2013), *The Image of God: Rights, Reason, and Order* (2010a), and *Is Dignity the Foundation of Human Rights* (2015); and in his paper *Dignity, Rights, and Responsibilities* (2010b), Waldron argues that dignity as a legal status may be seen as the genuine basis for

the derivation of our rights. In what follows, I briefly present the main points of his theory as it is developed in the previous works.

Initially, Waldron claims that given current failures and disagreements regarding human rights, we must examine them from scratch in order to understand them in depth (2010b). He then claims that to begin with morality is not the best way to understand human rights; rather we should start from the *law* of human rights. That is to say, if we want to see what human rights are, we should just focus on human rights law, namely the Declarations and Conventions – not on the moral idea underlying them. According to Waldron, the best way to understand human rights law is to start by examining its official justification, human dignity, treating it as a legal rather than a moral concept in the first instance (2010a; 2012).

More specifically, Waldron regards dignity as a type of status[1] traditionally connected with rank (2012, 14, 17, 21, 22, 24, 30; 2015, 133; 2013, 336)[2]. He writes: "My own view of dignity is that we should contrive to keep faith somehow with its ancient connection to noble rank or high office" (2012, 30). In his Tanner Lecture, *Law, dignity, and self control*, Waldron defines status as a "high-ranking status, high enough to be termed a *dignity*" (2012, 59). He claims then that to treat someone as dignified is to treat her as royalty (2010b); hence one whose dignity is not respected today should be seen as a prince or a duke who had not been respected in the past (2010b, 34-36).

In particular, Waldron argues that dignity is a *legal* status, namely something accorded to people by law. A legal status in law is, in principle, as Waldron says, a package of rights and duties accorded to a person or to a group of persons by law (2015, 134). For example, infancy is a legal status. Yet, Waldron considers a legal status not just as a given package of rights and duties, but as a *special* package that also entails a deeper explanation of *why* these rights and duties are accorded to persons (2015, 135). That is to say, a legal status, as Waldron understands it, can explain *why* rights and duties are attributed to

1] The status conception of dignity comes from the so called meritocratic dignity. That is a kind of dignity which was protagonist in archaic societies (750-479 BC). For example, in the Homeric epics dignity is accorded to persons with high status. The latter is the status arising from the possession of characteristics regarded as meritorious. Special characteristics such as family and friendships determine where one stands. Dignity, in this sense, is determined by the *place* people hold in the social hierarchy, and is ascribed only to those of high rank. We also see the status conception of dignity (*dignitas*) in the Roman world in which it is similarly associated with the respect and honor due to a person with an important position. For instance, in Cicero's *De Inventione*, dignity denotes one's role in the society, as well as the honors and the respectful treatment that are due whoever has this role. One possesses an elevated social status in the social order if one belongs, for example, to the nobility or to the church. The status conception of dignity continued to exist until the Enlightenment, in the 18th century. Then it started gradually to fade away, as a result of the abolition of aristocratic 'dignities' associated with aristocratic 'status'. But after years of latency, dignity as a status seems to have a sudden revival nowadays especially within the context of the human rights discourses. Waldron is today one of its main supporters. See, for example, Cicero, M. T. (84BC). De *Inventione*. Available at http://www.classicpersuasion.org/pw/cicero/dnv2-1.htm (accessed 24 October, 2017).

2] Waldron seems to have been influenced by Gregory Vlastos, a great classical scholar, who has argued that we organize ourselves like an aristocratic or caste society; see (Vlastos 1984, 41).

citizens by law. Eventually, this kind of dignity, namely dignity as a legal status, might be seen, as Waldron claims, as the foundation of human rights (2015, 133-36; 2012, 14).

Ultimately, Waldron sees the above legal status conception of dignity as perfectly combined with equality (2012, 14, 31, 33). That is to say, he claims that *all* people in modern democracies possess it. More specifically, he points to the Kantian passage in the Metaphysics of Morals (Ak. 6:329) according to which: "no human being in a state can be without any dignity, since he at least has the dignity of a citizen" (Kant 1996, 471). Waldron argues then that dignity as a legal status is combined with equality in the case of the "dignity of citizenship". He then claims that the legal status of the dignity of 'citizenship' can legitimately be proposed as the foundation of the rights of *all* citizens in modern democracies (2012, 60; 2013, 327-43). Recently, Waldron has examined the extension of the notion of legal citizenship from the domestic to the global level. To support this, he invokes the Kantian 'universal citizenship', which he understands as a "realization of the dignity of moral being" (2013, 332; 1996, 281).

2. Waldron's Claims Refuted

According to the above analysis, these are Waldron's main claims:

1. Because of the disagreements regarding human rights today, we must reconsider their grounds.

2. We should begin from human rights law, not its moral background, in order to understand the former.

3. We should start by examining the official justification of human rights law, that is, human dignity treating it as a legal rather than a moral concept in the first instance.

4. Dignity is a type of status traditionally connected with rank.

5. In particular, dignity as a *legal* status may guarantee our human rights.

6. The legal status conception of dignity can be combined with equality, in the sense that *all* citizens in modern democracies are considered to be dignified.

7. The extension of the notion of legal citizenship from the domestic to the global level through the Kantian 'universal citizenship' is feasible.

In this section, I respond to Waldron's claims, explaining why dignity as 'legal status' is not the foundation of human rights.

Initially, Waldron is right that the controversies regarding human rights law today force us to understand it in depth. For instance, many argue that human rights law is just a Western concept imposed by western cultural circles on non-western states, with other values, customs and mores, only for serving West's special interests. Hence, we should understand the *nature* of human rights law in order to cope with such disagreements. Surely, the need for the determination of the nature of human rights leads us to the examination of their grounds. However, I disagree with Waldron's claim that we must ignore morality, and start from law, in order to understand the latter. Here are two reasons: First, human rights law is not detached from the *moral* idea of human rights. I see the latter

as moral rights, with political connotations, which are protected by law. Consequently, in order to understand the *law* of human rights, we must not separate it from its *moral* nature and grounds. Incidentally, focusing on morality in order to understand human rights law, and not vice versa – as for example Hohfeld (1919) did and Waldron suggests – is, I think, a significant philosophical stance that brings philosophy at the heart of human rights' discourse connecting the two disciplines.

Secondly, let's say that we follow Waldron and posit human rights law as our starting point. The issue here is whether the law of human rights can effectively answer the question of what human rights are. Unfortunately, in most cases today there is a large controversy regarding the interpretation of our fundamental human rights. One example is the interpretation of the right to freedom of religion initiated by the landmark case of the ECJ, that is, Kokkinakis v. Greece (1993)[3]. As Rivers says: "There is no doubt in the minds of the Strasbourg judges that freedom of religion is important, but exactly what it is and why it matters remains elusive" (2013, 405). It seems then that the law itself – the article 18 of the UDHR in this case – does not provide the grounds for a deep understanding of the relevant right. But there must be an exact and correct interpretation of rights because without prior interpretation, there cannot be proper application of law (2013, 381). Hence I think we must not separate the law of human rights from morality. Only an investigation of the latter could enrich our understanding of the former.

Moreover, the fact that dignity seems to be the official justification of human rights in the major documents (UDHR, ICCPR, ICESCR), is not sufficient reason for us to conclude that dignity is *actually* the foundation of human rights. The meaning of dignity is not clearly delineated within the context of the major human rights documents. For instance, dignity does appear as the moral basis of human rights in the ICCPR, in which it is written that the rights it contains "derive from the inherent dignity of the human person"[4]. But, dignity does not appear as a straightforward foundational concept in the *Universal Declaration of Human Rights*, in which it is written that "all human beings are born free and equal in dignity and rights" (UN General Assembly, 1948). Hence, given the great confusion regarding the meaning of dignity in the fundamental human rights law, I think we should not focus on the latter to shed light on the obscure concept of human dignity. Incidentally, Waldron himself admits this claiming that law is not noted for its philosophical rigor (2015, 118). Hence, it is unclear why he still insists that the legal rather than the moral meaning of dignity should be our starting point (2012, 13-15; 2015, 118, 121, 123).

In addition, we must bear in mind that even though the legal concept of dignity is 'autonomous', it is still a rich concept with deep philosophical and theological roots.

3] See Human Rights Constitutional Right. 1993. *Kokkinakis v. Greece*. http://www.hrcr.org/safrica/religion/Kokkinakis.html (accessed 23 October, 2017).

4] See United Nations Human Rights. 1966. *International Covenant on Civil and Political Rights.* http://www.ohchr.org/EN/ProfessionalInterest/Pages/CCPR.aspx (accessed 26 October, 2017).

Thus, if we want to understand its legal meaning, we should not ignore, but examine in depth the moral theology (e.g. Aquinas), the natural law theory (e.g. Hobbes, Pufendorf, Rousseau), and of course the Kantian moral philosophy in the background (2015, 124). Hence, I think one should go back to dignity's *moral* roots treating it as a *moral* idea in the first instance (Waldron 2013, 383). However, this does not mean that we won't be able to return to the legal concept of dignity.

Furthermore, I disagree with the idea of grounding human rights in dignity understood as a status with references to aristocracy. I regard this anachronistic idea that cannot be applied in the post-Enlightenment world. Incidentally, we must not ignore the fact that there are many examples in history which show that not all princes and dukes had high status[5].

Also, I disagree with Waldron's connection of the legal status conception of dignity with human rights. This connection might lead one to the unreasonable claim that in political systems in which, say, legal status (i.e. dignity) is conferred only to men but not to women (e.g. in Saudi Arabia), only men's human rights are justified. But this would be an unacceptable claim. Therefore, I argue that dignity as legal status and human rights should be kept distinct from one another.

Further, I do not understand how Waldron's notion of the 'dignity of citizenship' can be the foundation of human rights. Here, I see four problems: 1) The dignity of citizenship which Waldron suggests refers by definition only to the citizens of democratic countries excluding all other people residing in it, for example the immigrants, refugees, and 'apatrides'; 2) it does not include the citizens who live in countries which have not the characteristics of a democracy (e.g. North Korea is a democratic republic only on paper); 3) it does not include all those who live in non-democratic countries (e.g. Qatar, Saudi Arabia); 4) it does not include those who still live in isolated jungle tribes in the world. Therefore, given that citizenship is a status not assigned universally, it cannot support any argument for the justification of human rights, namely the rights of men, not citizens (Waldron 2010b).

Finally, I cannot but disagree with the extension of 'legal citizenship' from the domestic to the global level via the Kantian notion of 'universal citizenship'. Contrary to what Waldron argues, I do not think Kant favors such a citizenship of the world. Even though Kant speaks about "the relation of theory to practice […] from a cosmopolitan perspective", surely he does not favor the notion of a universal citizenship (Kant 1996, 281). A universal citizenship presupposes a world or global state which Kant clearly rejects. For instance, in 8:367 in *Perpetual Peace* he writes:

> […] the *separation* of many neighboring states independent of one another […] is nevertheless better, in accordance with the idea of reason, than the fusion of them

5] See for instance the *Earl of Shrewsbury's Case*, 12 Co. Rep. 106, 77 Eng. Rep. 1383 (1612) cites the terms of an act of Parliament in the reign of Edward IV for the formal degradation of George Nevill, Duke of Bedford.

> by one power overgrowing the rest and passing into a universal monarchy, since as the range of government expands laws progressively lose their vigor, and a soulless despotism, after it has destroyed the seed of good, finally deteriorates into anarchy (Kant 1996, 336)[6].

Consequently, the rejection by Kant of the possibility of a world government makes it impossible further to be argued that there could be a kind of world citizenship. Hence, Waldron's conception of the dignity of the citizens of the world is not feasible[7].

II. THE VALUE (CATHOLIC AND KANTIAN) CONCEPTIONS OF DIGNITY AS THE FOUNDATION OF HUMAN RIGHTS

Apart from the status-conception of dignity, there are also two other significant value-conceptions of dignity: the Catholic and the Kantian. They are too suggested as the grounds of human rights. In this section I critically discuss both of them starting from the Catholic conception of human dignity.

1. The Catholic Value Conception of Dignity as the Foundation of Human Rights

According to the Catholic notion of the dignity of humanity, with roots in Stoic ideas, each and every person has intrinsic *value* (dignity), because he or she incarnates God's image (*Imago Dei*) (Waldron 2010b; Berman 2008). Aquinas defines dignity as the goodness of something on account of itself, that is, as the intrinsic value of something. This Catholic conception of dignity is very often used in the human rights discourse today. Recently, it has been argued that dignity in the Catholic sense grounds the rights of every human being, including embryos, even from the moment of conception (Ruston 2004, 10-12; Aquinas 1945, 220, 223; Lee and George 2008)[8].

