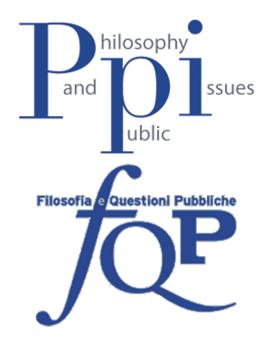# EDITORIAL PREFACE

BY

SEBASTIANO MAFFETTONE – LEAD EDITOR

GIANFRANCO PELLEGRINO – EXECUTIVE EDITOR

MICHELE BOCCHIOLA – MANAGING EDITOR

[THIS PAGE INTENTIONALLY LEFT BLANK]

Editorial Preface

Ethics and political theory are sometime described as forms of 'applied' philosophy because they focus on how people ought to live their lives, or how to regulate public behaviors. However, ethics and political philosophy are rarely considered outside academia: notwithstanding they are far from the high level of abstraction and pure speculation of other parts of philosophy (such as metaphysics or philosophy of language), most people don't think that these disciplines have anything to say about so-called 'real world issues.' Ethics and political philosophy are considered sophisticated intellectual enterprises, good for university classroom, with no practical impact whatsoever on people's life.
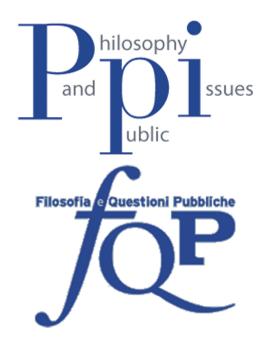
In his *Ethics and Public Policy: A Philosophical Enquiry* (London: Routledge, 2011), Jonathan Wolff explains how philosophy has a lot to say about concrete problems of everyday like, and he does it in a remarkably clear way. This Symposium of *Philosophy and Public Issues* tries to show how professional philosophers reason about practical issues, discussing ideas, comparing and contrasting different theories, trying to elaborate explanatory exhaustive and coherent philosophical views.

This Symposium hosts a discussion among Jonathan Wolff and Annabelle Lever, Enes Kulenovic, Fabienne Peter, Gerald Lang, Steven Smith, and Elvio Baccarini.

Sebastiano Maffettone

Gianfranco Pellegrino

Michele Bocchiola

Editors of *Philosophy and Public Issues*

A PRÉCIS OF
ETHICS AND PUBLIC POLICY

BY

JONATHAN WOLFF

A Précis of
Ethics and Public Policy

Jonathan Wolff

E*thics and Public Policy* was written for several different audiences: philosophers, academics in related fields, policy makers, students, and the 'interested general reader' if there are any left. It was also written for several reasons. On one level it was an attempt to account for what I had been doing for the previous decade or so. My first serious engagement with issues of public policy started around 1999 when I had the privilege of being invited to join a government committee looking to modernise gambling regulation in the UK (something I remain involved with to this day). After that ended, I was asked to work with the Railway Safety and Standards Board, who were grappling with the question 'how safe is safe enough?' given that it is always possible to spend more on safety and that the industry was suffering intense media criticism after a number of high profile railway accidents. Other invitations followed and I have found myself thinking about the regulation of recreational drugs, the law of homicide, accounting practices and sustainability, value-based pricing of pharmaceuticals, and many other issues, some of which I did not know existed until I was asked my opinion on them. In the great majority of cases the impetus to work on an area came from an invitation, rather than my own research agenda. Not all of these areas have led to chapters in *Ethics and Public Policy*, but my most extensive engagements have.

In one way I regarded this work as a side line to what I consider as my main work: philosophy and social policy, as distinct from public policy, by which I mean the study of disadvantage, equality and social justice. (There are some overlaps, in the case of health and of disability, for example.) Nevertheless, I found the topics I looked at fascinating, and working with civil servants, lawyers, journalists, world leading academics in other disciplines and people in charities and business provides a variety of refreshing perspectives, and a reminder that universities are very far from having a monopoly on intelligence or research skills. In each case I studied I came to my own view about the rights and wrongs of the subject under consideration, but also came to see how little my own view mattered. Policy debates are hemmed in by practical considerations of the politically feasible, the need to bring others with different views with you, and the irritating but fascinating ability of our fellow citizens to see ways round, or in the limit simply ignore, whatever regulations you recommend if they don't agree with them.

One purpose of the book, as I have said, was to record and bring together my activity over the years, some of which had been written up in papers, and some of which had not. Another, obviously enough, was to try to contribute to a number of substantive debates, such as the debate over the justification of punishment, picked up in Gerald Lang's contribution to the symposium. But a third, and perhaps the most important aim, was to reflect on what it is to make a philosophical contribution to a public policy area, which is the topic discussed, one way or another, by the other papers in this symposium. Prior to looking at this question in detail I had a fairly low opinion of applied moral philosophy. A typical contribution might identify a moral dilemma, attempt to work out what various forms of consequentialism and deontology would entail for the subject

area, declare one of the theories the winner and come to an end. It seemed to me formulaic and unhelpful, and not a way in which I wanted to spend my time. No doubt this is very unfair, but it nevertheless reflects the prejudice against applied philosophy that I held and the position that many philosophers continue to hold now.

The more I became engaged in public policy the more a different approach methodology forced itself upon me. First of all it is vital to understand the dilemma presented to you in the way that those it grips in real life understand it. What, exactly, are people worried about? Why has anyone asked a philosopher for help? Second it can often be very useful to find out how we got to the situation we are in. Very often I have heard policy proposals which, if implemented, would take things back to a situation which was previously found unacceptable and in need of urgent reform. Third, what empirical evidence is there concerning the field? How many people are dying in train crashes? What physical harm does heroin do to a human being if given by a safe method in a regulated dose? How do people gamble on slot machines? 'Common sense' in these and other areas is often distorted by pre-judgement and, worse, by media reports, which are motivated by the need to sell more copies or increase viewing figures, and stories are chosen on the basis of their headline potential rather than their significance for the subject area. Philosophers breathe in the same air as everyone else and need the same correctives from objective research.

To work in an area of public policy, then, can be time-consuming. It requires understanding policy, regulation, history, political pragmatics, and the role of the media and of campaigning groups, which are sometimes funded by 'sinister interests' on an apparently objective moral crusade but with the ultimate goal of knocking competitors out of business. Once all
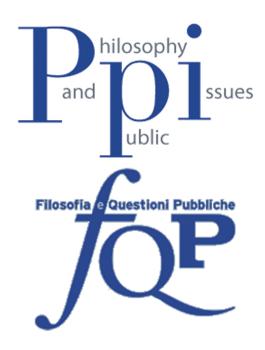
this is understood, it raises the question of what room is there left for philosophy? My answer is: not much, at least as traditionally done in the sense of appealing to a theory that will be 'sort out' a subject area. But on the contrary I think there is a lot of room for philosophy in the sense of the application of philosophical reasoning, training and intelligence. What this means in particular cases will vary according to the circumstances. One has to observe, listen, reflect and look for a way in. And one should not expect to be able to move things very much, although ultimately large and very significant changes can happen. *Ethics and Public Policy* is intended as one example of how philosophers can engage in a range of topics in public policy.

University College London

If you need to cite this article, please use the following format:

Wolff, Jonathan, "A précis of *Ethics and Public Policy*," *Philosophy and Public Issues (New Series)*, Vol. 4, No. 3 (2014), 3-6, edited by S. Maffettone, G. Pellegrino and M. Bocchiola

When the Philosopher Enters the Room
Comments on Jonathan Wolff's Approach
to Philosophy and Public Policy

BY

ANNABELLE LEVER

[THIS PAGE INTENTIONALLY LEFT BLANK]

When the Philosopher Enters the Room
Comments on Jonathan Wolff's Approach to Philosophy and Public Policy

Annabelle Lever

W hat can philosophy tell us about ethics and public policy? What can the ethics of public policy tell us about philosophy? Those are the questions that Jonathan Wolff addresses in his wonderful little book.[1] At one level, of course, the answer is straightforward—ethics is a branch of philosophy, so philosophy can tell us about the ethics of public policy, understood as a matter of deciding 'what we should do' in a manner that is institutionalised and collectively binding. But at another level, as Wolff shows, there is something deeply puzzling about the idea that philosophy can tell us anything very useful about public policy and about the ethical dilemmas that it raises. Those dilemmas arise as a result of political constraints, struggles and resources which generally have little to do with philosophy and, for the most part, rather little to do with ethics. In fact, as Wolff shows, philosophers have as much to learn about ethics and public policy as other people. What they have to offer in return is less some nugget of pure gold, to be offered to a grateful public eager for enlightenment, than a few useful distinctions, qualifications and questions (201) which might help intelligent, conscientious, public-spirited people to decide some tricky, but usually unglamorous, questions of policy. And that, basically, is what is so appealing about Wolff's book: that it departs so

---

[1] Jonathan Wolff, *Ethics and Public Policy: A Philosophical Inquiry*, (London: Routledge, 2011). Page numbers in the text are to this book.

strongly from what we might call 'the philosopher as mouth of God' approach to ethics, and situates itself in a variety of committee rooms, where tables are stacked with reports and statistics, enlivened only by the odd foray for investigative visits to a laboratory that does experiments on animals.

As Wolff shows, the difficulty with the 'philosopher as mouth of God' approach to ethics and public policy is less that God is dead—if s/he is—but that however true, elegant and persuasive our philosophical theories, they have to be implemented in a world where truth, elegance and philosophical persuasion are not the most important values. Wolff's aim, therefore, is not to present a general theory about how philosophers should approach public policy, still less to present a previously-worked out theory about ethics for application in particular cases. Rather, his book seeks to give us a sense of what it is like to think about giving advice on public policy to others who will bear the responsibility for carrying it out; and to help us to see both the interest and appeal, but also the very real limits, of that role.

Wolff is a distinguished political philosopher, with long experience as a teacher of political philosophy, as well as a researcher, seeking to push the boundaries of his field. So, while Wolff makes it clear that his approach to ethics and public policy is 'bottom up,' rather than "top down," (9), in many ways Wolff's approach in this book is a development of themes and concerns which can be found in his previously published work.