Contrary to those who argue that the Catholic value conception of dignity is the foundation of human rights, my claim is that the Catholic and any other religious understanding of dignity is an inappropriate basis for human rights. Here are two reasons: firstly, as Jeremy Waldron indicates, we must not ignore the fact that many people are atheists, namely persons who do not share a religious worldview (2010a). There are also many people who are followers of other religious traditions than the Jewish-Christian one, or who are committed to an approach to rights that favors a multi-faith society. Waldron

---

6] For the rejection of the possibility of a world government by Kant, see also 8: 354 and 8: 357.

7] Similarly to Kant, I argue that serious problems would arise from the creation of a global political entity. There would be not only practical problems regarding the cohesion between the different political entities, but also the crucial problem of despotism as derived from a 'Leviathan-type' global governance. Thus, I think states should better exist autonomously in a soft-cosmopolitan model.

8] See, also, John Paul II, Original Unity of Man and Woman: Catechesis on the Book of Genesis, 1981, cited in Coughlin J.J. 2003. Pope John Paul II and the Dignity of the Human Being. *Harvard Journal of Law & Public Policy*, 27 (65): 72-4; Pope John Paul II's encyclical *Evangelium Vitae*, March 25, 1995, available at: http://www.vatican.va/holy_father/john_paul_ii/encyclicals/documents/hf_jpii_enc_25031995_evangelium-vitae_en.html (accessed 20 October, 2017).

argues then in Rawlsian terms that if we wanted to ground human rights for *all*, we should not ground them in a concept that cannot by definition be shared by all. Incidentally, the grounding of human rights in dignity interpreted via a comprehensive dogma might lead one, for example, to torture another person on the basis of the religious deviation of the latter, and her alleged lack of dignity. But surely this would be an unreasonable, unacceptable, and punishable criminal act.

Secondy, dignity understood as the value attributed indiscriminately to all human beings including embryos from the moment of conception might lead to the unreasonable claim that abortion is impermissible even in serious pregnancy complications, e.g. eclampsia, in which the mother's life is threatened. Additionally, the association of the Catholic value conception of dignity with human rights contradicts certain women's rights such as their right to reproductive freedom. Therefore, I can hardly see how a religious understanding of the moral concept of human dignity, and human rights can be compatible (Waldron 2010a, 216-35).

2. The Kantian Value Conception of Dignity as the Foundation of Human Rights

But the Catholic is not the only value-conception of dignity suggested today as the foundation of human rights. There is also the Kantian moral concept of human dignity which is very often presented as the genuine basis of rights[9]. In this section I critically discuss the Kantian human dignity as the foundation of human rights.

To begin with, Kant discusses thoroughly the dignity of humanity in the first moral work of his critical period, namely in the *Groundwork of the Metaphysics of Morals*. Here is Kant's famous passage concerning human dignity (Ak. 4:434-35):

> In the kingdom of ends everything has either a *price* or a *dignity*. What has a price can be replaced by something else as its *equivalent;* what on the other hand is raised above all price and therefore admits of no equivalent has a dignity/What is related to general human inclinations and needs has a *market price;* that which, even without presupposing a need, conforms with a certain taste, that is, with delight in the mere purposeless play of our mental powers, has a *fancy price*; but that which constitutes the condition under which alone something can be an end in itself has not merely a relative worth, that is, a price, but an inner worth, that is, *dignity*./Now, morality is the condition under which alone a rational being can be an end in itself, since only through this is it possible to be a lawgiving member in the kingdom of ends. Hence, morality, and humanity insofar as it is capable of morality, is that which alone has dignity (1996, 88)

Also, in 4:436, Kant argues that every rational human being has intrinsic value (dignity), because he or she has autonomous will. Kant writes:

9] The Kantian value conception of human dignity is invoked as the foundation of human rights in several cases in courts today; see for example The German Airliner Case, available at: http://www.bverfg.de/e/rs20060215_1bvr035705en.html (accessed 21 October, 2017).

> But the lawgiving itself, which determines all worth, must for that very reason have dignity, that is, an unconditional, incomparable worth; and the word *respect* alone provides a becoming expression for the estimate of it that a rational being must give. *Autonomy* is therefore the ground of the dignity of human nature and of every rational nature (1996, 85).

Now, this Kantian value concept of dignity is proposed as the foundation of human rights. That is to say, Kant is understood as saying that we all have rights, namely entitlements to be respected and treated as ends, not as means (4:429)10, in virtue of the fact that we all possess an intrinsic value or worth, that is, dignity, which is further based on autonomy. Eventually, from our autonomy, namely our capacity to make our own decisions independently, without guidance or coercion from others, our dignity is derived, in which, through the Formula of Humanity (4:429), our human rights are further grounded (Habermas 2010, 464-80; Griffin 2009, 50; Bennett 2015, 76). Within this context, Kant is understood as claiming, for example, that slavery is wrong because it is incompatible with the idea that human beings, and all other rational beings, have to be treated as ends, namely as persons with dignity and autonomy; or that torture is wrong because the tortured man is not being treated as an end, namely as a dignified person with autonomy.

Nevertheless, despite its resonance, I see the above argument for the justification of human rights based on the Kantian moral concept of human dignity as mistaken. As it has been mentioned above, in 4:434-6 in the Groundwork of the Metaphysics of Morals, Kant explicitly says that human dignity is a value grounded in autonomy. In 4:440, he defines autonomy (of the will) as "the property of the will by which it is a law to itself (independently of any property of the objects of volition)…" (1996, 89). Hence, according to Kant, one is considered as a dignified person in virtue of her autonomy or autonomy of the will. Now, it might be argued that human and other rational beings possess human rights in virtue of the fact that they are persons who have to be treated as ends, that is, with dignity because they are autonomous, that is, capable of being independent and free to make their own decisions without supervision or coercion from others.

Contrary to the above interpretation of the Kantian moral principle of autonomy – on which the dignity-based Kantian argument for the justification of human rights is based – I see the Kantian autonomous person not just as a free and independent person who does as she pleases, but as a freely self-legislating person who, applying reason to her inclinations, passions, social conventions, religious beliefs, ideologies etc., does her (moral and legal) duty for duty's sake. Under this definition of autonomy, I see the Kantian human dignity, which is based on it, as the value of those who are eventually capable of doing their duties. That is to say, the dignified person is not the person who is simply independent and

10] According to Kant's 'Formula of Humanity': "So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an *end*, never merely as a *means*" (1996, 80).

free from coercion, but the person who does her duty. Consequently, the Kantian human dignity is a value possessed not by the right-holders, but the duty-bearers.

Therefore, I consider the Kantian argument, according to which dignity is the foundation of human rights, as incorrect, because it is constructed upon the mistaken idea that dignity is a value possessed by right-holders who have the right to being treated as ends in virtue of the fact that they are autonomous, namely free and independent to act as they see fit. But dignity is actually a value attributed to autonomous duty-bearers who do their duty treating *others* as ends, that is, with respect. Hence, human beings are respected not because they have dignity as rational agents, but because others, namely those who (must) respect them, have dignity. Within this context, slavery and torture are wrong not because they are incompatible with the idea that slaves and the tortured persons have to be treated as ends, namely as dignified and autonomous persons themselves; rather such transgressions are wrong because they are irrational (decisions and) acts performed by heteronomous and non-dignified persons who fail to do their duties towards others. Eventually, I argue that if we simply respected others because of their dignity/autonomy/rationality, then babies, children, the insane, the mentally disabled, animals, and plants, which lack all these properties, would be left unprotected.

In addition, I do not consider the dignity-based Kantian argument for the justification of human rights as a *truly* Kantian *argument* given that it disregards the centrality and priority of the concept of 'duty' over the concept of 'right' within the Kantian duty-based ethics in general. Autonomy and duties are concepts which cannot be seen as detached from each other; rather they are concepts very closely related. My Kantian argument for the justification of human rights below has been constructed upon the idea that duties have full priority over rights. If we focus on rights, duties are not guaranteed; but if we focus on duties, then this, even if it is not politically 'catchpenny', surely it is legally more efficient. I think such a legal efficiency is needed more than ever in global justice today.

Finally, I would like to mention one more objection to the idea that the Kantian human dignity is the foundation of human rights. Any Kantian dignity-based argument for the justification of human rights disregards the fact that the Kantian moral concept of human dignity does not apply to the *external* domain of law. Rather, it is a moral concept belonging exclusively to the *internal* domain of morality. But if we restricted our 'justificatory horizons' only to the latter, without any recourse to the external domain of law, in which human rights typically reside, how could we plausibly and legitimately justify the latter?

III. CONCLUSION: A NEW PHILOSOPHICAL ACCOUNT FOR THE JUSTIFICATION OF HUMAN RIGHTS

Given that the status conception of dignity seems to protect only the rights of the privileged few, and promote 'meritocratic' societies and Court rooms[11]; and further in

11] See for example the recent 'Lavinia Woodward case': http://www.telegraph.co.uk/

view of the fact that the Kantian dignity-based argument for the justification of human rights is based on a mistaken interpretation of the moral concept of human dignity, I do not regard either the status or the value concept of dignity as appropriate grounds for human rights. In this section, I suggest a different Kantian, duty-based justification of human rights inspired by Onora O'Neill's theory of duty (1996). My starting point is Kant's supreme principle of morality, autonomy, as it is presented in the *Groundwork of the Metaphysics of Morals* (1996, 89).

Autonomy is the property of the will by which it is a law to itself (lawgiving) (Kant 1996, 89). Hence, autonomy is about lawgiving and further obeying our duties constraining our will in accordance with moral law's commands despite any other inclinations we might have. Hence autonomy is not simply identified with sheer independence or freedom from coercion. The question now is what that lawgiving function of morality is. In the *Metaphysics of Morals*, Kant distinguishes between two types of 'lawgiving'. On the one hand, there is the ethical lawgiving, "which makes an action a duty and also makes this duty the incentive" (1996, 218-21); and on the other hand, there is the juridical lawgiving "which does not include the incentive of duty in the law and so admits an incentive other than the idea of duty itself" (1996, 218-19). The juridical lawgiving, as Kant argues, refers only to external duties, while the ethical lawgiving refers both to ethical internal/non-enforceable duties of virtue and ethical external/enforceable duties of right (1996, 383-4).

Yet, Kant says that, in the case in which the ethical lawgiving takes up external duties, the internal incentive to action, that is, the idea of duty must not be present in the relevant lawgiving[12]. That is to say, by removing the condition of the ethical motivation (in the external domain of law), and allowing for externality, we can derive *external* duties from the *ethical* lawgiving (Baiasu 2016). Eventually, the ethical lawgiving can be considered as the 'interface' between Kant's moral and his legal and political philosophy.

Further, the ethical external/enforceable duties, derived from the ethical lawgiving function of morality, are divided into four main categories[13]: 1) universal perfect duties of right, that is, duties held by all and owed to all; 2) specific perfect duties of right, that is, duties held by some and owed to specified others; 3) universal imperfect duties of virtue, that is, duties held by all, yet they are owed to none; and 4) specific imperfect duties of virtue, that is, duties held by some and owed to none. Specifically, the two first types of duties (universal perfect and specific perfect), as Kant argues, "[…] the capacity for putting others under obligation, that is, the concept of a right can afterwards be explicated" (Kant 6: 239) – while from the 3) and 4) no rights derive (O'Neill 1996, 152). What must be pointed out here is that the word 'develop' is a better choice, rather than Mary Gregor's word 'explicate' because: 1) the original Kantian word 'entwickelt' is closer to the English

news/2017/05/16/oxford-student-spared-jail-extraordinary-talent/ (accessed 24 October, 2017).

12] Incidentally, this is the main difference between ethical and juridical lawgiving (Kant 1996, 384).

13] I am following here Kant's schema of duties of right in 6:240-42, (1996, 395-97); also O'Neill's diagram of the four Kantian types of duty or obligation (1996, 152).

word 'develop' rather than Gregor's 'explicate', which simply means 'correspondence' or 'correlation' or 'connection', and 2) the word 'develop' expresses a deeper foundational relation between duties and rights. The latter are developed, or generated, or derive from the former. It is not accident that Kant talks about 'duties of right'. The word 'right' is itself entailed in a 'duty of right'. Eventually, I argue that **human rights are generated from** (see above, the division of the external duties of right, number 1) **the 'universal perfect duties of right'**[14].

According to this derivation of human rights from duties, one understands why rights are both moral (because they derive from ethical duties) and legal (because they derive specifically from the external ethical duties of right). Incidentally, only human rights, along with the so-called socioeconomic rights derive from morality – not other rights. Kant explicitly denies such a possibility. In 6:396 in the *Metaphysics of Morals*, he argues that the principle of right as analytic is not related to or derived from the principle of morality which is synthetic. Apparently, here Kant means rights in the strict sense such as those which he defines in 6:232. I see the strictly juridical lawgiving as the source of the latter. A right in the narrow sense is for example the right of the seller of a property in England to use the buyer's deposit between exchange of contracts and completion, in order to purchase another property. The latter is not a moral human or socioeconomic right such as those mentioned in the major human rights documents. Rather, it is a strict or legal right or right in the narrow sense, with practical scope that just renders transactions more efficient (Lestas 2014).

Here is a rough diagram depicting the development of human (and socioeconomic) rights, as we find them in the major human rights documents today (UDHR, ICCPR, ICESCR), from the Kantian duties of right:

Practical Reason

Autonomy (4:440) > ***Ethical Lawgiving (6:219)*** >

[Internally: Dignity (4:435)] ***[External Duties to Others (6:219)]***

(6:239)

- *Universal Perfect Duties of Right (6:240) → Universal Human Rights*
- *Specific Perfect Duties of Right (6:240) → Socioeconomic Rights*
- *Universal Imperfect Duties of Virtue (6:240) → No Rights*
- *Specific Imperfect Duties of Virtue (6:240) → No Rights*

Apparently, dignity is not the basis of rights. Rather, it is the value attributed to those who obey their duties – from which human rights are afterwards generated. Eventually,

14] While socioeconomic rights, namely the rights that need the establishment of certain institutions to determine the duty-bearers and their (negative and positive) 'distributed special obligations' are generated from the 'specific perfect duties of right' as it shown in the schema. See also (O'Neill 1996, 131-34).

Waldron's notion that dignity and rights might be *coordinate* ideas rather than ideas deriving one from the other seems to be true (2015, 118). Ultimately, UDHR's article 1 formulation is the correct one: "All human beings are born free and equal in dignity and rights…" (UN General Assembly, 1948).

msl22@kent.ac.uk

REFERENCES

Aquinas, Thomas. 1945. *Basic Writings of Saint Thomas Aquinas*. Edited by Anton C. Pegis. New York: Random House.

Baiasu, Sorin. 2016. Right's Complex Relation to Ethics in Kant: The Limits of Independentism. *Kant-Studien* 107 (1): 2-33.

Bennett, Christopher. 2015. *What is This Thing Called Ethics?* New York: Routledge.

Berman, Joshua. 2008. *Created Equal: How the Bible broke with Ancient Political Thought.* Oxford University Press.

Habermas, Jürgen. 2010. The concept of human dignity and the realistic utopia of human rights. *Metaphilosophy* 41(4): 464-80.

Hohfeld, Wesley N. 1917. Fundamental legal conceptions as applied in judicial reasoning. *The Yale Law Journal* 26 (8): 710-70.

James, Stephen. 2009. *Human rights*. John Wiley & Sons.

Kant, Immanuel, and Allen W. Wood. 1996. *Practical Philosophy*. Edited by Mary J. Gregor. The Cambridge Edition of the Works of Immanuel Kant. Cambridge: Cambridge University Press.

Lee, Patrick, and, Robert P. George. 2008. The nature and basis of human dignity. *Ratio Juris* 21 (2): 173-93.

Letsas, George. Review of *The Heart of Human Rights*, by Allen Buchanan, Oxford: Oxford University Press, 2013.

O'Neill, Onora. 1996. *Towards Justice and Virtue*. Cambridge: Cambridge University Press.

Rivers, Julian. 2013. Justifying Freedom of Religion: Does 'Dignity' Help? In *Understanding Human Dignity*, edited by Cristopher McCrudden. Oxford: Oxford University Press.

Ruston, Roger. 2004. *Human Rights and the Image of God*. SCM Press.

United Nations. 1948. Universal Declaration of Human Rights. http://www.un.org/en/universal-declaration-human-rights/ (accessed 23 October, 2017).

Vlastos, Gregory. 1984. Justice and Equality. In *Theories of Rights*, edited by Jeremy Waldron. Oxford: University Press.

Waldron, Jeremy. 2010a. The image of God: Rights, reason, and order. In *Christianity and Human Rights: An Introduction*, edited by John Witte, Jr and Frank S. Alexander. Cambridge: Cambridge University Press.

———. 2010b. Dignity, Rights, and Responsibilities. NYU School of Law, Public Law, Research Paper No. 10-83. https://ssrn.com/abstract=1710759 (accessed 27 October, 2017).

———. 2012. *Dignity, rank, and rights*. Oxford: Oxford University Press.

———. 2013. Citizenship and dignity. In *Understanding Human Dignity*, edited by Cristopher McCrudden. Oxford: Oxford University Press.

———. 2015. Is Dignity the Foundation of Human Rights?. In *Philosophical Foundations of Human Rights*, edited by Rowan Cruft, S. Matthew Liao, and Massimo Renzo, Oxford: Oxford University Press.

Public Reason 8 (1-2): 65-79

The Problem of Fake News

M R. X. Dentith
Institute for Research in the Humanities

Abstract: Looking at the recent spate of claims about "fake news" which appear to be a new feature of political discourse, I argue that fake news presents an interesting problem in epistemology. The phenomena of fake news trades upon tolerating a certain indifference towards truth, which is sometimes expressed insincerely by political actors. This indifference and insincerity, I argue, has been allowed to flourish due to the way in which we have set the terms of the "public" epistemology that maintains what is considered "rational" public discourse. I argue one potential salve to the problem of fake news is to challenge this public epistemology by injecting a certain ethical consideration back into the discourse.

Key words: fake news, social epistemology, polite society.

You would think – from the outcry and hullabaloo – that fake news, alternative "facts" and disinformation are becoming more and more prevalent in public discourse. Now, it may well be – in that we are certainly very alert to, or aware of its presence and use – especially now that certain populist, anti-intellectual politicians are on the rise. But what is troubling about fake news – if such a thing can be measured – is its centrality in certain political campaigns and regimes. Yet whether or not we happen to be concerned about populist politicians, SJWs (social justice warriors), or "virtue signaling", the phenomenon that is fake news is not new. After all, fake news, alternative facts and misinformation were part of the story certain Western powers tried to sell about weapons of mass destruction in Iraq, the basis and need for a Cold War, and the like. What might be new is the centrality of claims "That's just fake news!" in our political discourse.

So, what makes "fake news" *fake* news? What threat does fake news present to our polities? What are the epistemic and ethical concerns that occur when someone alleges some piece of news is fake? In section I a definition of fake news will be presented. I will then argue, in section II, that the power of claims of "Fake news!" come out of our understanding of what we typically think the "news" represents. Using *the Illustrated Police News* and the *Journal of American Physicians and Surgeons* as examples, I argue that there is a salient difference between sensationalised, sometimes fictitious "news" and "fake news", one which trades upon the difference between claims being insincere, and agents being indifferent about the truth of their utterances. I then, in section III, discuss the problem of balanced reporting, looking specifically at the way in which subjects which are controversial in one sense get reported as controversial in some other. Then, by looking at how, in section IV, the U.S. and the U.K. downplayed objections to the claim Iraq was developing weapons of mass destruction back in 2003, I posit that the threat that is accusations of 'That's just fake news!' comes out of worries that it is a merely a rhetorical device used by the powerful to crush dissent. I then posit, in section V, that a partial salve

to this worry is to challenge the standard of acceptable public discourse in most Western-style democracies by re-injecting a little ethics back into our *public* epistemology.

I. WHAT IS "FAKE NEWS"?

So, what is fake news? Well, fake news is an allegation that some story is misleading – it contains significant omissions – or even false – it is a lie – designed to deceive its intended audience.

That is, fake news is a *purported* fact. Such *purported* facts are either entirely false (and thus not facts at all) or they are only *partial* truths; the *purported* fact lacks some context or additional piece of information which, when revealed, undermines either its truth-value, or saliency to some broader claim.

This raises the question: is fake news simply disinformation relabeled? Disinformation is, after all, the introduction and use of *fabricated* information in order to make some otherwise suspicious explanation or theory look warranted.[1] The term was coined in the late 1930s by agents of the U.S.S.R., as a way of challenging the findings of the *Commission of Inquiry into the Charges Made against Leon Trotsky in the Moscow Trials* (AKA the Dewey Commission). A more modern example of the form is arguably the "Dodgy Dossier," which *allegedly* justified the invasion of Iraq in 2003, but turned out to have been doctored by political operators in the U.K.

Disinformation is *fabricated* or *purported* (because in some cases it is alluded to, but not presented to the public) information which is said to warrant some conclusion which also just happens to discredit some rival hypothesis. That is, it is often news which is fake.

However, the rhetorical frame 'That's just fake news!' can be ambiguous. That is, it goes beyond simply being disinformation. Sometimes the claim is, after all, sincere – what has been reported really is news which is fake – whilst other times the issue is that such a claim is being used as a baseless dismissal of stories, or of media outlets. For example, the allegation certain stories were "fake news" was co-opted by 45th President of the United States of America, Mr. Donald J. Trump during his run in the Republican Primaries in 2016, and has largely become a right-wing rhetorical device ever since.[2]

The claim "That's just fake news!" is also sometimes associated with a grander claim about the "fake media," such as Trump's tweet that:

> The Fake Media (not Real Media) has gotten even worse since the election. Every story is badly slanted. We have to hold them to the truth! (Trump 2017b)

1] For a fuller discussion of the definition of disinformation, see Matthew R. X. Dentith's "The Philosophy of Conspiracy Theories" (2014, 128-29).

2] Historically, the claim "That's just fake news!" was largely a left-wing broadside against certain knowingly partisan right wing news outlets, such as Fox News, and Infowars, which were peddling news stories which were, if not outright fabrications, reported in such a way to overstate their case.

Now, if we take the allegation of something being fake news, or fake media seriously, then – if we assume one side of the debate is correct – then someone is misleading the public. So, is the allegation of "That's just fake news!" just the exposure of a lie? Not quite; the allegation something is fake news is a *rhetorical device*, one designed to cast doubt on what would otherwise be some received story. That is, currently allegations of fake news, or fake media, focus on challenging mainstream media discourses.[3]

For example, at least according to the president and his aides, claims that the crowds at President Trump's inauguration were smaller than those of his predecessor Barack Obama was a clear-cut case of the media deliberately deceiving the public (David 2017). Similarly, claims that operatives within the Russian Federation hacked the Democratic National Committee, and leaked information which helped the Trump campaign, were labelled as "fake news" by the Trump administration (Trump 2017a).

II. "REAL" NEWS

What makes *fake* news so galling is our expectation that the news (what we might call the "*real* news" here) should be, if not factual, the most plausible and justified account of some event. While we typically accept *breaking* news might be subject to revision, and some events might even be misreported, we expect that those who investigate and report the news to *sincerely* be doing their best to get the story straight, and to report accurately their findings.

Media – the Fourth Estate – is often thought of as having an ethical duty to present to the public the facts. Allegations of fake news, then, cut into this account by suggesting either news reporters are being insincere – and reporting falsehoods as purported facts – or that they are indifferent to rival accounts of events, and simply reporting things according to some agenda.

Now, the ethical duty of the Fourth Estate is a curious one, because it might be a case of the public (the Third Estate) expecting of journalists a duty that members of the Fourth Estate themselves do not necessarily subscribe to.

For example, media coverage of supposedly contentious issues (particularly in the U.S., it seems) have turned out to be, for decades, highly partisan. Topics like anthropogenic climate change, the link between smoking and lung cancer, and the like, have all been examples of debates where one side (usually a particular media network) presents news which is contrary to some other side (typically another media network).[4] In this respect, the expectation that the news is an attempt at an *unbiased* reporting of the facts (or the most justified beliefs of the reporters) might seem naive to the sociologist, media

3] In this way, the history of claims of fake news are not very helpful to us, because the term was initially used to label alternative or non-mainstream news sites as fake. However, now the term is largely used to do the opposite; CNN is fake media. Breitbart and Infowars are "real" media.

4] We will leave to one side at this juncture the media vs. experts.

studies scholar, or social epistemologist. Given the disagreement amongst media outlets on certain issues, if journalists, and the like have a duty to the public, it is not clear that journalists and their contemporaries necessarily see that duty as being one of purveying facts. Rather, they might well see themselves as purveyors of opinion, with some news outlets claiming their opinions are more or less justified than those of some other.

As such, the idea that the news is unbiased, or apolitical, is a curious one; even the most casual or cursory examination of the media landscape shows that news outlets are not objective purveyors of truths, but, rather, subject to viewpoints and agendas. Now, like historians, journalists and their ilk, they may try to temper such inclinations or limitations in their reporting, but the idea of journalistic integrity meaning journalists are not subject to ideologies, politics, et cetera, seems odd at face value.

The idea the news is some attempt at the objective truth – that it is a factual endeavour – might also be a rather modern, and possibly fleeting, invention. History as a source-based, "as it happened," discipline was arguably the invention of Leopold von Ranke in the 19th Century. As such, the facts-based approach to history we saw in the last two hundred years may well be an anomaly. Ancient historians, for example, were well aware that their discipline was considered a particular branch of fiction, where it was permissible, even normal, to invent details in order to fill out, or improve the historical narrative. Newspapers, up until the early 20th Century, often reported fake news alongside real news, often with little to distinguish the two types of story (if, indeed, there was a real distinction to be made).

This is to say that perhaps the problem is not fake news *per se* but, rather, that we might have lived through a period of *real*, AKA objective, *factual* news, which was the real (historical) anomaly.