In *Disadvantage*[2] Wolff suggests an alternative to two standard ways of thinking about political morality—the one, which would reduce it to a matter of cost-benefit analysis; the other which turns every question of political morality into a problem of

---

[2] Jonathan Wolff and Avner de-Shalit, *Disadvantage* (Oxford: Oxford University Press, 2007).

justice. In their place, Wolff asks us to take seriously the needs of those who are disadvantaged in our society, and the difficulty of determining the nature, severity and remedies for their situation, whether or not meeting those needs is economically efficient, or required by our favoured principles of justice. It is not that we should be indifferent to questions of economic efficiency, according to Wolff, nor to questions of justice. Resources are scarce and it is wrong to waste them. But, for Wolff, cost-benefit-analysis is not the right way to think about the ethical use of scarce resources, because there is no universal metric we can use to measure and compare the costs and benefits of different policy proposals, or even different states of affairs. (89-107) How bad it is to be a drug-addict rather than homeless, or to be paraplegic rather than hungry just isn't the sort of thing which we can answer by measuring amounts of badness, as though we were comparing quantities of water, or depths of colour in some scientific experiment. So cost benefit analysis—whether as practiced by economists or by utilitarians—is going to have only a subsidiary role, at best, in the ethical analysis of public policy.

Likewise, Wolff argues, it is a mistake to reduce questions of political morality to questions of justice, even when we are concerned with the plight of people whose situation may reflect individual or social injustice. Even if people are poor, hungry, homeless, addicted or injured because of their own foolish or self-destructive behaviour, Woolf argues, we should be morally concerned about their plight. Duties of justice are certainly important—the minimum that we owe others—though that minimum, unfortunately, is often so demanding that we are unable to meet it individually or collectively. But for Wolff, we have duties to each other which are not simply duties of justice— duties of compassion, consideration, decency, humanity, respect and solidarity. There are many different ways that we might describe such duties, but their point is that these are things we

owe to each other whether or not our society is just. As Wolff and de-Shalit put it, à propos those homeless and sleeping rough on our streets:

> Perhaps when this phenomenon started people were genuinely moved. But it seems fair to say that Western societies are becoming apathetic….But we should not allow ourselves to get used to it. In other words, we want to claim, the state of the least advantaged is a mark of shame and speaks poorly of Western societies. It should be a call to action not only because some people get or own less than others, but also because it implies that social relations have deteriorated to an inhuman state (pp. 7-8).

Wolff's adoption of a 'bottom up' approach to ethics, then, does not mean that we can do without theory, or that we must be radically particularist as philosophers. The point, rather, is that issues of ethics and public policy are not reducible to questions of justice. And so, when we are concerned to know what types of animal experimentation, if any, should be legally tolerated, or what forms of gambling, if any, should be legal, we need to know what injustices follow from actual policy and are likely to occur with the main policy alternatives to the status quo. But we *also* need to consider the range of moral and political ideals in our society, and to consider how, if at all, they bear on the presentation and evaluation of the main policy alternatives. As this range is only partly a matter of philosophy—being influenced also by tradition, by chance, by struggle and by the unintended outcomes of past events and choices—philosophers need to ask themselves how best they should approach the constraints and opportunities created by the circumstances within which their advice has been sought, and for which it is supposed to work. Precisely because this latter topic has received so little attention from philosophers, we are especially lucky to have Wolff's reflections on this matter, presented with his customary clarity, modesty and humour.

Wolff's conclusions can be briefly summarised. Philosophers must be prepared to get their hands dirty, by studying the policy area in question, in order to understand why there is an ethical problem that needs resolution, to understand what remedies for resolving it have been tried, as well as to understand the constraints on any possible solution (81-82, 196-7). This is all hard work, and the results can be disappointing. "Most members of the British public seem happier with a policy that legalizes alcohol but not ecstasy, inconsistent though it seems, than they would be with a consistent policy that treats them both in the same way. To the philosophical mind, this is infuriating, but it is the everyday world of public policy" (198). Philosophers must therefore learn to accept that the test for good law and policy is not the same as the test for a good philosophical argument, because consistency in policy is a matter of psychology and institutional design—whether people can actually do the different things that laws require of them without generating counterproductive consequences—and not a matter of what people can believe without logical contradiction. (82, 200)

Moreover, "policy failure is always a threat," because policies often have unanticipated consequences. "One reason for this can be that policies are advanced on the basis of insufficiently examined empirical assumptions," and because policies often depend on false assumptions about human motivation and, particularly, about the enforceability and effectiveness of legal regulation (199-100). Attention to the nature and quality of the examples used to illustrate and test competing policy positions is therefore essential to the ethical quality of public policy, and is one area in which the traditional skills of philosophers can be used effectively to improve public policy (102-108). Because the world is an unpredictable place, "if an argument is to be based on empirical premises, speculation may well run ahead of evidence […] Empirical claims need empirical support, not the support of

common sense, however common and however sensible it may seem" (60). Above all, it is important to realise that while we can be moved by moral arguments to question our behaviour and to feel guilty about it, getting people to change their behaviour is a more complex matter and, generally, requires structural change of a rather concrete (and often expensive) kind (201). "Many people would like to drive less, but they will not do so unless there are safe cycle routes or reliable and quick forms of public transport. Progress in this area, as in so many others, requires social and material change to accompany moral argument."

I am sympathetic to all of these claims and, particularly, to the importance of trying to understand the historical background against which a particular problem of public policy occurs. Partly what is at issue, here, is our response to the threat of "false necessity," as Roberto Unger calls it;[3] but what is partly at issue is our understanding of the stakes in a particular policy. Attention to the circumstances of past choices, the reasons for which they were made, and the differences between expected and actual results are an indispensable antidote to mistaken, and often, self-serving, claims of necessity. However, they also highlight the importance of attending to what people say they need and want, even if people can be mistaken about what these are, or about their moral and political weight.

Wolff's discussion of the limitations of philosophical approaches to disability, which conceives it as an issue of distributive justice, illustrates the importance of listening to what people say they want. Conceptualising the claims of the disabled through the lens of distributive justice alone makes it seem as though we are morally obligated to compensate people for those forms of disability which we cannot remove. But this is to

---

[3] Roberto Mangabeiro Unger, *False Necessity: Anti-Necessitarian Social Theory in the Service of Radical Democracy* (Cambridge: Cambridge University Press, 1987).

misunderstand what disabled people want policy change to do for them, as well as how it might do it. Their demands of public policy are not that they be able to do all the things that they would be able to do if they did not have the disabilities they have, or to be compensated for the differences in their life chances, but that they not be treated worse than other people, that the world should not be *especially* unjust to them. (154, emphasis in the text). Wheel-chair friendly environments, for example, do not erase the real and significant differences between needing and not needing a wheel-chair. But they do mean that wheel-chair users are not excluded from activities that they would otherwise enjoy, from which they might benefit and to which they have much to contribute, because we wrongly assume that justice for the disabled is primarily about the distribution of special resources, or of 'fixing' those who are disabled, instead of removing handicaps created by familiar ways of thinking and behaving.

Still, there a couple of gaps in Wolff's approach to ethics and public policy which I hope he will fill in a sequel to this book. The first concerns his relative indifference to questions of process and procedure, even though he urges philosophers to be more attentive than they generally are to what is involved in changing the world (192). The second concerns problems of evidence and how we should handle them. The two issues are partly connected, in so far as the ways we approach problems, and the people who are involved, can create, or alleviate, informational problems—as Wolff recognises when highlighting the significant contributions that disabled activists have played in changing the terms of debate on mental and physical disability. But procedural and evidential issues are not identical and it helps to approach them separately.

Most of us will not be asked to participate on committees designed to formulate or alter public policy and one's primary concern, if asked to do so, is to address the substantive problem

on which one has been consulted. Reflections on political or administrative procedure may look like a distraction from this task. Just as it would be a mistake to confuse what is good policy in an ideal world from what is good policy right here and now, so it would be a mistake to confuse idealised forms of procedure with those which are possible here and now. Still, it is striking that Wolff has nothing to say about the latter, and its potential relevance to the former, or any advice to give philosophers about how they should think about the procedural aspects of public policy making. In part because Wolff's presentation of the issues starts once he is on the committee, or in the room, questions about the differences between the ideal and the 'realistically utopian' aspects of process and procedure do not appear. And yet who is asked to participate on a decision can matter to the decisions that are made, and how decisions are presented and how much time there is for discussion, also matter to outcomes. The questions here are not just 'who should be consulted?' or how far, as an advisor, one is free to bring up such issues. In addition, there is the question of how far transparency about selection procedures and the reasons they were adopted (because they are desirable, or a compromise, or a matter of habit) affects the freedom of advisors and, thereby, the decisions that they reach.

Take, for example, the Citizens Council, which forms a part of the process by which the National Institute for Health and Clinical Excellence (NICE) reaches its decisions about what forms of medical care ought to be available on the National Health Service in the United Kingdom.[4] The Citizens Council is

---

[4] The following paragraphs are based on P. Littejohns and M. Rawlins (eds.), *Patients, the Public and Priorities in Health Care* (Oxford: Radcliffe Publishing, 2009) and on my "Democracy, Deliberation and Public Service Reform," in Henry Kippin, Gerry Stoker and Simon Griffiths (eds.), *Public Services: A New*

deliberately selected in order to provide a microcosm of citizens with no special knowledge or experience of the NHS. Every effort is made to make sure that financial concerns, caring responsibilities, disability or distance do not prevent people from volunteering or participating; and to ensure that discussion is not dominated by a few members. But it seems likely that members of the Council are not informed about the efforts that have been made to ensure that they are a more descriptively representative body than is typical of most advisory or decision-making bodies, nor does it seem likely that they are told why such descriptive representation might matter to their advisory role, or to the advice that they are asked to provide. As a result, both their report on age as a factor in the distribution of healthcare resources and their report on inequalities in healthcare were marred by a simplistic equation of equality with 'identical treatment.' Thus, the view that 'positive discrimination [is still discrimination]' led the Council to reject preferential spending on children rather than the elderly, and to reject efforts to target health care resources in order to minimise the very significant effects of social inequality on life expectancy, general health and access to healthcare in the UK.[5]

Reform Agenda (London: Bloomsbury Academic Press, 2013), ch. 6 pp. 91 -106; also available at

http://clients.squareeye.net/uploads/2020/documents/0921TWE_ESRC_democracy_050730%20C.pdf

[5] See Brian Brown, "The View of a Council Member," P. Littejohns and M. Rawlins (eds.), *Patients, the Public and Priorities in Health Care*, pp. 125-8, 128. He is reporting not only his own view of equality, but that of all but three of the thirty Council members. The problem, of course, is that if you define 'positive discrimination' as intrinsically bad, you still have to show that differential distributions of resources are an example of positive discrimination so understood, rather than, say, a positive attempt to maximize the impact of scarce resources, or a remedial attempt to mitigate unfairness.

Of course, if part of the problem here is an inadequate conception of equality, it could be addressed by ensuring that the expert advice to which members of the Council listen, prior to their deliberations, includes advice on the different ways that key concepts like freedom, equality and rights can be understood. But what is striking in this example is that a bunch of intelligent, thoughtful people were not aware of the limitations of this way of thinking about equality as a result of their discussions. One of the findings of an ethnographic study of Council members, commissioned by NICE, was that "A Lack of clarity about the grounds on which citizens could legitimately speak, and pressures to not generate conflict, meant that while differences of class, ethnicity, gender, disability and age were visible to all, these identities were not […] explored."[6] That makes me wonder whether those reports would have been the same had Council members been aware of the reasons why NICE thought it so important for the Council to be a descriptively representative body. I would therefore have been interested to hear what ideas, if any, Wolff had about the relationship between procedure and outcome in the ethics of policymaking, and how far he thinks that philosophers have anything useful to contribute as advisors by standing back from questions of substance in order to consider the ethics of procedure.

The second issue on which I would have liked more discussion concerns the ethical issues raised by the politics of knowledge and information. As Fenton, Brice and Chalmers note, patients' and clinicians' priorities for research very rarely match those of researchers. Whereas the former frequently want to know about the likely results of physical therapy or surgery,

---

[6] See C. Davies, M. Wetherell and E. Barnett, "A Citizens Council in the Making: Dilemmas for Citizens and their Hosts," P. Littlejohns and M. Rawlins (eds.), *Patients, the Public and Priorities in Health Care*, 129 -138, p. 133.

researchers overwhelmingly study the effects of drugs, and pay little attention to patient interests in access to good information on how to cope with chronic or disabling conditions. But "in a research world where perverse incentives often determine what research will be done" it is unclear how "the information needs of patients and clinicians can achieve more prominence."[7]

What counts as knowledge is not wholly independent of politics, and whether they are philosophers or policy makers, people are rarely in a position to evaluate the quality of the evidence with which they must work. How then should we handle competing evidence if we are concerned with the ethics of policy, and how far may we solicit additional information from people we trust?

Until recently, the gold standard for clinical trials of new medicine was what is called the 'double-blind trial,' on which doctors and scientists are as ignorant as patients about who is receiving the treatment to be tested and who is receiving a placebo.[8] But what it takes to make sure that the only difference between two groups of people are the differences between a medicine and a placebo means that these sorts of studies are generally of very short duration, and are only done on a very

---

[7] M. Fenton, A. Brice and I. Chalmers, "Harvesting and Publishing Patients' Unanswered Questions about the Effects of Treatment," in P. Littejohns and M. Rawlins (eds.), *Patients, the Public and Priorities in Health Care*, 165-180, 166-9.

[8] The following paragraphs are based on John Worrall's "Why There's no Cause to Randomize," *British Journal of Philosophy of Science* 58.3 (2007), 451-488, his "Evidence in Medicine and Evidence-Based Medicine," in *Philosophy Compass* (2007) and "What Evidence in Evidence-Based Medicine," *Philosophy of Science* (Sept. 2002). See also Sir Michael Rawlin's *Harveian Oration to the Royal College of Physicians*, October 2008, which can be found at http://www.rcplondon.ac.uk/pubs/contents/304df931-2ddc-4a54-894e-e0cdb03e84a5.pdf . Sir Michael was Chairman of NICE from its foundation in 1999 until 2012.

special population—typically, relatively young men, who suffer from one disease, but with no complicating factors to distract. These highly artificial conditions have increasingly come to seem an epistemologically unreliable source of knowledge about how a drug will behave when prescribed for people who do not fit such narrow criteria and circumstances. Not all forms of useful medical treatment can be subject to double-blind trials and, as Wolff notes, some of the best forms of health-care and maintenance are not specifically medical at all—things such as clean air, water, sanitary houses, safe workplaces, equality for women and efforts to combat the coercion and poverty which often blight their lives and the lives of their children and grandchildren (130). Thus, the idea that double-blind trials are the top of a pyramid of knowledge inevitably privileges some types of medical care and some types of health needs over others, for no good reason.

Abandoning the idea of a hierarchy of knowledge, however, highlights the difficulty of determining what weight to attach to different types of evidence—qualitative or quantitative—and what, if anything, one can or ought to do when faced with the potentially self-reinforcing links between power and knowledge. Wolff's book has little to say about the ethical significance of power differences for public policy and, unfortunately, offers little guidance when, as with issues of information and procedure, these differences of power frame public policy, rather than being its object more directly.

To see the point, let's return to the ecstasy/alcohol comparison. In the late nineteenth century the association of the brewing industry with the Conservative Party was well-established. Nowadays, representatives of the brewing industry probably give money to both political parties, and try to make it difficult for either one of them to oppose the damage caused by

alcohol in our societies. As with the National Rifle Association, which has successfully fought efforts seriously to limit legal access to guns in America, so every effort at treating alcohol as a dangerous drug is equated with Prohibition, with hostility to pleasure, and with a willingness to sacrifice the interests of the many because of the foolishness or bad luck of a few. If inconsistency in our policy preferences and ignorance about relevant information are not acts of nature, but the results of politics, what difference should this make to our search for common ground in the ethical formulation of policy?

Wolff's book is a welcome addition to the literature on method in political philosophy and to philosophy and public policy. I hope that there will soon be a sequel, in which Wolff extends his analysis of what, to paraphrase Rawls, we may call "the circumstances of policy,"[9] in light of what happens before the philosopher enters the room.
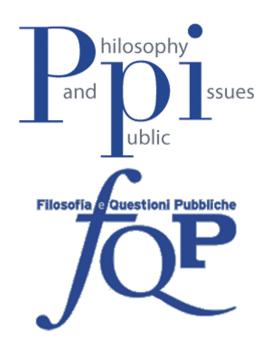
University of Geneva

---

[9] John Rawls refers to "the circumstances of justice" in *A Theory of Justice*, (Cambridge (MA): Harvard University Press 1971), Ch. 3, § 22, 109 -112.

Political Philosophy and Public Policy: Six Models

BY

ENES KULENOVIC

Political Philosophy & Public Policy: Six Models

Enes Kulenovic

I will begin my presentation the same way Jonathan Wolff begins his book *Ethics and Public Policy*: there is a phone call from a certain government agency asking you to participate in a team creating a new public policy—on gambling, treatment of animals, health reform, railway security—take your pick. After a long deliberation—about 30 seconds—you say yes, but then a moment of reflection kicks in: what can political philosopher actually contribute to the process of creating a public policy? What can philosophers offer that economists, sociologists, political scientists, legal scholars and experts in that specific field—can't? Obvious answer is that we have a better mastery of abstract concepts, ideas and arguments employed in policy-decision processes. But what does that actually mean? How can we translate this mastery into something useful to policy making? In short, what is the role of a political philosopher involved in designing a new policy or in defending the existing one?

In this article I'll try to answer this question or, to be more precise—as the title suggests—I'll offer six different ways we can answer this question. Each of these six answers will correspond to a model of political philosopher involved in discussing public policy issues. To make things a bit more coherent, I will look at three different things in each of these models: 1) what is the goal that this specific model of political philosopher is trying to achieve, 2) what are the advantages and 3) what are the disadvantages of each model. In my discussion I will rely on

insights and examples from Wolff's book *Ethics and Public Policy*. Although Wolff in his book poses a same question I do in this article—what role should philosophers play in policy-making process—his focus, it is worth nothing, is somewhat different from mine. His main goal is to provide a number of practical insights derived from interplay between philosophy and public policy and also to tell us how these practical insights might transform some of our normative presumption, making them more relevant for the world we live in. Wolff has a clear preference for more realist, 'bottom-up' approach rather than more idealist, 'top-down' approach. However, this preference is based on his understanding of the most fruitful way for philosophers to think about and contribute to policy issues. In what follows I will leave that question aside and try to look at how different (and sometimes conflicting) views on what are the more general aims political philosophers should strive to achieve translate into different models philosophers follow when contributing the policy-making agenda. Major part of the article will deal with describing and comparing these different models. The concluding part will offer a suggestion how we could go about evaluating each of these models.

I
Syracuse Model

The Syracuse model's answer to the question posed above— what is a philosopher's role in policy-making process—is the following: bringing the truth of a rational argument to the table. The political philosopher who embraces this model is on an enlightenment mission: he wants pubic policies to be directed by a philosophical principle or set or principles (think of Kant's "treating persons as ends, never as means," Bentham's "greatest

happiness of the greatest number" or Mill's harm principle). What should be avoided is political bargaining, populist tendencies or incoherent claims of influential moral traditions. Syracusean (if we can call him that) philosopher is a true philosopher-king that brings the light of rational and philosophically coherent principles to dirty business of policymaking. This model take its cue from Leo Strauss' view essay "What is Political Philosophy?" where he argues that "[p]olitical philosophy will then be the attempt to replace opinion about the nature of political things by knowledge of the nature of political things" but also "the attempt truly to know both the nature of political things and the right, or the good, political order."[1] Social scientists and experts on policy issue being discussed can tell us more about efficient means necessary to achieve certain ends, but it is the political philosophers who can tell us what ends are worth pursuing. It is their job to explain what a just policy that is compatible with an ideal of common good should be.

So, what are some of the advantages and disadvantages of this model? Syracuse model political philosopher, when it comes to deciding policy issues, should be, at least in his own view, the main gal or guy in the room. Any revision or radical change in existing policy will have to begin by taking into account the abstract principle (or set of principles) that philosopher puts forward. In top-down process of policy-decision making we can't avoid looking at the concepts and principles that are on very top and therefore, can't avoid giving the philosopher a central role. However, possible disadvantage is that this model seems to work only if we take a top-down approach, but as Wolf points out in his book, we should be skeptical about this type of approach.

---

[1] L. Strauss, "What is Political Philosophy?," in M. Cohen and N. Fermon (eds.), *Princeton Readings in Political Thought* (Princeton: Princeton University Press, 1995), pp. 643-644.