1. Old News

Take, for example, *The Illustrated Police News*, published between 1864 and 1938 in the U.K. A tabloid publication, *The Illustrated Police News* contained sensational and sensationalised news stories about crimes both historic and contemporary. Not everything *The Illustrated Police News* published was true; some of the stories within its pages were edited for dramatic effect, whilst some others were (probably) entirely fictional. *The Illustrated Police News* contained and peddled some stories that we would consider now to be fake, in order to please an eager public who wanted "penny dreadful" tales which, plausibly, took place in the world in which they lived.

Fast forward to the present day, and you have the *Journal of American Physicians and Surgeons*, a publication of the Association of American Physicians and Surgeons. The Association of American Physicians and Surgeons was founded in 1943 as a non-profit association. On the face of it both the Association of American Physicians and Surgeons, and the *Journal of American Physicians and Surgeons* sound perfectly normal. However, the Association of American Physicians and Surgeons was founded to specifically fight back

against "socialised medicine" in the U.S., and combat a purported takeover of medical practice by the government. The *Journal of American Physicians and Surgeons* has – over its long tenure – published a raft of articles on what are considered *discredited* medical hypotheses. These include papers arguing that the HIV virus does not cause AIDS, the link between vaccines and autism link, and abortions being a significant causal factor in the development of breast cancer. However, these papers are neither well-researched, nor well-evidenced, and thus do not survive peer review outside of the gated community which subscribes to the *Journal of American Physicians and Surgeons*.

Examples of publications like *The Illustrated Police News*, and the *Journal of American Physicians and Surgeons* are not hard to find. Indeed, as *The Illustrated Police News* shows, publishers of *purported* facts have been part of our media landscape for a long time. Yet *the Illustrated Police News* and its sensationalised stories was a kind of open secret; the writers of that periodical were not covertly making up stories. Rather, they expected that their readers would be aware of the sensationalised nature of the stories within its pages. However, the pieces published in the *Journal of American Physicians and Surgeons* are meant to be taken seriously.

2. A Salient Difference

The salient difference between the *Illustrated Police News* and the *Journal of American Physicians and Surgeons* is surely one of intent. The editors *and readers* of the *Illustrated Police News* knew that the content of the gazette was sensationalised; the entire point of the publication was to titillate and entertain. However, the *Journal of American Physicians and Surgeons* does not present its "findings" (such as they are) as entertainment, but, rather, as *purported* facts, ones which challenge findings elsewhere in the medical sciences.

What should we make of this? It is tempting to think that perhaps this is an example of what Richard Feldman has described as "reasonable disagreement." Feldman considers whether it is possible for "epistemic peers" – epistemic agents roughly equal with respect to intelligence, reasoning ability, and privy to the same background information – to *reasonably* disagree with one another (2006; 2011).

Reasonable disagreement differs from standard disagreement with respect to the way in which the disagreement is managed. When two or more people *reasonably* disagree with one another, they effectively *agree to disagree*, rather than continue to try and persuade the other party to switch positions. Using examples from the Philosophy of Religion, Feldman argues that atheist and theist philosophers seem able to agree to disagree as to whether the gods exist. This, he notes, should strike us as odd. In part this is because it is just a fact of the world as to whether the gods exist, and also because philosophers of religion, by-and-large, share the same kind of training and background knowledge – they are epistemic peers, after all – and so you should expect there to be reasonable agreement on the matter, not reasonable *disagreement*.[5]

5] There could, of course, be unreasonable disagreement on the issue, if one party or the other de-

One explanation for this seemingly reasonable disagreement is to claim that some epistemic peers have access to a special kind of evidence – private evidence – which cannot be shared with some other epistemic peer (Feldman 2006, 222-3). When epistemic peers cannot share, disclose, or make public all the evidence for their views, they cannot enter into a state of full disclosure (where peers will express all *salient* reasons for their beliefs on a certain topic) and thus can then argue over the *relevance* of their evidence (2006, 220). If the peers still disagree with one another when under full disclosure, then either the disagreement must be unreasonable – because one peer is sticking to their point of view *in spite of the evidence* – or they are not epistemic peers at all.

Now, part of the problem here is that working out who epistemic peers are – in many cases – is tricky, and the problem comes in two forms. The first is that it can be difficult for highly educated people to work out whether they are epistemic peers with other similarly educated folk. The second issue is that if it is hard for epistemic peers to recognise one another, this difficulty is compounded when we ask lay members of the public to both recognise and trust in the work of such peers. Indeed, we see this in public debates about the existence of the gods, which are often *unreasonable*, in that we see supposed epistemic peers square off against each other, and *disagree to disagree*. In at least some of these such cases, the peer relationship is not epistemic peer versus epistemic peer. Rather, two or more prominent individuals from each camp, who are not roughly equal with respect to intelligence, reasoning ability, and privy to the same background information, are mistakenly or falsely portrayed as being peers in an epistemic sense.

With respect to the *Journal of American Physicians and Surgeons*, is this a case of reasonable disagreement when it comes to the medical sciences? Do the authors and editors reasonably disagree with their peers? Or, are they are acting insincerely? That is, are they trading upon the notion of being epistemic peers in order to muddy the waters of medical research? Are the editors failing to do their due diligence because they have political views which trump their academic interests? Or do they simply not have the requisite background to do their job properly?

For the editors of the *Journal of American Physicians and Surgeons* to claim papers in their journal are the products of reasonable disagreement in the field of medical research would require that researchers publishing work elsewhere (their academic peers) agree with that assessment. Yet the work in the *Journal of American Physicians and Surgeons* is largely only cited within a subset of journals, journals which are similarly (and often overtly) slanted against the vast majority of work published elsewhere in the medical sciences. This calls into question the findings of the journal because it suggests that the journal exists to further a political, rather than scientific agenda. That is, the worry about the *Journal of American Physicians and Surgeons* is not just that it probably is just a platform for the dissemination of fake research/news, but it also exists to portray research to the contrary as examples of "That's just fake news!"

---

cides to stick with their position despite the wealth of arguments and evidence against it.

III. BALANCED REPORTING

Let us look at another example, a journal of record which *purports* to combat fake news, but seemingly engages in promulgating it as well. In April of 2017 the *New York Times* hired Bret Stephens as a columnist. This was a controversial choice; Stephens is a noted anthropogenic climate change denier and earlier in the year the *New York Times* campaigned for new subscribers on the basis that they sought to combat fake news generally, and *specifically* accused those of pushing the denial of anthropogenic climate change as purveying *fake news*. Yet, when asked why they had hired Bret Stephens, the editors of the *New York Times* claimed, in their defence, that there are "millions of people who agree with him." (Calderone and Baumann 2017)

The cynical will claim that this is an entirely economic decision: in order to increase the number of subscribers to the *New York Times* the editors and the managing board need writers who reflect the multiplicity of views found amongst its readers. In this respect the idea that the *New York Times* was going to fight the spread of fake news was itself most likely an economic decision, one designed to increase subscriber numbers. That is, the editors and the managing board of the *New York Times* were, if not insincere, indifferent about the threat of fake news, at least if we treat seriously their claim Bret Stephens was a worthy hire just because there are "millions of people who agree with him."

This worry about the indifference, or insincerity of agents when it comes to activities the public take it have associated epistemic or ethical duties presents a challenge to how we conceive the public sphere. It suggests, as we will see, that what we might term the "public epistemology" which underlines much contemporary political debate is oddly formed. However, at least in defence of the *New York Times,* we can claim that as the *New York Times* is a journal of record, their employment of Bret Stephens – who defies their stated mission – might be the result of needing to engage in balanced reporting of controversial or contentious issues.

1. False balance

One duty we typically place upon the media is that of presenting controversial issues in a balanced way. We expect – if a contentious issue is being discussed – that all sides will get to present their story. What is the motivation or source for this balance? It might be motivated purely by an appeal to freedom of expression, or the principle that all ideas should be appraised on their merits, rather than shutdown or kept out of public debate due to social niceties, or political agendas. That is, we expect reporting on controversial or contentious topics to be balanced, because – in the end – it serves the public good for all sides to be well-informed. In this respect, the purpose of balanced reporting aids in the task of making us epistemic peers with respect to the issue at hand.

Yet balanced reporting has become (if, indeed, it was not always) something of a problem. This is because the *appearance* of a controversy does not tell us something is controversial in an *interesting* or *salient* sense. Take, for example, Bret Stephens and his

denial of anthropogenic climate change. It is true that anthropogenic climate change is a controversial topic, but it is not controversial because the science behind such a claim is unsettled (and thus up to debate). Rather, it is controversial in some other sense: it either commits us to actions some of us would rather not engage in, or the science is in conflict with some ideology.

Claiming something is controversial does not tell us *in what way it is controversial*. A topic which is controversial politically might not be controversial scientifically. A topic which is controversial religiously might be considered uncontroversial in a sectarian setting. Yet, the kind of reporting we routinely see when it comes to anthropogenic climate change creates the impression that it is controversial in a scientific sense. The "balanced" reporting we see associated with it is not sociologists and psychologists debating what it is that causes a scientific thesis to be dismissed by some on political grounds. Rather, the balanced reporting often pits climatologists – who are arguing as to why the science says the climate is changing – against talking heads, *as if they are epistemic peers*.

This is "false balance," where an issue is presented as controversial in a sense that it ultimately is not. Whilst scientists might disagree with some of the minutiae of just what anthropogenic climate change will entail, the science is settled as to the fact it is occurring. Yet by presenting the *political* controversy as a *scientific* one, the reporting on anthropogenic climate change creates a sense of *false* balance.

Balanced reporting can, of course, be difficult. After all, there are different kinds of reporting, which come with different epistemic burdens. There is – at least it seems – a salient difference between science reporting, and the "simple" reporting of, say, what happened at a town hall meeting. Determining who the experts or appropriate witnesses is trivial in the latter case, at least with respect to the former. Then there is the worry that concerns about balanced reporting rest on strong personal preferences: **I** am worried about anthropogenic climate change, so **I** am perturbed by how it gets reported, and thus see problems when it comes to balance. **You** might be similarly disturbed by reports of low level corruption that do not dismay me. None of this gets us away from the problem of appraising and judging expertise, but it does suggest the problem of balance has both a (purely) epistemic *and* a social aspect to it.

Whatever the case, *false* balance arguably allows fake news to flourish. By presenting issues which are controversial in one sense as being controversial in some other, we either let insincere agents present disinformation as news, or people who are indifferent to certain research methodologies or paradigms to present their views as being *en par* with those of others when it turns out they are not.

2. News vs. Opinions

Sometimes balanced reporting is defended as *merely* being a competition of opinions, in that the opinion of one expert is taken to be equivalent to that of another, whether or not it has been established whether said experts are epistemic peers.

Now, opinions, unlike facts, are taken to be personal and subjective. Indeed, a feature of contemporary public discourse (although surely a feature throughout human history) is the defence "That is just your opinion!" When someone says "That's just your opinion!" they are typically claiming that no matter how factual your utterance is, they are entitled to some other belief because whatever you say, it is just an opinion. The idea behind "That's just your opinion!" is that if everything is an opinion, then all opinions are equally warranted, or everyone is entitled to their opinions over those of others.[6]

The allegation that some piece of news is "fake news" could be seen as just another way of saying "That's just your opinion!" However, in the case of alleging something is "fake news," the implication is that the only correct opinion is that of the president, the minister, or the pontiff. That is, alleging something is fake news *and expecting it to have some weight* requires implying that whatever your opinion is, the opinion of the person in power is not just equal. No, rather it is the stronger claim that what you have been told is a lie. Not just that, but what you have been told has likely been fabricated for political, or commercial purposes. As such, you should give it no epistemic weight whatsoever.

IV. APPEALS BY AUTHORITY

Let us return to the Dodgy Dossier, and the *purported* weapons of mass destruction Iraq was alleged to be developing. In 2003 the governments of the U.S. and the U.K. claimed they had credible evidence that the Saddam Hussein regime in Iraq was developing weapons of mass destruction, despite U.N. weapons inspectors claiming otherwise. The media entered the debate, largely questioning the official narrative coming out of the White House and Downing Street, a view which subsequently became popular with the publics in the U.S., U.K., and most of the West in general.

The W. Bush and Blair governments continued to claim they had evidence that Iraq was developing weapons of mass destruction (WMDs), and that their detractors were wrong. The existence of this evidence – the public was told – justified the actions of the W. Bush and Blair governments, even in the absence of support by the international community. What was interesting about this evidence was that it was *secret*.

Now, it might be appropriate, in *some* cases, to keep certain evidence secret. We might want to protect sources, for example, or some of the evidence is simply not salient – perhaps even distracting – to the issue at hand. However, whenever some hypothesis or view is only said to be justified with respect to secret evidence, the worry is that the person or people keeping the secret are manipulating the evidence, or being insincere.

6] As Patrick Stokes has argued, talk of opinions and justifying said opinions can be a fruitful start to a discussion, but we are not automatically entitled to our opinions. Opinions need to be backed up with arguments, and reasons if we are to expect others to take them seriously (2012). As such, the antidote to "That's just your opinion!" is to either say "It's not my opinion," or "Here's why this opinion matters, however…"

For example, it has been argued that Rumsfeld and associates knew the confidential intelligence they received about the Iraqi WMD programme only weakly suggested that Iraq might still be producing weapons of mass destruction, but they treated these claims are having much more evidential weight than they deserved because it was politically convenient. When confronted about this, they called their opponents "conspiracy theorists" who were peddling "conspiracy theories," characterising the rival views as belonging to some extreme, non-centrist position. The effect was the same as labelling these contrary views "fake news." Admittedly, this rhetorical tactic by the governments of the U.S. and the U.K. largely failed, in part because the many of the very people they were accusing of now being conspiracy theorists had been portrayed as trusted political operators only weeks before. The official theory about the WMDs was revealed to be subservient to some political agenda and not upon an appeal to the evidence.

1. Influence

Worries about the Dodgy Dossier point to the motivating concern about claims of fake news here-and-now, which is how allegations of "That's just fake news!" are abused by those in power. It is not unusual for politicians to accuse people, and to be accused themselves, of getting their "facts" wrong. "Fact," as a term in political debate, often has a different meaning to that in epistemology; "facts" can be wrong, "facts" can be contested, and "facts" can change from situation to situation. Politically, "facts" are something between opinions and justified beliefs. That is, politically "facts" are *purported* facts. But what is most striking about much *fake news* is the sheer indifference many purveyors of such "news" have when it comes to presenting and defending such purported facts.

Take, for example, senior Trump aide Kellyanne Conway. When she was confronted about her defence of White House Press Secretary Sean Spicer's obviously false claims about the crowds at the inauguration of President Trump,[7] she claimed Spicer was simply presenting "alternative facts" (Gajanan 2017). Later, she defended the first travel restriction executive order of the Trump presidency (AKA the "Muslim Ban") with reference to a wholly fictitious Muslim terror attack, the "Bowling Green Massacre" (Blake 2017). In both cases Conway showed little concern that the information she was presenting or relying upon was false.

Were Conway and Spicer being stupid, or were they being insincere in their proclamations, in the hope someone would believe them?[8] If it was the latter, then therein lies the seductive nature of claiming some opposing report is "fake news;" if you can get away with using allegations of fake news in your discourse, then you can change the nature and direction of public debate. In both cases – the inauguration crowd size, and the

7] In that obvious photographic and eye-witness evidence was readily available to show Spicer's claims to be "trumped" up.

8] With respect to the Bowling Green case Conway's numerous references to the entirely fictitious event indicates insincerity on her part.

"incident" at Bowling Green – Conway was alleging that the media was not presenting the *real* news: Trump *really* had record crowds at his inauguration despite what the media was telling the public, she was claiming, and the Muslim Ban was justified because of incidents like the Bowling Green Massacre that the media had refused to report on.

Was Kellyanne Conway a lone actor, presenting "alternative facts" in isolation? No. The newly-minted Trump presidency was simply continuing a policy of relying upon claims of "That's just fake news!" and presenting "alternative facts" which had served them well during the presidential primaries, and then the presidential election campaign. That is to say, it was a long-standing policy within the Trump campaign team to present mainstream news coverage about the Trump campaign, and then presidency, as fake news.

Now, fake news and allegations of "That's just fake news!" are not solely a problem of governance; fake news was part of the Brexit campaign, which was – at least ostensibly – non-partisan. Putative politicians outside the mainstream political parties made allegations of fake news part of their campaigns in the election in the Netherlands and France in 2017. Indeed, Trump used allegations of fake news well in advance of being president. So, even though the threat of fake news seems worst when it comes from people in power, the allegation some stories are fake news can emanate from non-governmental sources, and be potentially as dangerous. However, a spokesperson for government labelling some news as "fake," in order to cast doubt on its veracity *sans* any good reason to do so, is an abrogation of the duty of care we expect of our elected representatives.

V. THE PROBLEM OF FAKE NEWS

The allegation "That's just fake news!" – as we have seen – is a problem because claims some piece of information in public discourse are fake often are made insincerely (the agent is lying), or are the result of indifference towards truth (the agent does not care whether the allegation is true or false). Now, labelling something as "fake news" does not necessarily mean people will regard it as false. However, alleging something is "fake news" is still a problem because it has become a rhetorical ploy, used mostly by those in power. That is, the claim "That's just fake news!" exists to create the illusion of a controversy.

Some have blamed a culture of post-modernity for the phenomena of fake news (take, for example, historian Richard Evans (Evans 2017; Sandham 2017)). Whilst it is true that many of the proponents of allegations of fake news – at least in the White House – were schooled at a time where post-modernist approaches in education were popular, the idea a disdain for the facts is a relatively new phenomenon seems quaintly ahistorical; how would we explain earlier examples of the same phenomena, such as *The Illustrated Police News*? It does not help that many of the proponents of "fake news" seem utterly opposed to what they consider to be a "leftish" notion, that of post-modernist approaches.

A more promising angle on this phenomenon can be found in critiques of contemporary political liberalism. As Jack Z. Bratich has argued, a liberal philosophical

outlook can become a remarkably intolerant political position. Because liberalism focuses on normalising political discourse, and moving away from extremism, the sensible middle is where political agents should desire to be. As such, marginal voices (those on the extremes) are often chastised for not adhering to the politics of the centre, *even if the centre is failing to address the issue at stake* (Bratich 2008). That is, maybe the *contemporary* problem of allegations of fake news is its apparently sudden centrality to political debate?

Indeed, one of many disturbing consequences – at least to the liberal – to rhetoric such as "That's just fake news!" is the way in which it shifts the centre of political debate. If positions which should be considered uncontroversial are presented as *merely* differences of opinion, or as insincere, or even falsehoods, then this licenses – at least to some – grounds for debate on topics which do not need it.

Another worrying consequence is that by alleging some media outlet purveys fake news this – either explicitly or inadvertently – lends itself to a narrative wherein outlets known for sensationalising, or even fabricating "news" stories (such as the Alex Jones' sites *Infowars* and *Prison Planet*) are normalised, if not exalted as being *real* or *proper* news.

1. What to do about Fake News?

So, what can we do? The following suggestion as to how to curtail both the power of fake news and the allegation of "That's just fake news!" is – at best – only a partial salve to an age-old problem.

Part of the problem of fake news might well be a product of what we might call the "Polite society." In a polite society there are certain things which are not talked about. Some truths which might be considered toxic[9] should they be discussed openly are *politely* ignored or glossed over. For example, we might have all been aware that, in the 1970s, the police routinely planted evidence in order to secure convictions, but as those criminals were thought to be obviously guilty of *something*, we *politely* ignored the specific cases of evidence tampering. It would be impolite to talk about the matter, or think of raising it because the intentions of the police – keeping the streets safe – was a public good.[10]

Politeness is one reason why the claim "That's just your opinion!" ends up having what appears to be *apparent* epistemic weight; we are often polite in the face of dissent, in order to not cause further dissent, or embarrassment.

The obvious objection to the Polite society hypothesis is to claim that this hypothesis is only true of some societies, and only true at certain times in those societies. I do not disagree; the existence of a polite society is likely the result of a variety of different cultural and political factors. Whilst I think much of the West, for example, currently exists in a state of politeness, how that politeness will be expressed will differ from place to place, have come into existence at different times, and might be – in some polities – less and less tenable.

9] The idea of truths which are toxic I borrow liberally from the work of Lee Basham (2017).

10] For further detail about the "Polite Society" hypothesis, see "Conspiracy theories on the basis of the evidence" (Dentith 2017).

Indeed, we are often told that political debate has no place for impoliteness. That is to say that emotions, or feelings are said to have no epistemic weight, or worth, and are just likely to cloud our judgement; they are not part of our *public* epistemology. Now, some might be tempted to argue either there are norms which license this. However, civic discourse does not exist in a vacuum; it echoes and often reinforces pre-existing structures and hierarchies. An awful lot of challenging political speech – talk of institutional sexism, racism, misogyny – is confronting, and there is a tendency to downplay such talk, or excuse oneself from the analysis. A consequence of this is for people adversely affected by these structural pieces of discrimination to get angry, or upset.[11] The standard of debate we find which has allowed allegations of fake news to flourish, I would argue, has come out of requirement that we be dispassionate and polite in our discussions. This allows for accusations of "That's just fake news!" – to proliferate. This situation should give us reason to pause for thought. Depending on what you think the relationship between epistemology and ethics is, the idea that rational debate should be conducted, or based outside of ethical norms (whatever they might be), is startling. After all, if you think epistemology is the study of what we *ought* to believe, then it is curious that it should be considered divorced from ethics, which considers how we *ought* to behave.[12]

After all, it is easy to be dispassionate about events which do not directly affect you, or you are largely indifferent to. That is to say, much political debate seems to assume that we are discussing opinions. As such, when someone claims "That's just fake news!" our initial response is often to laugh about it, because that is the polite thing to do. Would the problem of fake news dissipate if we were more impolite when it came to public discourse? Probably not. However, if we were more impolite – which is to say we challenged the politeness which infects much public debate – it would make it harder for certain sorts of allegations of fake news to hold sway.

VI. CONCLUSIONS

Allegations that some view or position is an example of fake news need to be taken seriously. This is because such allegations typically exist to marginalise or delegitimise views, in a way in which we are meant to think that the expressions of the powerful are at least equal to, or better than, the news (at least as we typically understand it).

It is tempting to appropriate blame for the existence, and power of claims of "That's just fake news!" on a variety of sources, whether that be post-modernists, false balance in the media, and the like. However, the spectre of fake news – no matter its history – is worrying simply because it speaks to a certain indifference or insincerity amongst agents

11] This is to say nothing about explicit cases of sexism, racism, or misogyny.

12] Indeed, for some of us, epistemology might be seen as either a branch of ethics, or complementary to it.

in contemporary political debate. How might we challenge such allegations? Well, one potential solution – although it is likely only a partial salve – is to change the terms of the debate by challenging the polite way in which we discuss issues pertinent to our polities.

After all, alleging something is fake news is an abrogation of the duty of care we expect of people in power; by being either indifferent or insincere in their utterances, such influential people or institutions fail to uphold an *expected* duty. Rather than present such discussions as a war of opinions, and claim that – as they are opinions – you have no right to be upset about them, let alone dispute them, we need to include more room for impoliteness when confronting issues in our contemporary public discourse.

m.dentith@episto.org

Acknowledgements:

Thanks to Jeremy Seligman, Martin Orr, and Mihail-Valentin Cernea for helpful feedback on an earlier version of this paper.

REFERENCES

Basham, Lee. 2017. Joining the Conspiracy. *Argumenta*. DOI: 10.23811/55.arg2017.bas

Blake, Aaron. 2017. Kellyanne Conway's 'Bowling Green Massacre' Wasn't a Slip of the Tongue. She Has Said It Before. *Washington Post*. https://www.washingtonpost.com/news/thefix/wp/2017/02/06/kellyanne-conways-bowling-green-massacre-wasnt-a-slip-of-the-tongue-shes-said-it-before/ (accessed 28 July, 2017).

Bratich, Jack Z. 2008. *Conspiracy Panics: Political Rationality and Popular Culture*. New York: State University of New York Press.

Calderone, Michael, and Nick Baumann. 2017. Hiring Anti-Trump Conservative Is Part of New York Times' Effort to Expand Opinion. *Huffington Post*. http://www.huffingtonpost.com/entry/bret-stephens-new-yorktimes_us_58f12c80e4b0b9e9848bed3e. (accessed 25 July, 2017).

David, Javier E. 2017. White House Press Secretary Sean Spicer Rips Media, but Facts Contradict His Arguments. *CNBC*. http://www.cnbc.com/2017/01/21/white-house-press-secretary-sean-spicer-rips-media-for-false-reporting-of-inauguration-day-crowd.html (accessed 23 July, 2017).

Dentith, Matthew R. X. 2014. *The Philosophy of Conspiracy Theories*. Palgrave Macmillan.

———. 2017. Conspiracy theories on the basis of the evidence. *Synthese*. DOI: 10.1007/s11229-017-1532-7

Evans, Richard. 2017. If I Am Wrong, and Postmodernist Disbelief in Truth Didn't Lead to Our Post-Truth Age, Then How Do We Explain the Current Disdain for Facts?. *Twitter*. https://twitter.com/RichardEvans36/status/823828364280664065 (accessed 29 July, 2017).

Feldman, Richard. 2006. Epistemological Puzzles About Disagreement. In *Epistemology Futures*, edited by Stephen Cade Hetherington. Oxford: Oxford University Press.

———. 2011. Reasonable Religious Disagreements. In *Social Epistemology: Essential Readings*, edited by Alvin Goldman and Dennis Whitcomb. Oxford University Press.