Applying general philosophical principles to policy issues can often backfire: either 1) because we are trying to make facts fit our theories or 2) because we are refusing the see the full consequences of applying abstract theories to real-life challenges. Now, the advocate of Syracuse model can argue convincingly that his approach doesn't need to sweep facts under the rug or deny the complexities of real-life challenges. He can even accept a certain level of compromise when it comes to formulating the policy in question due to these kind of challenges: the government that wants to be re-elected will be reluctant to push forward a policy that is fair and just, but goes against the opinion of the majority of voters; fiscal limits that determine the limits of applicability of the policy; discrepancy between the levels of inconvenience that voters are ready to accept and what new policy requires; etc. The Syracusean philosopher can take all of these factors into account and still consistently argue that just principles might not always be realizable, but they still offer a measuring standard by which we can judge the success or a failure of certain policy.

More important disadvantage of this model arises from the fact that it rests on two premises that most political philosophers would reject today: 1) that there are objective principles that work as knock-down arguments when applied to policy issues and 2) that we can convincingly show that there is a single theoretical principles (or set of principles) that is superior to all other options. In the first case, take Wolff's example of calculating the costs of ensuring higher standards of railway safety. We can use a moral standard of sanctity of human life which would exclude the option of putting a monetary value to human life, but at the end of the day when we have to make a decision how much are we ready to pay to lower the chance of preventable deaths (by, for example, putting additional barriers next to train tracks or train doors that open only when trains stop). Such a decision,

inevitably, leads to putting a monetary value on human life. The second case is even more troubling, because the whole enlightening mission of the Syracusean philosopher is put into question if we take value pluralist epistemology seriously and accept that, for example, both a consequentialist and a Kantian approach to a certain policy issue can be seen as philosophically coherent and just, although they are also mutually exclusive.

II
Rawlsian Model

This model of political philosopher—both epistemologically and ethically—is less ambitious than the previous model, but it is also more realistic. First, Rawlsian doesn't seek to affirm the Truth with a big T and proclaim *sub specie aeternitatis* what objective principles certain policy has to embody to be considered justifiable, but relies on underlying core values and concepts that are imbedded in the moral and political tradition of specific community, such as particular understanding of equality or liberty. Second, this model allows much more flexibility than previous model in a sense that it doesn't strive to impose one particular principle (or set of principles) on the issue that is being debated, but rather tries to exclude certain arguments from the policy-debate by declaring them unreasonable or irrational. As the name suggests, this model is, basically, what you get if you apply philosophy of later Rawls to policy-making process.[2] The stating point for this model is to call upon basic values that political community identifies with (usually in the letter of the constitution), rather than offer a unique insight into philosophical truths. This allows for the next step where those policy proposals that are contrary to these basic values are dismissed.

[2] J. Rawls, *Political Liberalism* (New York: Columbia University Press 1996).

Philosopher's main job is to make sure that whatever policy solution we decide upon in the end, that solution is compatible with society's shared understanding of what is just, fair and right. Also, philosopher should point out which kind of arguments should be taken as valid and which should be discarded in policy debates (for examples, in secular and rationalistic societies scientific arguments should carry much more weight than religious arguments).

The advantage is that, unlike Syracuse model, Rawlsian model can accommodate more then just one option and even strive to reach a compromise between these different options. Also, it doesn't necessarily rest on top-down approach; it can more easily incorporate empirical insights without interpreting them so they are compatible with abstract philosophical principles. Again, philosopher's job is to deal with ends rather than means of specific public policy. However, this time around he is not the one who necessarily has to propose the goals we should strive to achieve, his main task is to filter different proposals offered by public or experts and explain why some of these proposals are unacceptable.

Possible disadvantage is that this model inevitably has to establish a standard—in case of Rawls and his followers it's reasonableness—which allows it to exclude certain options and certain types of arguments invoked to justify these options. The fact that advocates of this model often disagree what that standard should actually be—just take existing disagreements that exist among the leading advocates of Rawls' political liberalism approach on how we should define reasonableness[3]—suggests that decision which options are in and which options are out can

---

[3] G. Gaus, *The Order of Public Reason* (Cambridge: Cambridge University Press, 2011); J. Quong, *Liberalism without Perfection* (Oxford: Oxford University Press, 2011).

seem rather arbitrary. Do we allow religious arguments to enter policy debates or not? Do we insist that justice should always have priority over stability? Also, when it comes to discussing the basic values we should rely on, this model seems to have in-built bias for status quo: we are looking at values that we have traditionally relied upon and that are already widely shared by citizens. This historical and democratic perspective on values can lead to exclusion of new perspectives advocated by minorities. Take the example of animal welfare, which is one of the issues that Wolff devotes a whole chapter in his book. Advocates of animal liberation or animal rights can find themselves in disadvantage because the existing policies that regulate testing on animals do not seem to go against neither the tradition nor the moral views of the majority.

III
Value-Pluralist Model

The guiding light for this model is one of the last sentences of Isaiah Berlin's famous essay "Two Concepts of Liberty" (taken from Schumpeter): "to realize the relative validity of one's conviction and yet stand for them unflinchingly is what distinguishes a civilized man from a barbarian."[4] Unlike the Rawlsian model, it's goal is not to exclude certain positions and arguments from the debate, but to take them all into account and then advocate one of them, while pointing out that every choice, even the one pluralist advocates, entails that something of value will be lost: what Berlin called "the tragic choice". Value-pluralist's mission is threefold: 1) to give the best philosophical defense for different options being discussed and then 2) show

[4] I. Berlin, "Two Concepts of Liberty," in *The Proper Study of Mankind* (New York: Ferrar, Straus and Giroux, 1997), p. 242.

that there is no one superior principle that can make our decision easier, that different options are, basically, incommensurable and 3) give solid arguments for one option, while pointing out what is going to be lost.

Let us turn to advantages and disadvantages of this model. Value pluralist model allows different options to be discussed without eliminating any one of them *a priori* and giving them all moral weight (or most of them, some options are just ridiculous, deeply immoral or irrational), so that even if one option has to be taken, others are not just swept under the rug. It also takes into account that introduction of new policies, even if policy proves to be a success, comes at the price and that promoting some values means that some other values will have to be sacrificed. Take another example from Wolff's book—that of gabling laws. More regulation that the state imposes on gabling, the more likely will certain negative social problems that are usually connect with gambling decrease. On the other hand, such regulation inevitably limits the personal freedom of citizens and promotes a more powerful and paternalistic state. The main disadvantage of this model arises from a criticism that was often directed towards Isaiah Berlin: if different options are incommensurable, on what grounds do we chose one option over all others? Value pluralist can tell us what can be lost when we indorse one policy option over another, but he can't really tell us which policy to choose.

IV
Proceduralist Model

This model is very similar to the value-pluralist model, in that it also accepts that there is no one best way to deal with a policy dilemma and that we might not able to rank different points of view on some pre-determined scale. One important difference

from the previous (Berlin inspired) model is that it doesn't strive to find what he thinks is either the best or the least painful option, its goal is only to give different options a same chance to be heard. Its main motto, following Stuart Hampshire, is *audi alteram partem*.[5] In his book titled *Justice is Conflict*, an extended version of his Tanner lectures, Hampshire summarizes his main argument in the following way: "Particular institutions, each with its specific procedures for deciding between rival conceptions of what is substantially right and fair, come and go in history. Only the one most general feature of the process of decision is preserved as the necessary condition that qualifies a process, whatever it happens to be, to be accounted as an essentially just and fair one: that contrary claims are heard. An unjust procedure, violating this necessary condition of procedural fairness, is unjust always and everywhere and without reference to any distinct conception of the good."[6]

The role of proceduralist political philosopher involved in public policy reform or creation process is to make sure all the relevant sides involved in the issue regulated by policy in question have their say. One could ask: why do you need a political philosopher to do that? In democratic societies where free speech is guaranteed isn't it better to allow different interest or social groups to voice their own concerns about a certain policy? That might as well be true, but there at least three cases where a presence of a proceduralist on policy-making body can prove useful: a) where the group whose interests might be endangered is not mobilized enough to let it's voice be heard and b) when power relations between different parties are so skewed on one side that the voices of those on the other side are completely

---

[5] S. Hampshire, *Justice is Conflict* (Princeton: Princeton University Press 2000), p. 8.
[6] Ibid., pp. 16-17.

muffled and c) when interests of a minority groups are not fully voiced by the spokesperson or representatives of that minority. Here, Wolff gives an example how a discussion of disability in the US has been dominated by war veteran invalids, skewing the policy discussion towards one type of disability (persons in wheelchairs). He offers a following warning: "just as we must pay attention to examples of people with disabilities we must also not allow the debate to become completely dominated by those with a greatest public presence or sympathy or strongest lobbying group."[7] In short, proceduralist, when faced with these three cases, can act as a mouthpiece for powerless or can make the playing field more even.

This model has the same advantages as the value-pluralist model, without having to explain why we choose one option over other incommensurable options. Its main tasks is not to point out what can be lost by implementing certain policy, but to make sure all interested parties had a chance to contribute to decision making process. Therefore, the outcome is of no concern to proceduralist, her only preoccupation is with just procedure. However, that is also its biggest weakness: in policy-making process we do have to make a decision in the end, even if that decision is to stick to *status quo*. Proceduralist doesn't have an answer to a question: which policy should we choose? This model limits the impact of political philosophy only to ensuring that proper procedures have been followed, but it denies philosophers the opportunity to argue for one policy solution over others.

---

[7] J. Wolff, *Ethics and Public Policy* (London: Routledge 2011), p. 167.

V
Wizard-of-Oz Model

In the famous scene in *The Wizard of Oz* Dorothy's dog Toto pulls the curtain and reveals that the wizard is a mere con-man: in this model the political philosopher takes the role of Toto. He sees his mission not in promoting certain principles, or making sure that all sides have their say, but in pulling the curtain on concepts, principles and arguments invoked in public-policy debates and revealing them in full. It's a mission of avoiding, what Wolff calls, the dangers of hidden assumptions. One example of applying this model is when Wolff talks about deterrence theory in the chapter on crime and punishment and revealing the underling logic behind it: it rests on economic (cost-benefit) model of human behavior which might or might not be the best way to understand individual motivation. This model will most readily be embraced by those philosophers who rely on critical theory in their philosophical work. Their involved in policy-making process can be seen as an extension of their theoretical commitment of seeing philosophy, first and foremost, as a great debunker. Different policy solutions most often have hidden agendas and reflect specific interests, prejudices and biased assumptions. Before we make an informed decision which policy to endorse we should take all of these agendas, interests, prejudices and assumptions into account.