Gajanan, Mahita. 2017. Kellyanne Conway Defends White House's Falsehoods as 'Alternative Facts'. *Time*. http://time.com/4642689/kellyanne-conway-sean-spicer-donald- trump-alternative-facts/ (accessed 27 August, 2017).

Sandham, Alice. 2017. Denial: An Interview with Sir Richard Evans. *Woolf Institute.* https://woolfinstitute.blog/2017/02/07/denial/ (accessed 15 August, 2017).

Stokes, Patrick. 2012. No, You're Not Entitled to Your Opinion. *The Conversation.* https://theconversation.com/no-youre-not-entitled-to-your-opinion-9978 (accessed 13 August, 2017).

Trump, Donald J. 2017a. The Fake Media (Not Real Media) Has Gotten Even Worse Since the Election. Every Story Is Badly Slanted. We Have to Hold Them to the Truth! *Twitter.* https://twitter.com/realdonaldtrump/status/853945633903923200 (accessed 15 August, 2017).

———. 2017b. James Clapper and Others Stated That There Is No Evidence Potus Colluded with Russia. This Story Is FAKE NEWS and Everyone Knows It! *Twitter.* https://twitter.com/realDonaldTrump/status/843772976151642112 (accessed 23 August, 2017).

Public Reason 8 (1-2): 81-89

Political Legitimacy and the Unreliability of Language

Eva Erman & Niklas Möller
Stockholm University, Royal Institute of Technology

Abstract: Many political theorists in current debates have argued that pragmatist theories of mind and language place certain constraints on our normative political theories. In a couple of papers, we have accused these pragmatically influenced political theorists of misapplication of otherwise perfectly valid ideas. In a recent paper, one of the targets of our critique, Thomas Fossen, has retorted that we have misrepresented the role that a pragmatist theory of language plays in these accounts. In this paper, we claim that Fossen's attempt to chisel out a role for his account in normative political theory rehearses the same problematic view of the utility of theories of language as his previous iterations. We argue that Fossen's account is still guilty of the fallacious claim that a pragmatist theory of language (in his case Robert Brandom's account) has implications for the form and justification of theories of political legitimacy. We specifically focus on three flaws with his current reply: the idea that criteria and conditions are problematic on a pragmatist outlook, the idea that a pragmatist linguistic account applied to a particular political context will have a distinct political-theoretical payoff, and the idea that a fundamental linguistic level of analysis supplies normative guidance for theorizing political legitimacy.

Key words: political legitimacy, pragmatism, Robert Brandom, theories of language.

Many political theorists in current debates (Mouffe 1999, 2000; Tully 1989, 2002; Norval 2006; Fossen 2013) have argued that the pragmatist theories of mind and language put forward by Ludwig Wittgenstein (or ideas, if you prefer not to use the word 'theory' for his thoughts) and Robert Brandom place certain constraints on our normative political theories. In a couple of papers, we have accused these pragmatically influenced political theorists of misapplication of otherwise perfectly valid ideas (Erman and Möller 2014, 2015). In a recent paper, however, one of the targets of our critique, Thomas Fossen (2017), has retorted that we have misrepresented the role that a pragmatist theory of language plays in these accounts:

> Pragmatism's significance for thinking about political legitimacy does not lie in the normative conclusions it justifies but in the way it re-orients our thinking toward political practice. This shift in orientation does not refute standard theories of political legitimacy, as such, but it renders problematic their narrow focus on criteria of legitimacy, in abstraction from the forms of political practice in which such criteria are at stake (Fossen 2017, 2).

Fossen further argues that we presuppose, in our critique, an "overly narrow view of what 'normative political theory' consists in ... that pragmatism calls into question" (2017, 2). The task, he claims, "is not just to articulate criteria of legitimacy, but more fundamentally to explicate the ways in which the question of legitimacy manifests itself in practice, and the forms of activity through which we might engage it" (2017, 11).

In what follows, we will argue that Fossen's attempt to chisel out a role for his account in normative political theory rehearses the same problematic view of the utility of theories

of language as his previous iterations. If anything, he now makes even more of a strawman out of the accounts from which he attempts to distinguish himself. We will argue that Fossen's account is still guilty of the fallacious claim that a pragmatist theory of language has implications for the *form and justification* of theories of political legitimacy, focusing specifically on *three* flaws with his current reply: the idea that criteria and conditions are problematic on a pragmatist outlook (first section), the idea that a pragmatist linguistic account applied to a particular political context will have a distinct political-theoretical payoff (second section), and the idea that a fundamental linguistic level of analysis is of normative guidance for theorizing political legitimacy (third section).

Before we address these flaws, let us say a few words about Fossen's project. In Fossen's view, his account – as well as the accounts of Mouffe, Norval, and Tully making use of Wittgenstein – aims "to open up conceptual room for a different way of looking at a problem" (2017, 4). Utilizing Brandom's seminal theory of language, Fossen proposes an account of the specific role of the concept of legitimacy within a certain form of practice: the encounter between political subject and authority. A key claim is that the role of the concept of legitimacy is to express one's political stance toward the authorities:

> In calling an authority legitimate, one attributes an entitlement to rule to that authority, while also undertaking a commitment to treat it in ways appropriate to its status (say, as a source of reasons), and attributing such commitments to other subjects. In this way … one can explain what it is to *be* legitimate in terms of taking something *as* legitimate (Fossen 2017, 6).

Hence, while Fossen acknowledges that Brandom's theory is fully general, he still stresses "its significance for thinking about political legitimacy": its specific application to the political context has a theoretical payoff as it "draws our theoretical attention toward political practice" (2017, 7; 2013, 426-50). Fossen concludes that while Brandom's theory alone does not entail any constraints on political theories, "[i]f the content and justification of criteria of legitimacy is bound up with ongoing practice ... then *practice* may place such constraints" (2017, 7). According to Fossen,

> pragmatism about language *does* suggest ... that the purview of a theory of political legitimacy is often construed too narrowly. Theories of legitimacy are usually taken to consist in the articulation and justification of criteria of legitimacy (what Erman and Möller refer to as a 'substantial' normative theory). Typically, theorists proceed as if one can settle the content and justification of such criteria in abstraction from the forms of practice through which legitimacy is politically contested – just as certain theorists of language consider meaning (semantics) in abstraction from use (pragmatics). From a pragmatist perspective, that is a problematic form of abstraction because it fails to do justice to the ways in which concepts and criteria are bound up with practice. So, the difference a pragmatist approach makes here is that it problematizes the failure to attend to politics – a lack of realism, if you will (2017, 7).

With these basic ideas on the table, largely in Fossen's own words, let us turn to the first flaw of Fossen's attempt to save his Brandomian approach to normative political theory.

I. FIRST FLAW: THE 'CRITERIA' STRAWMAN

Fossen structures his reply around the question of whether his account is guilty of the 'pragmatist fallacy,' i.e., the fallacious idea that one can infer normative conclusions from pragmatist linguistic theories. He insists that it is not, and as first order normative conclusions are concerned, we agree. *Pace* Fossen's explicit denial and somewhat shifting terminology, however, his account is still guilty of an analogous *meta-normative* pragmatist fallacy, namely, to put constraints on the *form and justification* of normative political theories of legitimacy through a theory of language.

In his original account, he claimed that on a pragmatist analysis, political judgment is not a philosophical problem "calling for a general solution" (Fossen 2013, 442). We demonstrated in our previous article that Brandom's theory of language entails no such anti-generalist conclusions. In his current reply, Fossen downplays the anti-generalist constraints of his previous iteration, now putting a stronger emphasis on *criteria of justification* as the feature that traditional political theorists allegedly are obsessed with but from which his account frees us. His account, he now insists, "re-orients our thinking toward political practice" (Fossen 2017, 2) and thus "free[s] us from captivity by a picture" (Fossen 2017, 4). This picture is political philosophy as a theoretical endeavour with a "narrow focus on criteria of legitimacy, in abstraction from the forms of political practice in which such criteria are at stake" (2017, 2).

But Fossen's new emphasis on criteria of justification is just a sheep in wolf's clothing: it faces the exact same objection as we demonstrated in our previous critique. There we showed that Brandom's account, in which conceptual content is grounded in our actual practices – the commitments and entitlements we assign to ourselves and our fellow linguistic participants – is in no way 'anti-theoretic:' it entails no bars on general principles or norms, holding for all eternity and for all and everyone. It is in fact a feature of Brandom's account that it accommodates, in principle, true claims on any level of generality and within any domain. Mathematical theory, the laws of physics, and even Brandom's own systematic account of language, which aims to be perfectly general – what he claims about language does not hold for his fellow Americans only, but for all linguistic practices and practitioners – are all safe, as far as his theory of language is concerned (Erman and Möller 2014, 490-494).

This ecumenical feature of Brandom's theory includes the concepts used in forming claims. We may frame our claims in very particularist language, free from the use of any conditions or criteria and without referring to any particular norms ("Their [*pointing at an entity*] power is legitimate in the here-and-now"), or we may frame them in terms of necessary and sufficient conditions ("The power of an entity X is legitimate if and only if condition Y holds"). Brandom provides a pragmatist account of how a practice counts as a linguistic practice and when a performance counts as belonging to it, together with an inferentialist semantic account of the conceptual content of such performances. That analysis is valid for all sorts of claimings; it does not promote one sort and reject another.

Whether any specific claim holds has to do, as Brandom likes to put it, with the 'giving and asking for reasons': as long as we have better reasons to believe a claim, at whatever level of generality and whatever form it happens to have, we should endorse it; otherwise not (Brandom 1994).

Apart from being unsupported from a Brandomian viewpoint, the mentioning of 'giving reasons' brings us to the strawman aspect of Fossen's painting of mainstream political philosophers in general (and us in particular) as narrowly focusing on 'criteria,' 'standards,' or 'necessary and sufficient conditions' for what is legitimate (Fossen 2017, 2-3, 5, 8; 2013, 430). By focusing on *one* way of making a normative argument – that of putting forward criteria and conditions for what is legitimate – and claiming that it is a narrow-minded view of normative political theory, Fossen merely conceals the problem at hand: the lack of *any form* of normative reasoning in his account. While a normative argument does not need to make reference to criteria, standards or conditions, it needs to give *reasons* for (in this case) why an entity is legitimate. These can be contextual or general, but there need to be reasons. Consequently, it is a fundamental problem that Fossen's account "bracket[s] the question of what makes a political authority legitimate" (Fossen 2013, 432), arguing instead that an account which tells us to focus on how persons or groups *take* it to be so is enough to constitute a "genuine alternative" (2013, 428).

Judging from his insistence on a *normative* role for a 'Brandomian' theory of legitimacy, we suspect that Fossen mistakes *inspiration* for *argument*. Here, the case of the utilization of Heisenberg's uncertainty principle springs to mind. Heisenberg's uncertainty principle is a famous principle in quantum mechanics, which states that we cannot at the same time measure the exact position of a particle and its exact speed (or rather momentum, i.e. the product of its mass and speed); the more precisely we determine the position, the more uncertain we are about the speed with which it travels. This principle has been used time and again by human and social scientists – through what Douglas Hofstadter calls "careless paraphrases" – in attempts to substantiate how for example the observer always interferes with, and thus changes, the phenomenon under scrutiny (Hofstadter 1985, 455-57). But this is clearly a misapplication of Heisenberg's principle, which relates only to *quantum phenomena*. In the macroscopic world, there is typically no problem of measuring the exact position of, say, a vehicle and its speed at the same time (Hofstadter 1985, 463-64). Unless it is a case of fallacious reasoning, the use of Heisenberg's uncertainty principle should only be *inspiratory*; there is no actual *argument* connecting it to the (often reasonable) idea that, say, the conducting of an experiment has an effect on the subject.

Similarly, pragmatist political theorists are often sceptical about general claims such as principles with universal application. But however reasonable that scepticism is, we should not invent an argumentative connection where there is none. Although we agree that this scepticism is often warranted, it would be a severe mistake to think that Brandom's theory (or any theory of language for that matter) *shows* or *suggests* that it is so. One virtue of Brandom's account of linguistic practice is arguably that it convincingly

demonstrates that we do not *need* general theories for meaningful discourse. Although this insight is certainly not new, and its consequences have been heavily debated in the particularist-generalist debate in moral philosophy (Hooker and Little 2000), we think that it may still serve as an *inspiration* for theorists aiming to argue for a contextual account of some key concept in political theory.

Fossen, however, thinks there is an argument present. In his reply, he seems to acknowledge that Brandom's *general* theory alone is insufficient to draw his conclusions. But we miss, he claims, all of the *particulars* of his Brandomian account, and it is from *them* that his conclusions allegedly are drawn (Fossen 2017, 6-7). So let us now have a look at these particulars.