It's worth noting that this model works as a two-way road: by revealing what's behind the curtain, philosopher finds out how and if different theories and arguments, when applied to public policy realm, work or not. After all, Oz was a con-man, but he also helped Lion find his courage, Tin-man his heart and Scarecrow his brain. I can't think of a better illustration of this than Wolff's book in which every chapter ends with a short *Lessons for Philosophy* section. So, for example, in the chapter on

scientific experiment on animals, Wolf convincingly shows that dominant philosophical thinking on this topic—treating animal as either having full or no moral concern at all—is misguided. Taking the argument out of the ivory tower and into policy arena can reveal not only that policy decisions rely on hidden assumptions, but also that our philosophical theories do to.

What are some of the advantages and disadvantages of this model? One advantage is that this model gives philosophers role where then can really shine: putting doubt in presumptions that are considered common-sense or intuitively true has always been philosopher's strong suit. Economist, sociologist and political scientists might be much better on collecting and crunching the data, but philosophers are usually better in giving us a broader picture that goes beyond sheer data. Policy decisions are often made without reflecting on some of the assumptions behind these decisions. Philosopher's role is to make us reflect more on what certain values and arguments that are taken for granted mean. How about disadvantages? Revealing the hidden assumptions behind certain policy solutions might not be the most popular, or, for that matter, the most effective way in tackling policy challenges, so the philosophers that advocate the Wizard-of-Oz model might not be most welcomed to contribute to policy-proposals. Also, policy-makers might not be too interested if philosopher involved in policy-making process has new insights for his fellow academics. They are interested in policy, not philosophy.

VI
Bullshit Model

Bullshit model takes one step further from Wizard-of-Oz model: if there are no knock-down arguments, no great truths to

be realized through public policies, if all concepts are inherently contested, then it might make sense for a political philosopher involved in public policy making-process to behave as a modern-day sophist. If we start perceiving all the sides in some public policy debate as nothing more than different elements in the same power game, nothing stops us from becoming a gun-for-hire, offering our philosophical expertise to the highest bidder. Knowing that there every side in the argument can be shown to be true or untrue, right or wrong, just or unjust, reasonable or unreasonable, consistent or inconsistent, all that remains is a power game. Hence the name bullshit model: not carrying if what we are arguing for is true or not is, as Harry Frankfurt points out, the essence of bullshit.[8]

The advantage of this model is mainly personal: policy makers will probably be thrilled to hire a philosopher who is ready to defend with philosophical arguments their preferred position and discredit the position of their political opponents. Therefore, the political philosopher who embraces the bullshit model can expect many phone calls from the government. Of course, it's questionable if advocates of this model should call themselves philosophers at all: sure, they have the whole intellectual arsenal at their disposal, but as Plato pointed out in *Gorgias*, it's the goal we are striving for and not the tools-of-the-trade that make the philosopher. There is a reason why sophists have such a bad reputation. Of course, from sophist's perspective such reputation is undeserved: if we take moral relativist position as seriously as sophists do, bullshit model might be a most sincere way for a philosopher to contribute to policy-making process. If calls for truth or justice are mere masks that power wears then pretending otherwise is not only naïve, but also dangerous.

[8] H. Frankfurt, *On Bullshit* (Princeton: Princeton University Press, 2005).

VI
Conclusion

In conclusion I want to say a few words on the relationship between these six different models. The fact that I referenced Jonathan Wolff's book when describing some of these models, suggests that most of them are not mutually excludable. They work as ideal-models, while it's fair to assume that philosopher involved in making public-policy recommendations will take more then just one of these models into account. After all, that is exactly what Wolff does in his book. Of course, some models— such as a Syracuse and bullshit models—are mutually exclusive because they rest on opposing epistemological and moral positions and, therefore, define the role of the philosophy and its relations to public philosophy in contrary way. On the other hand, it is quite possible for a philosopher to embrace goals of value pluralist model, but also compliments them with proceduralist model: we can talk about what sacrifices choosing a certain policy solution might entail, but also make sure that all interested parties have there say before the final decision is made. Also, Rawlsian could, when taking about position and arguments that should be taken seriously or disregarded, make his position stronger by using some of the Wizard-of-Oz model's insights about assumptions embedded in these positions and arguments.

In the end, our understanding of how these models relate to each other will in large part depend on which model we find the most convincing. For those philosophers - like Leo Strauss - who embrace the Syracuse model the rest of the models I've described will tell a tale of decline, each next model losing a bit more of what true philosopher's calling should be from its sight, reaching the very bottom with a bullshit-sophist model. From the perspective of those that subscribe to bullshit model the story is just the opposite: it's a narrative of philosophers' *hubris*
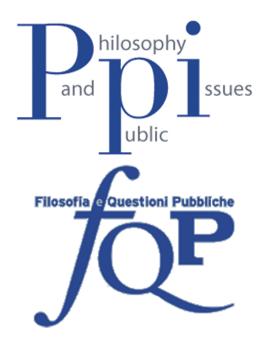
culminating with a smug Syracuse model. For the other four models—all of which show their face in Wolff's book—the story is one of avoiding the extremes: we should avoid boarding that ship to Syracuse, but also keep away from becoming the philosopher's for hire in the bullshit land. The answer to the question we started with—what role should philosopher play in policy-making process—will depend on our understanding of what the proper role of political philosophy should be. It is in the nature of philosophical enquire that there will always be more than one answer to this question.

University of Zagreb

Agreement-Based Practical Justification: A
Comment on Wolff

BY

FABIENNE PETER

[THIS PAGE INTENTIONALLY LEFT BLANK]

Agreement-Based Practical Justification: A Comment on Wolff

Fabienne Peter

I t is tempting to describe Jonathan Wolff's *Ethics and Public Policy* as a book about sex, drugs, and rock n' roll. But it is about much more—gambling, for example, and crime, health and safety, disability, animal experimenting, and the limits of free markets.[1] Even this longer list fails to describe the book accurately. One of its central aims is to further our understanding of the role of moral and political philosophy in addressing public policy concerns. More to the point, it presents a diagnosis of why this role is currently so limited and—under the rubric "lessons for philosophy"—it explores how moral and political philosophy could change to become more relevant for public policy.

In my brief comment on Wolff's wonderfully written book, I want to engage with his diagnosis of the malaise and his lessons for philosophy. As Wolff analyses it so well, the problem is not just created by the fact that philosophers disagree and that disagreements are in tension with the need for a decision—either in favour of a new policy or in favour of the status quo—that characterizes public policy-making. The problem is, rather, that the controversies that are typical in moral and political philosophy, unlike those in the sciences or social sciences, are often sterile in a public policy context. As I interpret Wolff's diagnosis, they are sterile in the sense that they do not give policy-makers reasons to change their beliefs about what public policy—

---

[1] J. Wolff, *Ethics and Public Policy: A Philosophical Inquiry* (New York: Routledge 2011). Unless otherwise specified, parenthetical references are to this text.

existing or new—they should adopt. I agree with Wolff that this is a lamentable state of affairs. One would have hoped that, of all the sub-disciplines of philosophy, moral philosophy and political philosophy would be apt to address public policy questions.

I
Wolff on Public Policy, Philosophy, and Disagreements

What explains the lack of relevance of much moral and political philosophy for public policy debates? According to Wolff, there are three main reasons. One is that public policy debates always start from the present situation. The underlying question is: where do we go from here? By contrast, until recently, at least, contributions in moral and political philosophy have tended to focus on a different question: what is the best or most just state of affairs or institutional arrangement? Wolff follows Amartya Sen who, in *The Idea of Justice*, has famously argued that an answer to the latter question is neither necessary nor sufficient for answering the former. Comparisons between a new policy and the status quo that start from a blueprint for the best or most just state of affairs or institutional arrangement are likely to be mired in second-best problems.

Wolff argues that adopting a "bottom-up," issue-based, approach instead of proceeding "top-down," in theory-driven fashion, could make moral and political philosophy more relevant to public policy-making. The problem, of course, is that an issue-based approach requires careful attention to the features that characterize the status quo and to the details of a new policy proposal. Acquiring the relevant knowledge is time-consuming, but inevitable if philosophers want to play a role in public policy debates. To capture the thought, Wolff turns a famous Marxian

slogan on its head: wanting to change the world is not sufficient; philosophers also ought to try harder to interpret it (191).

Wolff points out, rightly, that the bottom-up, issue-based approach has become more common in moral and political philosophy in recent years. But he adds a well-placed warning to those pursuing this approach. Starting with the issues is not sufficient. It is also necessary to gain an understanding how practitioners themselves think and talk about those issues or one finds oneself once again sidelined in public policy debates.

A second reason has to do with the practical nature of public policy-making. As already mentioned, policy-making needs decisions. Philosophy, by contrast, is a theoretical enterprise, concerned with what we should believe. Seen in this light, it is clear that there is necessarily a gap between philosophy and public policy-making. But philosophical inquiry can, of course, stretch to the subject matter of public policy. There is nothing wrong with asking what we should believe about an existing policy or a new policy proposal. Indeed, this is exactly the kind of philosophical inquiry that Wolff's book so admirably engages in. Wolff asks, for example, what the right way forward is for policies about animal experiments, what to think about the inconsistencies that affect drug policies, or what constitutes the purpose of punishing criminals.

A problem arises, however, insofar as the culture of philosophy is happy to leave our beliefs unsettled. Philosophy, Wolff writes, "thrives on disagreement" (3). Much work in the sciences and social sciences, Wolff argues, aims at advancing a recognized "state of the art"—what everyone should believe. Because this more cooperative research culture produces a body of work that is organized around a core of settled beliefs, it is more compatible with the needs of public policy-making. I shall come back to this issue below.

Beyond this reluctance to settle beliefs and the penchant towards ideal theory, a third main factor that impedes moral and political philosophy's relevance in the public policy context, according to Wolff, is that the search for the correct moral judgment with regard to a particular issue may lead philosophers in the wrong direction. The moral view that will have most traction in the public policy context is not necessarily the correct one, but the one that is most widely shared. As Wolff puts it (196): "If we can achieve intersubjective agreement [on a moral judgment] then we have all we can reasonably hope for."

Wolff explains this need for agreement on the basis of what John Rawls has called the burdens of judgment.[2] The burdens of judgment are a list of factors that play a role in our judgment formation, especially in a moral context. They include items such as the complexity of the evidence about the relevant circumstances, problems that affect the weighing of that evidence, the vagueness of moral concepts, and the difficulty in making hard choices. The joint effect of these burdens of judgment is that even rational deliberation about which policies we should favor may not succeed in settling our beliefs. Wolff argues that because of the burdens of judgment, there is "little prospect of demonstrating that any [moral] view is correct" (5).