II. SECOND FLAW: PRAGMATICALLY INDUCED PARTICULARS

As we have seen, Fossen argues that his *specific* analysis of political legitimacy has the power to open up alternative ways of looking at normative political theory. To rehearse Fossen's main idea: "In calling an authority legitimate, one attributes an entitlement to rule to that authority, while also undertaking a commitment to treat it in ways appropriate to its status … and attributing such commitments to other subjects" (Fossen 2017, 6).

This is fine as far as it goes; the problem is that this is not very far. Apart from the direct application of Brandom's expressive analysis (attributing a certain combination of commitment and entitlement to the participants), the specific role of legitimacy-claimings as expressing entitlements to rule is almost wholly uncontroversial, since it is more or less the lexical meaning of the term. Thus, we cannot see that any "conceptual room" is opened with this analysis. Few political theorists would deny that legitimacy has to do (at least partly) with entitlement to rule, and that in claiming that X's power is legitimate, a speaker undertakes some set of commitments. This room, it seems to us, is already visited by virtually all political theorists.

Moreover, since the *specific expressive role* of legitimacy-claimings (as opposed to the general role of claimings) is nothing we get from Brandom's account (or the account of any other philosopher of language for that matter), even in the case of the theorist who would *disagree* with the role Fossen ascribes to the concept of legitimacy – arguing for example that political legitimacy-claimings specifically express entitlement to use coercive power if the participants disobey – it is hard to see that she would disagree with anything that Fossen's account gets from the *pragmatist analysis* as such. That is, when the particular role of legitimacy-claimings is switched to whatever the specific theorist favours (X), it is hard see how she would *not* endorse the idea that one attributes entitlement to X, while also undertaking a commitment to treat it in ways appropriate to its status as well as attributing such commitments to others.

Consequently, Fossen's pragmatist linguistic analysis of legitimacy – his utilization of *general* aspects of assertion (Brandom's account) and his suggestions for the *specific* role of the concept – adds nothing to normative political theory that previously has been

missing. Let us therefore move to the third flaw in his reply, the idea that a fundamental linguistic level of analysis is of normative guidance for theorizing political legitimacy.

III. THIRD FLAW: CONFLATING DIFFERENT LEVELS OF ANALYSIS

In his reply, Fossen complains that we underestimate the relevance, in normative political theory, of thinking about how we theorize our concepts (Fossen 2017, 9). Certainly, Fossen is right when he says that "we (political theorists) often presume that we know what we are talking about when we speak of legitimacy, justice, and the like, and that we know what we are doing in talking about them," but that "the task of the political theorist is surely also to question such taken-for-granted notions, and to ask whether the ways in which we theorize them do justice to the phenomena in question" (Fossen 2017, 9).

The problem is that Fossen batters at an open door. Part of philosophy has always been to clarify the concepts we are using, and political philosophy is no exception. If we do not agree, at least approximately, about what our terms mean, we are not talking about the same things when we utter them. Therefore, the justification of a normative argument will always depend on how we should interpret the terms involved. An important part of normative analysis typically consists in specification, in more or less clear terms, of the concepts used, so what is claimed becomes as clear as possible. But this, again, has nothing to do with *pragmatism*, and particularly not with pragmatist theories of language as such.

The point of our distinction, in our previous paper, between substantive normative accounts and semantic accounts of a political concept, was to highlight how the latter can never constitute a normative argument. You need normative premises to make a normative argument. Fossen acknowledges this, but still tries to argue that pragmatist theories of language have normative relevance. If all he meant was that conceptual analysis may help us make a good normative argument by being clearer, we would have no objection, other than that this door is already open. However, this is not all he means. He further insists that his Brandomian-inspired conceptual interpretation of legitimacy somehow re-orients us towards political practice, that "pragmatism offers a promising, if as yet unfulfilled, avenue for pursuing a more practice-oriented approach to political theory" (Fossen 2017, 2).

We cannot see what motivates this conclusion. It seems as if Fossen thinks that the idea that we express commitments and entitlements (ours and those of others) by our legitimacy-claimings somehow calls for another perspective, one that is more focused on the practice on making such claims than what theorists have previously taken into account, and which precludes us from thinking that what is legitimate may have a general solution. However, this is not only to misunderstand Brandom, but to misunderstand the debate between pragmatists and other theorists in the philosophy of language. It is *not* contested that we express commitments and entitlements by our

use of various concepts, such that calling someone a 'terrorist' rather than a 'freedom fighter' entails expressing different sets of commitments towards that person. *Everyone* agrees on this. This is why, as we argued above, no political theorists we know of would disagree with Fossen's analysis of legitimacy in a way that matters. What *is* contested in the philosophy of language is what ultimately *explains,* or *accounts for,* meaning. Traditionally, philosophers of language do not think that the idea to start in the 'takings' of participants (what they *take* to be the case) in order to analyse 'beings' (what *is* the case) holds water. While traditional theories of language take the notions of representation and truth as primitives, Brandom's theory takes socially conferred normative statuses (such as commitments and entitlements) on the pragmatic side of language, and inference on the semantic side, as the basic building blocks.[1] These opposing camps thus face different justificatory challenges, and while they disagree about its theoretical *relevance*, the fact that we express commitments and entitlements by the use of various concepts is *not* controversial (Brandom 1994, xvi). At this point it should be clear why Fossen's account of legitimacy does not take us very far. What he attempts to argue is, in effect, that since every physical object, on a fundamental level, consists of atoms and molecules, a correct account of how we should arrange our houses must study the molecular structure of furniture. Naturally, we are *constrained* by the molecular structure of furniture, such that arranging a house using a material too sensitive to pressure or indoor temperature would not be suitable for, say, the kitchen chairs. Still, we need not conduct any chemical analysis to know that chairs made of ice would not be suitable in the normally tempered house. Similarly, as long as we reasonably may be interpreted as talking about legitimacy in our normative arguments, whether or not we should further study the practices of legitimacy-claims is an *open question*. Sometimes we might not need to know more than we already do to make judgments about the legitimacy status of attempts to rule, such as when we look at slave practices or other practices of domination and oppression and judge them illegitimate. Similarly, tyrannical rule may not become legitimate no matter how systematic and multifaceted our analysis is of the political activity enabling subjects to *regard* this authority as such. In other words, it seems that we cannot "explain what it is to *be* legitimate in terms of taking something *as* legitimate" in these cases (Fossen 2017, 6).

Once we move from the 'inspiratory' level of analysis to the level of first order normative theorizing – which, as we have seen, does *not* have to involve either criteria or necessary and sufficient conditions – what is required of the theorist is to offer normative reasons in support of an account of political legitimacy. Located in the space of reasons with this normative task at hand, however, the "fundamentally *critical*" (Fossen 2017, 10) role of Brandom's theory seems to be of no assistance at all to the theorist. Fossen is

1] Brandom, *Making it Explicit*. Cf. pp. xiii-xvii in the preface for an introduction of Brandom's overall strategy, and chapters 3 (esp. sections II and III), and 5 (esp. section I) for central parts of the more detailed account.

right that a focus on stance-taking always opens the *possibility* of "articulating a new way of understanding what we are doing," but he is wrong in believing that this will in itself enable us to do things *better* (Fossen 2017, 10). Because in certain social and political contexts, it may turn out that 'better' entails (in Fossen's 'taking-something-as'-sense) that an authority 'legitimately' eradicates people on the basis of the colour of their skin, or that actual power makes a rule legitimate. Fossen points to the affinity between pragmatism and political realism, but not even realists would accept that legitimacy can ever be 'might is right'.

By contrast, 'better' for the normative theorist means more justified from a normative point of view, e.g. that we ought to do X (better) rather than Y (not as good). As normative theorists, we are all in the same boat when it comes to first order theorizing: we need to argue by way of giving reasons, justifying our claims through other claims, eventually reaching 'the level of bedrock,' in this case the normative premises we take as valid but have no further reason for (Wittgenstein 1953, §217; Brandon 1994, 661).[2] So whether or not Fossen's pragmatist approach will turn out to be successful for theorizing legitimacy depends, rather than on Brandom's theory, on what limits to legitimacy he presumes are set by our social and political practices.

Numerous theorists in current theoretical debates share Fossen's commitment to stressing the importance of practices, but they explicitly attempt to answer these questions. According to political realists, for example, in order for a political order to be legitimate, the agreement or willing consent among the citizens must be perceived as free, and thus cannot rely on means that are too tyrannical or be the result of total deception (Horton 2010, 431-38; Williams 2005) According to practice-dependent theorists, likewise, the coercive practice of slavery can never be just since it is not capable of being "justified to all participants" (Sangiovanni 2008, 163).

To be fair, Fossen is clear about not having developed a positive account of political legitimacy yet. Still, he argues that developing such an account "would require developing a perspicuous representation of political practices, analogous to that offered by Brandom for discursive practices in general" (2017: 10). But, again, our point is that without normative glasses, such a perspicuous representation of political practices would be normatively toothless.

Hence, while we agree that conceptual analysis on a pragmatist outlook may have exactly the critical role that Brandom envisages, namely, explicating concepts to open them up to rational criticism, it does not help us in our endeavour to argue that female circumcision is wrong. For this normative task, language is entirely unreliable. For sure, the work of Brandom and Wittgenstein may still offer a theoretical framework for the type of political theory that attempts to find a more contextual answer to the question

2] Note that while each participant in discourse would eventually reach this bedrock state, for a Sellersian anti-foundationalist such as Brandom, there is of course no *justificatory* bedrock, no 'given' proposition that is justified by itself: c.f. e.g. Brandom *Making it Explicit*, 1994, 215-17.

of what is legitimate (Fossen) or democratic (Mouffe, Norval, and Tully), and so on. It may also supply a fitting terminology for describing what goes on in political action and discourse. However, we should not mistake this inspiration for an argument for such an account.

eva.erman@statsvet.su.se

Acknowledgements:

Eva Erman wishes to thank the Swedish Research Council and Marianne and Marcus Wallenberg Foundation for the generous funding of her research.

REFERENCES

Brandom, Robert. 1994. *Making it Explicit.* Cambridge: Cambridge University Press.

Erman, Eva, and Niklas Möller. 2014. Brandom and Political Philosophy. *Journal of Political Philosophy.*

———. 2015. What Not to Expect from the Pragmatic Turn in Political Theory. *European Journal of Political Theory* 14 (2): 121-40.

Fossen, Thomas. 2013. Taking Stances, Contesting Commitments: Political Legitimacy and the Pragmatic Turn. *Journal of Political Philosophy* 21: 426-50.

———. 2017. Language and Legitimacy: Is Pragmatist Political Theory Fallacious? *European Journal of Political Theory.* DOI: 10.1177/1474885117699977.

Hofstadter, Douglas R. 1985. Heisenberg's Uncertainty Principle and the Many Worlds Interpretation of Quantum Mechanics. In *Metamagical Thomas: Questing for the Essence of Mind and Pattern.* New York: Basic Books.

Hooker, Brad, and Margaret Little. 2000. *Moral Particularism.* Oxford: Oxford Clarendon Press.

Horton, John. 2010. Realism, Liberal Moralism and a Political Theory of Modus Vivendi. *European Journal of Political Theory* 9 (4): 431-48.

Mouffe, Chantal. 1999. Deliberative Democracy or Agonistic Pluralism? *Social Research* 66 (3): 745-58.

———. 2000. *The Democratic Paradox.* London: Verso.

Norval, Aletta. 2006. Democratic Identification: A Wittgensteinian Approach. *Political Theory* 34 (2): 229-55.

Sangiovanni, Andrea. 2008. Justice and the Priority of Politics to Morality. *Journal of Political Philosophy* 16 (2): 137-64.

Tully, James. 1989. Wittgenstein and Political Philosophy: Understanding Practices of Critical Reflection. *Political Theory* 17 (2): 172-204.

———. 2002. Political Philosophy as a Critical Activity. *Political Theory* 30 (4): 533-55.

Wittgenstein, Ludwig. 1953. *Philosophical Investigations.* Oxford: Blackwell.

Williams, Bernard. 2005. *In the Beginning was the Deed: Realism and Moralism in Political Argument*, ed. Geoffrey Hawthorn. Oxford: Princeton University Press.

Public Reason 8 (1-2): 91-95

Book Reviews

Joseph Fishkin, Bottlenecks. A New Theory of Equal Opportunity, Oxford University Press, Oxford, UK, 2014, Pp. 1+258, ISBN: 9780199812141

The theoretical landscape in which equality of opportunity occupies a central place has recently been enriched by the challenging perspective Joseph Fishkin presents in his book *Bottlenecks. A New Theory of Equal Opportunity*. The author, whose expertise lies in the area of political and legal theory, with a focus on discrimination and equal opportunity laws in contexts such as employment or voting rights, is defending an innovative interpretation of equal opportunity through the key concept of *bottlenecks*.