As I understand him, Wolff is not just claiming that intersubjective agreement on moral judgments is sufficient in public policy contexts. He is also claiming that it is necessary for the legitimacy of a public policy decision. Even if knowledge of which moral judgment is the correct one were available, that would not be sufficient to legitimize basing a policy choice on that judgment. What is also necessary is that one is able to convince others of the plausibility of that judgment, i.e. to show in a way that acknowledges the burdens of judgment why they

[2] J. Rawls, *Political Liberalism* (New York: Columbia University Press, 1993).

have reason to settle on that particular moral judgment. If that condition is not met, the moral judgment—even if correct—will legitimately be dismissed with the question "who says"? (196)

In sum, Wolff argues that philosophers tend to maneuver themselves out of public policy debates by paying insufficient attention to the factual details of those debates, by leaving disagreements unsettled and failing to cooperate with others in the effort to connect philosophical debates with public policy issues, and, finally, by privileging truth over intersubjective agreement. Of those three, I am—nowadays—particularly interested in the last one and will discuss it in more detail in the next section. Specifically, I want to address the following questions: why does agreement matter and what is the lesson for moral and political philosophy in that?

II
Agreement-based Justification

How can we make sense of the idea that agreement about a moral judgment matters more than the truth of that judgment in the sense that agreement is both necessary and sufficient to legitimize basing a policy choice on that judgment? A first way we might try to answer this question focuses on the motivational situation. Put in terms of a simple rational choice model, the thought is that people are comfortable with their beliefs and will not be motivated to change them unless there is an obvious advantage to doing so. Accommodating a judgment that clashes with a belief that one holds may require making adjustments elsewhere in one's system of belief. It is thus potentially costly. And so, even if truth is the main aim of belief, given the costs of making adjustments, the benefit of doing so needs to be big enough for people to me motivated to accommodate a clashing

judgment. Motivational inertia may thus block the influence even of true moral judgments.

Given motivational inertia, it is clear that appeals to truth are not sufficient for people to make changes in their belief systems. Vice versa, if there is enough of a benefit for enough people in accommodating a particular moral judgment, then the motivational barrier is broken and the judgment can take hold. The motivational strategy thus offers one answer to the question as to why agreement, not truth, is necessary to legitimize basing a policy choice on that judgment.

The problem with this way of answering the question is that it only works for a descriptive concept of legitimacy. The descriptive concept treats the beliefs people happen to hold as the relevant baseline for the justification of policy choices. A normative concept, by contrast, is based on the beliefs that people should hold about policy choices. Such beliefs may, for example, concern the normative reasons that favour one policy choice over another. Of course, as John Stuart Mill has argued, a growing convergence in people's actual beliefs about certain judgments may be taken as evidence for the truth of those judgments. But the point remains that only the normative concept of legitimacy can discriminate between the beliefs people happen to hold and the beliefs they should hold.

I do not think Wolff has the descriptive concept in mind, so we can disregard that way of answering the question why agreement about moral judgments matters for policy choices. Instead, Wolff is drawn to the Rawlsian political liberal view, which assumes a normative concept of legitimacy.[3] On this view,

---

[3] For a fuller discussion of the normative and the descriptive concepts of legitimacy, see F. Peter, "Political Legitimacy," *The Stanford Encyclopedia of Philosophy* (2010 Edition), Edward N. Zalta (ed.),
URL = <http://plato.stanford.edu/archives/win2014/entries/legitimacy/>.

because people tend to disagree about moral truths, appeal to truth should be replaced by another standard of justification: intersubjective agreement on the validity of a moral judgment.

There is, however, an important objection to this Rawlsian view. The objection is that it remains unclear as to why we should accord such normative weight to intersubjective agreement. The objection can be put in the form of a dilemma. Either it is true that agreement has that moral weight, at least in some circumstances, in which case this is a moral truth we cannot bracket, or we do insist on bracketing all moral truths, in which case it is not clear why agreement matters at all.

What should someone who is drawn to the Rawlsian view say in response to this objection? Based on my reading of Wolff's book, it seems that developing such a response is important to support the approach he advocates. But doing so requires that one look beyond Rawls' own work.

Many post-Rawlsian moral and political philosophers are drawn to a moral strategy for supporting the need for agreement in legitimizing policy choices. The most fully developed proposals in this regard are those of Charles Larmore and Gerald Gaus.

Larmore argues that Rawls paid insufficient attention to the problem of how to support the agreement-based justification of moral judgments, and the principles on which they are based, in the political realm.[4] Specifically, he identifies an ambiguity in Rawls' claim that such judgments and principles cannot "be understood as 'moral requirements externally imposed.'"[5] Larmore explains:

---

[4] C. Larmore, *The Autonomy of Morality* (Cambridge: Cambridge University Press 2008), p. 149f.

[5] Ibid., p. 151, quoting J. Rawls, *Political Liberalism*, p. 98.

> This phrase might mean more than one thing […]. If it means that basic political principles are not to be imposed upon a citizen by some external agency—by an enlightened monarch, for example—then I agree. But if instead it means that citizens should not regard their political principles as drawing upon moral requirements whose validity is external to their collective will, then I believe that Rawls went wrong.[6]

Larmore argues that what supports the kind of agreement that a political liberal theory of justification calls for is a moral principle of equal respect for persons. It is this principle that explains why searching for reasonable agreement in the context of the justification of public policy choices is important in the first place. Without such a moral foundation, political liberalism collapses.

Larmore's proposal embraces the first horn of the dilemma I described. Truth sets the standard of justification for the principle of equal respect, but agreement is the standard of justification for further moral judgments, at least in the political realm. Applied to the justification of public policy choices, however, this strategy is vulnerable to the following rejoinder.[7] If there are true moral judgments that do not depend on agreement, why assume that the principle of equal respect is the only such moral truth that is admissible? There might be other moral truths that are relevant. Further argument would be needed to show that in those other cases, truth cannot trump agreement.

Gaus' proposal seeks to avoid this problem.[8] Like Larmore, Gaus also defends the need for agreement based on a premise of

---

[6] Ibid., p. 151f.

[7] Larmore's focus is on the justification of coercion (see 2008: 147f), not on public policy choices in general, and in this context his proposal may not be vulnerable to this objection.

[8] G. Gaus, *The Order of Public Reason* (Cambridge: Cambridge University Press 2011).

persons as free and equals. But he interprets respecting persons as free and equals as a procedural constraint on justification that is grounded in moral practice, not in a moral principle. Gaus argues that everyday moral practice reveals that persons have both the ability and the willingness to adopt rules or principles that go against immediate self-interest. Acknowledging this practice commits us to seeing people as equals in the sense that each has the capacity to adopt behavioural rules or principles and as free in the sense that no one has moral authority over others—such rules or principles are voluntarily adopted, not externally imposed.

A key difference between Larmore's and Gaus' proposals is that Larmore offers a positive argument for agreement-based justification while Gaus offers a negative argument. Larmore resorts to a moral principle to support the equal authority of persons, at least in the political realm. Gaus, by contrast, aims to show how absence of moral authority supports agreement. On his view, an attempt to justify public policies that is not agreement-based would violate the premise of equal freedom; it would falsely suggest that there is an agreement-independent moral authority that could be invoked to justify such choices. Only agreement-based justification respects the equality and freedom on which our moral practice is based.

I have much sympathy for this negative approach to defending the need for agreement. By not relying on a substantive moral principle, it avoids the first horn of the dilemma I described above. I also think that it is on the right track in response to the second horn: agreement matters, not because there is a true moral principle that supports it, but because there is no agreement-independent authority to turn to when it comes to the justification of policy choices.

My worry about Gaus' proposal is that it is incomplete, for the following reason. The proposal rests on a sharp divide between

what he calls the sphere of social morality and the sphere of personal values.[9] The kind of agreement that is relevant for the legitimacy of public policy choices relates to the rules or principles that people put forward for the sphere of social morality in support of moral judgments. What strikes me as implausible about this proposal is the assumption that people's preferences for different rules or principles of social morality are substantively independent of their personal ends. Most likely, if I am a person who, as a matter of personal ends, values generosity, this will influence my preferences over alternative rules of social morality. If I am a person whose personal ends give an important role to security, that is likely to influence my preferences over those rules too.

If this is correct, as I believe it is, then appeals to personal values, and the moral truths on which people take them to be based, cannot be bracketed wholesale: the question which moral judgment is warranted is back. But if moral truths cannot be bracketed wholesale, then the proceduralist strategy to defend the importance of agreement crumbles. Substantive considerations are back in the picture and so is the question as to which moral truths appropriately affect the justification of public policy choices.

III
The Epistemic Circumstances of Agreement-based Practical Justification

The moral strategy for defending the need for agreement thus runs into difficulties. But we should not conclude from this that the case for agreement-based justification is lost. There is an

---

[9] Ibid., p. 2ff.

alternative strategy open to those sympathetic to the political liberal project. This strategy is, like that of Gaus, a negative one. But instead of explaining the requirement for some form of agreement on the basis of the absence of *moral* authority, it explains it on the basis of the absence of *epistemic* authority about which moral judgments are warranted.

While I lack the space here to develop the proposal fully, let me briefly explain what I have in mind. The main idea is this: agreement-based justification becomes appropriate when truth-based justification encounters epistemic obstacles; it is for reasons having to do with the epistemology of practical reasoning that the truth of a moral judgment is neither necessary nor sufficient for practical justification.

If we grant that moral truths can ground moral judgments, the legitimacy of basing a policy choice on such judgments depends on the reach of the epistemic authority of those making those claims. A true moral judgment is necessarily someone's judgment. But, for burdens-of-judgment type reasons, for example, we often disagree about moral judgments. If there is disagreement, the question arises whether it is a weighty disagreement, one that cannot be resolved by rational deliberation, or whether it is a mere disagreement, which would dissolve if everyone took on board the epistemic reasons that apply to them. The reach of epistemic authority settles this question. If the epistemic authority of those making claims about which moral judgments are true can be established—if there is moral expertise—then there are epistemic reasons for those who hold incompatible beliefs to take those claims on board and revise their judgments. If the epistemic authority of those making claims about the truth of certain moral judgments cannot be established, however, then such claims, even if true, do not give others epistemic reason to revise their

judgments. This is Wolff's "who says" worry that I mentioned earlier.