By "bottlenecks" Professor Fishkin understands the sensitive spots of the social structure where obstacles (many of them endogenous and not easily visible) are likely to block one's pursuit of socially valuable goals. For example, the specific screening procedures used by schools and employers to select individuals, as well as the criteria behind such procedures are likely to become "bottlenecks", unless regulated with a broader concern for social justice. In the author's own words, a "bottleneck" is defined as "a narrow place in the opportunity structure through which one must pass in order to successfully pursue a wide range of valued goals." (13) Therefore, not only the structure of college examinations and the content of admission tests or job interviews are likely to become "bottlenecks", but, more importantly, so is the weight given to morally arbitrary criteria such as race or social status.

The innovative character of Fishkin's view, however, lies in attempting to restructure the whole way philosophers and political theorists have taught us to interpret the topic of equal opportunities. In order to do so, he identifies the core reasons for which we are still facing arbitrary barriers blocking our access to opportunities, despite the professed commitment of many liberal democracies for substantive equality of opportunity.

According to Fishkin, the mainstream perspective on equal opportunities, shared by both scholars and policy makers is trapped within a "unitary model", where social order is mostly monolithic and characterized by a high degree of inertia. People have identical hierarchies of preferences and goals, which means they are competing for virtually the same prize. The same criteria are operational to define performance and desert with reference to this very narrow set of goals, and individual aspirations are nipped in the bud by social conformism and lack of significant alternative options. Because individuals are trapped in an opportunity structure which is "wholly external and fixed" (15), they are very vulnerable to incur losses as a result of zero-sum contests. Given that most desirable social goods and positions are accessible via the same paths for all individuals, and individuals are immensely diverse in terms of their capacities and initial life chances, the competition tends to favor those already favored by various contingencies. In fact, as he shows, this model would be very similar to the famous example of the warrior society given by the philosopher Bernard Williams in *The Idea of Equality*.

Simply adjusting local criteria of selection without aiming at comprehensively and profoundly addressing inequality of opportunity at its core fails to yield a just result, and, in contrast to the unitary model, Fishkin presents his own theoretical offer, *opportunity pluralism*. The pluralistic model is, first of all, sensitive to the real diversity

of individual capacities and aspirations (in fact, it is also an environment very conducive to maximizing this diversity, by encouraging individuals to manifest their potential knowing there would be ways for them to reap the fruit of their talents). It is also sensitive to the multitude of ways in which individuals shape their own goals, how they decide what to concentrate their energies for.

Unlike the pyramid structure of the unitary model, opportunity pluralism consists, in Fishkin's view, of four principles: (i) the necessity of having a plurality of social values and goals, (ii) reducing as much as possible the positionality of goods, and ensuring that the great majority of desirable social positions are non-competitive or less competitive, (iii) offering a diversity of paths which enable individuals to attain their goals, unlike in the unitary model, where they are constrained to follow the same path towards social success, and (iv) the existence of *"a plurality of sources of authority* regarding the elements described in the other principles." (131-32).

Much could be said of the feasibility of applying such a model in societies which do not function according to ideal conditions, and Fishkin is well aware of the difficulties faced by traditional democracies to ensure even a reasonable standard of social mobility. He is also aware of the challenge of limiting the conversion of goods so that there should not be dominant standards governing vast areas of social life, as made explicit by the case of instrumental – good bottlenecks, echoing Michael Walzer's concerns about dominant goods which alter the pluralism of a just society. Nevertheless, an important advantage of Professor Fishkin's theoretical proposal is that it helps the reader better understand why equality of opportunity should really matter.

This achievement should not be overlooked, since it is, in my opinion, one of the key contributions of Fishkin's theory. The reason we have been, so to speak, misled into defending various views of equal opportunities which are ineffective and generate untenable implications is because we have been lured by the perspective of distributive fairness. Two main authors whose views Fishkin describes and criticizes, John Rawls and Ronald Dworkin, have discussed equality of opportunity within theories of distributive justice, focusing on its significance for social cooperation and social cohesion, as well as for facilitating access to primary goods or resources. Nevertheless, such interpretations fail to capture the deeper relevance of equal opportunities, which Fishkin attempts to revitalize with the help of opportunity pluralism.

This relevance has to do with viewing opportunities not merely as a means of accessing resources, socially valuable goods and positions of advantage, but first and foremost as a means of expressing human individuality, a matter of one being capable to choose the life one really wants to live according to one's genuine values and aspirations. In other words, we should care about opportunities not only as an expression of distributive justice, but, perhaps more importantly, as an expression of human freedom, as a form of autonomy. This is how Fishkin invites us to think about opportunities:

Opportunities open up the freedom to do and become things we otherwise could not. […] Second, opportunities have a distinctive value because of the roles they play in shaping who we are. Opportunities shape not only the paths we pursue, but also the skills and talents we develop and the goals we formulate. We do not come into the world with fixed preferences, ambitions, or capacities, but develop all of these through processes of interaction with the world and with the opportunities we see before us. (2-3).

It is worth emphasizing that these two ideas, which form the basis of Fishkin's innovative contribution, were not alien to John Rawls, whose *Fair Equality of Oppor-*

tunity is, perhaps, the most complex normative model designed in political philosophy to this end. Although he accepted as legitimate the use of natural talents, efforts and motivation as criteria to distinguish between deserving and non-deserving individuals, Rawls could not fail to notice that, to some extent, even the ability to strive and make an effort, that is, to set goals, formulate a plan of life and develop the necessary motivation to pursue it, depends much on favorable circumstances. Moreover, he was well aware that addressing natural and social contingencies by mechanisms of equalization would inevitably clash with parental autonomy, and he expressed doubts that equality of opportunity could be feasible insofar as the institution of family exists. His strategy for tackling such issues was of restricting the scope of *Fair Equality of Opportunity*, but complementing it with the *Difference Principle*, designed as a mechanism to maximize the life prospects of the worst off after the lexically prior principles have been applied.

Although Professor Fishkin acknowledges Rawls's merits, he still thinks that he leaves large distributive inequalities to persist, and argues that mitigating these is not enough in terms of social justice. In fact, Fishkin argues for a possible way out of the conceptual problems affecting both Rawls and Dworkin by shifting the focus from the equalization paradigm to that of broadening opportunities (his inspiration seems to come here mainly from the liberal individualism of John Stuart Mill).

More importantly even, both Rawls and Dworkin are criticized for working with the underlying assumption of natural talents which we could identify and isolate with precision from various layers of interaction between individuals and the environment. From this perspective, Fishkin is radical in his suggestion that all there may be for theorists to compare are, in fact, "only different individuals with different combinations of characteristics and potentialities *every one of which* is the product of layers of past interaction between a person and her environment, with her developmental opportunities playing a central role in this interaction" (99).

In order to substantiate this hypothesis which would authorize the rejection of the distinction between natural and social talents, otherwise central in the arguments constructed to defend equality of opportunity, he relies, among others, on James Flynn's survey on the evolution of social intelligence and the criteria considered socially relevant for measuring QI. As a related step, he draws attention to the fact that assessing and recognizing others' capacities is likely to be encumbered by various cognitive biases and unconscious stereotyping, which affect not only social psychology, but also the labor market (111).

The third and fourth chapters of the book offer valuable insight into the mechanism of opportunity pluralism, which should reconcile the concern for social justice with cultivating human autonomy, as expressed by one's aspirations to flourish and one's genuine capacity of choice.

The third chapter takes a closer look at the distinction between the unitary and pluralistic opportunity structure. Unlike the unitary model, opportunity pluralism should create not only diverse motivations which would duly honor the individuality of human beings, but also diverse incentives, encouraging individuals to follow various paths in order to reach their goals. Anticipating perhaps the objection that, no matter how many paths society could open to individuals in order to satisfy the diversity of preferences entering life plans and the diversity of means they use in order to fulfill these, there is still no escape from competitive and positional goals, he uses the example of education in order to make his views clearer.

Opportunity pluralism will not diminish the value of education in a society where professional success is predicated on some amount and type of education, but, unlike in the unitary model, parents will make different choices in order to give their child advantage over other children. They will tend to pursue child development in an absolute, rather than relative sense, and will choose various combinations of means that help them reach such goals. They will have more flexibility and more independence in making choices for their children, because the level of social conformism has reduced and society offers at least a reasonable range of options to satisfy various aggregates of preferences.

Professor Fishkin also gives the example of advanced secondary schools in the German educational system, i.e. the Gymnasium, which is crucial in the path towards higher education. As he explains, this institution is conducive to bottlenecks because there are no alternative preparatory options along the road to higher education. Admission into Gymnasium is based on "strong academic performance and teacher recommendations in primary school", and the institution "dominates the path to higher education" (147-48). This virtually means that a child's academic prospects are more or less decided at a very early age, due to the fact that the educational system is such shaped, that it does not leave room for catching up on lost opportunities. Advantages or disadvantages capitalize and replicate many social contingencies which could not be properly leveled at entry stage, but which could be adequately and legitimately addressed once children have genuine opportunities to demonstrate their achievements without the constraints of such contingencies.

The example of US community colleges, which provide general education on a non-competitive admissions basis, so that those not willing or not able to enter a university could still benefit by a form of higher education and have access to better jobs is meant to contrast the case of the German Gymnasium. Such colleges respond better to an ideal of diverse preferences, as well as diverse capacities and are better equipped to support individuals which have not reaped the benefits of using their opportunities.

The chapter continues with a distinction between arbitrary and legitimate bottlenecks, and Professor Fishkin argues here that we should contextualize each time we are faced with this distinction by inquiring, for example, into the mission of a certain institution. The same requirement or criterion could function as a legitimate bottleneck if it mirrors a legitimate aim in the specific context of a competitive interaction but, on the other hand, it would be arbitrary should it exceed its sectorial application. This way of looking at bottlenecks helps us preserve the value of social efficiency occupying its own place in a theory which, having discarded the notion of natural talents, has nevertheless preserved those of competition and individual merit.

Going to the core of the problem, the chapter discusses control over the opportunity structure – again, one can appreciate the author's concern for a more practically-oriented discussion, testing the confines of normative theory. Merely claiming that pluralism should be observed as a core liberal value is not enough, Fishkin suggests, to generate genuine opportunity paths for individuals. It would be necessary to decentralize the opportunity structure, so as to avoid monopoly and misuse of power, rigidity and scarcity of the options that individuals should enjoy. Many different institutions should contribute to shaping the opportunity structure available to individuals, and, in fact, this role should not be restricted to institutions alone: individuals too should be allowed to use their creativity and capacity for autonomous decision-making in order

to define "new paths and even new roles and goods that did not exist before" (154). Such a measure, practically feasible, would also function as an additional guarantee that the value of ensuring broad opportunities to individuals would not be compromised by the same conservative interpretation which often sabotaged the equalization paradigm.

The last chapter of the book is dedicated to discussing various applications of opportunity pluralism. In presenting the implications of this theory for policy making, law and institutional design, Fishkin has mobilized a large quantity of fresh material from sociological studies, discrimination case law, empirical findings and theories in education. This is, doubtlessly, a significant advantage of his work, in that it helps the reader represent rather accurately some dilemmas likely to affect opportunity pluralism. Of particular interest is the section where Professor Fishkin discusses integration and segregation in schooling, as a salient example of how the effects of bottlenecks replicate across sectors, and also of the concatenation of developmental opportunities which shape one's life path.

The justice in education debate has much to gain from this part of the book, where Fishkin provides valuable insights into the complex interdependences which affect the opportunity structure: the peer effect problem and the "geography of opportunity" by which some residential locations are privileged in terms of resources, parents' input, the existence of networks spreading information and shaping choice, etc. Nevertheless, the complexity of the problems addressed does not find a correspondent in normative theory, but rather in a concatenation of practical measures: "an important part of the solution to inequality of opportunity may lie in public policy choices, such as progressive taxation, social insurance, and the provision of non-monetary endowments, that either reduce material inequalities or temper their practical importance, making the bottleneck they create less severe." (219).

It may be that this approach, attempting at a balance between theory and practical input, between the coherence of a conceptual framework and the interdisciplinary perspectives which opportunity pluralism opens, is at the same time the main innovation and the sensitive spot of Fishkin's work. A consistent theoretical commitment along the whole book, other than that of liberal individualism and human flourishing as a basis for broadening opportunities is difficult to identify. At the same time, the deeper the implications of opportunity pluralism in various areas of life, the more difficult it is to avoid the paradigm of distributive fairness against which the author has designed his own view.

Whether these difficulties are inherent in opportunity pluralism or, on the contrary, tend to affect any other conceptualization of the long cherished ideal of equal opportunities is open to debate. In this respect, Professor Fishkin's perspective is a significant achievement that not only reformulates the debate on such a salient issue, but by the diversity and quality of factual material gathered to illustrate the dilemmas of opportunity pluralism it has definitely enriched the interpretation of classical authors such as John Rawls and Ronald Dworkin.

Reviewed by Ileana Dascălu
University of Bucharest
ileana.dascalu83@gmail.com

PUBLIC REASON

Journal of Political and Moral Philosophy

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

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

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

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

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

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

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

ARTICLES

BOOK REVIEWS

http://publicreason.ro
ISSN 2065-7285
EISSN 2065-8958