One of the many things that Rawls saw clearly is that moral knowledge that is sufficiently robust to function as a guide in contexts such as policy-making is difficult to achieve.[10] Even if we grant that moral expertise is possible, it is easy to think of scenarios in which moral expertise runs out. Wolff's book has many examples. To mention just one of them, Wolff offers an insightful discussion of the difficulties in establishing expertise in moral judgments about disability. What is the relevant basis of expertise in this case? Knowledge of theories and measurements of well-being? The experience of the disabled? Generally speaking, epistemic authority can be undermined in two main ways. First, the content of the claims themselves may be validly contested. If each side to a disagreement has some justification for the claims they make about which moral judgment is true, but there is no standpoint from which it can be established as to which side is right, then the disagreement is a weighty one. With such disagreements, the epistemic authority of both sides to a disagreement is undermined. Second, the claim to epistemic authority itself may be validly contested. In this case, the problem is not so much that there are justifications for competing claims. The problem is, rather, that the right to make claims about which moral judgment is true is disputed. David Estlund calls this the problem of invidious comparisons.[11]

What happens if appeal to truth is not sufficient to settle people's judgments? In the purely theoretical case, the answer is:

[10] Rawls' list of the burdens of judgment is a good starting-point for making sense of the sources of uncertainty that affect practical deliberation, although a more fully developed epistemology of practical reasoning is necessary to fill out the gaps in Ralws' account.
[11] D. Estlund, *Democratic Authority* (Princeton: Princeton University Press 2008).

not much. If we are only concerned with the question whether a particular judgment—moral or non-moral—is true and the epistemic reasons that apply underdetermine the answer, then justification is suspended until the circumstances change. This might explain the tendency in philosophy to fail to resolve disagreements that Wolff highlights—in the theoretical case, there is no pressure to settle disagreements.

In the practical context, however, things are different. In the practical case in general, it is often impossible not to act. Even deciding to stick with the status quo is an action. In the kind of practical case that is typical for public policy-making, some decision will often be required, whether it is in favour of a new policy, or in favour of the status quo, or in favour of some intermediate strategy, e.g. commissioning further research. In other words, in the practical context in general, one's judgments can often not be left unsettled. And if the relevant practical case is one of collective action, as it is in the policy-making context, there is pressure to settle not just one's own judgments but to find a way of settling the judgments of different people.

I think we can make sense of this situation by distinguishing between two functions of practical reasoning. The first function is shared with theoretical reasoning: I call it the orientation function. That function seeks to establish what we have reason to believe or to do in light of our evidence about the facts. To illustrate what I have in mind with one of the examples that Wolff discusses, if there is new evidence that clearly shows that some drugs are less harmful than previously thought and than drugs that are legally available, this gives us a reason to rethink the relevant drug policies.

But practical reasoning has a second function; I call it the settling function. The settling function helps us decide on what to do when we do not know enough about the relevant facts to

form a judgment on that basis. The settling function can also be illustrated on the basis of Wolff's discussion of drug policies. One explanation that Wolff offers for the status quo bias in policy-making, including policy-making about drugs, has to do with the uncertainty surrounding changes to existing policies. Change is thus harder to justify than the *status quo* (79). As a result, Wolff argues, even inconsistencies in existing policies do not necessarily give us reason to change those policies—the reasons given by the facts are not the only reasons that legitimately influence policy-making. As Wolff puts it:

> if we find out—as we seem to have—that the treatment of ecstasy and alcohol is inconsistent, then so what? You can still follow the law. Ideally, of course, there would be no inconsistencies, but may laws are compromises between competing interests, and different laws were made by different people, at different times, for different purposes (82).

The need to settle people's judgments about how to proceed can support agreement-based justification. The thought can be explained by tweaking Gaus' proposal a bit. Recall that Gaus argues that agreement-based justification applies when there is no appeal to moral authority. My alternative proposal is that agreement-based justification applies when there is no epistemic authority about which moral judgments are true yet there is a need to settle people's judgments. We should not think that we are in a normative wasteland if we cannot settle our beliefs about which moral judgments are true. Instead, if truth-based justification runs out, some other standard comes into play, a standard that regulates the settling function of practical reasoning. And agreement-based justification, in some form, is the appropriate standard when there is no hierarchy of epistemic authority.

The epistemic strategy I have sketched here reveals that the dilemma that Rawls' critics put forward against the political liberal
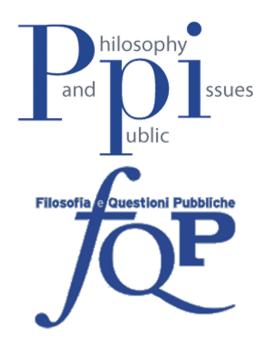
theory of justification is misleading. The dilemma is based on the presumption that a moral argument is necessary to defend agreement-based justification. What supports agreement-based justification, however, is not the moral appeal of agreement, but the epistemic limitations that truth-based justification runs into, at least in practical contexts such as policy-making.

The University of Warwick

If you need to cite this article, please use the following format:

Peter, Fabienne, "Agreement-Based Practical Justification: A Comment on Wolff,"
Philosophy and Public Issues (New Series), Vol. 4, No. 3 (2014), 37-51, edited by S.
Maffettone, G. Pellegrino and M. Bocchiola

Punishment and the Rebalancing of Status

BY

GERALD LANG

Punishment
and the Rebalancing of Status

Gerald Lang

Jonathan Wolff's *Ethics and Public Policy: A Philosophical Inquiry* offers both an insightful and lucid primer on several problems in applied ethics and political philosophy—the regulation of drugs and gambling, safety standards, the distribution of health-care, the treatment of animals, the proper operation of the free market, and so on—as well as the salutary reminder that there is more than just philosophical theorizing which must be taken into consideration if philosophical arguments, and the philosophers who advance them, are going to stand any chance of influencing public policy.[1]

The division between pure theory and troublesome practice actually appears to play a smaller role, however, in Chapter 5, where Wolff considers crime and punishment. As ever in moral theorizing about punishment, problems as to the effectiveness of criminal sentencing are acknowledged, particularly in respect of 'forward-looking' theories of punishment (118-24). Along this particular dimension, the main questions Wolff pursues are these. How and when does punishment deter? How effectively does it deter? What assumptions must we make about the motivational profile of those who are likely to be deterred by the prospect of punishment, and can these assumptions be squared with what we take to be the proportionality conditions on punishment? Perhaps theorists of punishment have sometimes been content to

---

[1] Jonathan Wolff, *Ethics and Public Policy: A Philosophical Inquiry* (London: Routledge, 2011). In-text page references will be to this book.

speculate about such matters from the armchair, rather than through the consultation of empirical data. Be that as it may, the significance of empirical questions about the sources of motivation and the effectiveness of punishment as a deterrent were always staring such theorists in the face.

Turn now to 'backward-looking' theories of punishment, to the ranks of which Wolff adds his own theory. (The appraisal of this theory will occupy the bulk of my attention in this article.) How can such a backward-looking account of punishment avoid the charge of being nasty or obnoxious? This is also a wholly familiar problem facing backward-looking theorists.

In sum, the questions about punishment which Wolff raises and grapples with are these largely familiar ones. What deserves a much closer look is Wolff's way of tackling those questions. As already indicated, I will be dealing mainly with the details of Wolff's backward-looking account.

This article will unfold as follows. Section I will outline Wolff's backward-looking account of punishment. Some critical questions for this account are raised in section II. The remaining part of the article will then attempt to show that Wolff is nonetheless indirectly getting at something that deserves to be taken seriously. In section III, I suggest that the prospects for an account broadly sympathetic to Wolff's aims will be brighter if we pay greater attention to the basis of an offender's liability and the connections between self-defence and punishment. Some brief conclusions are drawn in section IV.

I
Wolff's Rebalancing of Status Account

Wolff offers a striking and ambitious theory of punishment,

blending elements of a 'communicative' theory (117, 125) and retributivism, and I will now quote from it at length.[2] It begins to be outlined in this passage:

> [W]hat is so bad about crime, or at least some crimes for some people, is the fact of being made a victim. It is not so much that others attempt to treat you with contempt, but rather that they manage to do so. This is why… there is such a psychological difference between failed attempts and successful crimes. In succeeding in their crime against you, perhaps, they implicitly announce themselves as in some respect your superior. They have victimized you, and left you in a lower status. Even when there is no identifiable victim—as in the case of vandalism of public property—the successful criminal implies that in some sense he or she is above the norm, or, at least, above the rules. Crime communicates a message (124-5).

Wolff continues, in the paragraph below:

> If [that is] so, then punishment appears in a new light. For at least part of the purpose of punishment then becomes to re-establish some sort of proper status between all the parties. If a criminal is caught and adequately punished he has no longer got away with something. He can no longer implicitly claim to be of higher status, and those who were victims may feel that their victim is expunged, and they have their previous status restored to them (125).

As for crimes without living identifiable victims, either because the crimes were so-called 'victimless' crimes or because they were homicidal:

> The analysis […] still applies in modified form. In the standard case, where there is an identifiable surviving victim, punishment 'rebalances' status by raising the standard of the victim and lowering that of the perpetrator. In cases where the victim is dead, nevertheless the punishment can still show that we as a society still take that life very seriously… In cases of victimless

---

[2] For a detailed example of a communicative theory, see, in particular, Anthony Duff, *Punishment, Communication, and Community* (Oxford: Oxford University Press, 2001).

crime, all that can be done is lower the standing of the perpetrator, but this remains a significant matter. The echo of retributivism is, presumably, the greater the crime, the more is needed to restore the moral balance. (125)

As we can see, then, Wolff's account is focused on the status of the offender and the victim: in a criminal act, the offender's status is heightened, and the victim's status is lowered. Punishment then rebalances the offender's objectionably heightened status, and the victim's objectionably lowered status, so that there is, once more, parity of status between them. I will call Wolff's account the *Rebalancing of Status Account*, or the *Rebalancing Account* for short.

II
Problems with the Rebalancing Account

By Wolff's own admission, the Rebalancing Account is 'speculative rather than research based' (127). But that feature seems unavoidable in any serious backward-looking account of punishment, and is nothing to apologize for, at least as I see it. Punishment involves the intentional infliction of suffering, or at least hard treatment, on offenders, and that feature raises profound moral problems that cannot be settled simply by empirical research.[3] The basic problem here is that, even in the unconstrained speculative space in which philosophers feel most at home, there is nothing approaching a secure consensus among them as to why it is morally permissible to inflict such hard treatment on offenders. So Wolff's argument for the Rebalancing Account will not be taken to task for that reason. That leaves us, however, with a few other problems.

[3] I take the useful phrase 'hard treatment' from T. M. Scanlon, 'Giving Desert Its Due', *Philosophical Explorations* 16 (2013): 101-16.

In what follows, assume that X is the offender, and that Y is the victim. Consider the following claims:

(1) X criminally offends against Y at t_1.

(2) It is appropriate to punish X, at t_2, for his criminal offence against Y.

Claim (1) is true by assumption.[4] Claim (2) might strike us, or many of us, as intuitively attractive, but we clearly need an argument for it. In particular, what is the argument for (2) submitted by the Rebalancing Account?

Our initial task is to get a bit clearer about Wolff's notion of 'status'. The preferred interpretation of Wolffian status must be answerable to several different constraints, or features. First, we need an interpretation of it which accommodates its alterability: status can go up, and it can go down. Call this the *alterability feature*. The alterability feature thus excludes a purely normative understanding of status which simply records an individual's worth as a moral agent and moral patient. (Even if (1) is true, Y's worth surely does not decline, and X's worth surely does not increase.) Second, our interpretation of Wolffian status must provide for the *interconnectedness* between the alterations in X's status and Y's status: *because X's status increases, Y's status declines*. Call this the *interconnectedness feature*. Third, it would be pointless to settle for an interpretation of status which simply *re-described* (2) in slightly different language. We need claims which can genuinely *illuminate* or *explain* (2). Call this the *explanation feature*.

With this slightly fussy stage-setting in mind, we can proceed. The following claims form the substance of the Rebalancing

---

[4] Further assume, to ease possible concerns, that the particular offence X commits against Y is one which any reasonable or sane jurisdiction would criminalize: violent unprovoked assault, for example.

Account, and will surely have a large role to play in how the Rebalancing Account delivers (2):

(A) Before t_1, X and Y have the same status.

(B) After t_1, as a result of his criminal offence against Y, X has a higher status.

(C) After t_1, as a result of her subjection to X's criminal offence against her, Y has a lower status.

(D) After (proportional) punishment is imposed on X, at t_2, X's higher status is lowered to the status X possessed before t_1.

(E) After (proportional) punishment is imposed on X, at t_2, Y's lower status is raised to the status Y possessed before t_1.

Due to the interconnectedness feature, it makes sense to treat these claims primarily in pairs: (B) and (C) need to be jointly considered, as well as (D) and (E). We must also pay attention to the relationships between these pairs of claims.

How are (B) and (C) to be interpreted? It makes sense to isolate features of (B) and (C) which track, or reflect, the fact that X has wronged Y, since that will permit straightforward accommodation of the alterability and interconnectedness features; it will, furthermore, help us to see how (A) can be true; and it will help to pave the way for the explanation feature. But, to provide for the explanation feature, we also need to somehow *go beyond* these facts about wrongdoing, since the story about X's wronging of Y does not, by itself, demonstrate why it is appropriate, permissible, or desirable to punish X for having wronged Y. This is just another way of saying that the truth of (1) does not give us any obvious basis for (2). We are appealing to the Rebalancing Account, after all, to plug the gap between (1) and (2).

Look again at what Wolff says in the passage above: offenders 'implicitly announce themselves as in some respect [the victim's] superior' (125). In victimizing Y, X also implicitly conveys the attitude that Y is fit to be victimized, and that X is fit to victimize her. But even if this victimization story is true—even if it successfully captures X's underlying attitudes to Y[5]—it is not immediately clear what role it can play in justifying punishment. That is, even if the victimization story takes care of (A), (B), and (C), it does not obviously take care of (D) and (E). To put the underlying point more sharply, we do not need punishment to tell us that X's attitudes are morally flawed. But if we do not appeal to our practice of punishing murder and rape to reassure ourselves that murder and rape are wrong, then neither should we need to appeal to the practice of punishment to reassure ourselves that the attitudes which accompanied the original offences are morally flawed. Indeed—a stronger point still—it would be normatively back-to-front to appeal to the legitimacy of punishment to demonstrate that the offences which provoked punishment were wrong. So it seems odd to appeal to the practice of punishment to demonstrate that the attitudes which accompanied those acts were flawed. How could there be such a stark justificatory asymmetry at this point in the story? The prospects for the explanation feature now seem poor.

But perhaps Wolff's deeper point is not that we need to appeal to punishment to vindicate the wrongness of the attitudes that accompanied the offences, but that punishment gives us an appropriate way of *publicly expressing* our conviction that these attitudes were flawed. If these attitudes are morally flawed, then it is not inappropriate to seek some form of expression for our

---

[5] Wolff does not pretend to have demonstrated that this account about the offender's attitudes is necessarily manifested in his treatment of the victim (125).

condemnation. This is an important idea, and I shall be revisiting it in the next section. But it is not yet clear how the Rebalancing Account establishes that *punishment*—the imposition of hard treatment—is the proper expressive outlet. And it is far from clear how punishment can actually *rebalance* status. *Y* has been victimized, and this is morally objectionable; but how is *Y*'s status as a victim corrected, annulled, or cancelled out by *X*'s punishment? This is a significant lacuna for the Rebalancing Account. Wolff's story about rebalancing status is supposedly meant to tell us *why* punishment is the proper vehicle of that rebalancing; but it is none too clear how punishment achieves this rebalancing, because it is none too clear what, exactly, is being rebalanced.

We can concede to Wolff that a societal insult would be conveyed to *Y* if we failed to punish *X* for his offence against her, and yet continued to punish other tokens of that offence type against other victims. Here it is the *selectivity* of punishment which would indicate that we were taking *Y*'s victimhood less seriously than the victimhood of other individuals. In this connection, Wolff gives the example of victims from unpopular racial minorities, crimes against whom receive scant investigation or publicity (125). But that consideration alone will not serve to justify punishment *across the board*.

One further comment on the dialectic as it has unfolded so far. Some might think that *X*'s attitudes do not matter as much as *Y*'s attitudes. *Y* is the victim, and *Y* may have a very strong preference that *X* be punished for his offence against her. But that is not the story offered by Wolff, and for good reason. Unless we are committed, on independent grounds, to an industrial-strength version of an unrefined preference-satisfaction theory, it is implausible to hold that victims' preferences that

offenders be punished can justify punishment. We surely need a further account of *why* those preferences should be heeded.

Finally, a further problem concerns Wolff's treatment of the difference between attempted crimes and successful crimes. The challenge here is that a serious attempt at a criminal offence arguably reveals just as much contempt towards the would-be victim as a successfully completed crime. True, successful crime produces the injury to add to the insult, but Wolff's Rebalancing Account would seem, in any case, to be more concerned with the insult than with the injury.

Wolff happily admits that '[c]ontempt is shown even by an attempted crime' (115), but that '[a] successful crime cuts deeper' (115). He says:

> In both cases one has been victimized, but there is an important difference. In the failed attempt, one may be rather shaken but there may also be a rather triumphal feeling 'I got the better of him!' When the crime is successful, there is no such comfort (115).

Furthermore:

> [W]hen the attempt is successful perhaps one begins to harbour the thought that the contempt is deserved. If I am unable to protect myself, what sort of person am I? A successful crime seems, in at least some cases, to bring about a change in status and in self-respect (116).

These passages give expression to a variety of different ideas. Now it would be foolish to deny that the injury produced by successful attempts, combined with the contempt they express, can severely affect a victim's sense of self-respect and self-worth. This is truly a poisonous cocktail for victims to deal with. But there is also a hint of strain in the suggestion that victims of successful attempts will harbour 'the thought that the contempt is deserved'. If a victim does not deserve to be attacked in the first

place, it would be strange to think that the difference between an offender's failed attempt and his successful attempt could nonetheless reflect the victim's desert, and it would be surprising if victims standardly came to that distorted view of themselves. The difference between failed attempts and successful attempts is an important one, as we will see in the next section, but it seems to me that Wolff does not quite manage to exploit it in the right way.

So, to sum up. Wolff writes:

> If crime sends a message, then so does punishment, standardly an attempt to send a counter-message, cancelling out the first message. (125-6)

Even if there something attractive about this claim, there are at least two major problems for the Rebalancing Account. First, it is not clear why a 'counter-message' needs to be sent out at all; and second, it is not clear how that counter-message can 'cancel out' the first message issued by the offender.

III
The Extended Liability Account

I have suggested that there is no obvious interpretation of status in the Rebalancing Account that can serve Wolff's purposes. Even if we are strongly persuaded by the claim that it is appropriate to punish X for what he did to Y, we lack a convincing status-based explanation of that claim: we cannot plausibly secure (2) by saying that punishment reduces X's objectionably heightened status and raises objectionably lowered Y's status. We should therefore de-emphasize the notion of status in any putative justification of punishment.

What should we appeal to instead? The account I propose will retain Wolff's investment in the story about the communicative significance of punishment, but jettison his investment in the story about rebalancing status. However, we need to start yet further back. I propose that we start with an account that blends backward-looking and forward-looking considerations: namely, an account which seeks ultimate grounding for the permissibility of punishment in the permissibility of self-defence.[6] More specifically, I suggest that we begin with X's liability to self-defensive action. Now there are many ways of grounding defensive permissions and attacker's liability. I will not be concerned here with these intramural debates among self-defence theorists; my aim is simply to argue that, however the attacker's liability is to be grounded, it is implausible to hold that his liability evaporates altogether when his offence is completed.

Imagine, in some very nearby possible world to the world that was described by (1) and (2), that X is in the course of criminally offending against Y. (To be more specific, imagine that X is about to assault Y.) On the basis of X's attack on Y, we seem able to say the following:

(3) Y is permitted to take necessary and proportionate violent self-defensive action against X.

And, as an extension of (3), we can also say:

[6] See, for example, Warren Quinn, 'The Right to Threaten and the Right to Punish', *Philosophy & Public Affairs* 14 (1985): 327-73, and Daniel Farrell, 'The Justification of Deterrent Violence', *Ethics* 100 (1990): 301-17. David Boonin, *The Problem of Punishment* (Cambridge: Cambridge University Press, 2008), 192-211, provides a detailed critique of this 'self-defence solution' to punishment. I lack the room here to engage with the full details of Boonin's critique; I can only set the ball rolling in the right direction, by arguing that X's liability to defensive attack establishes a presumption that X is also liable to punishment.

(4) *Z* (a third party) is permitted to take necessary and proportionate other-defensive action on behalf of *Y* against *X*.

And, as a way of explaining how these permissions arise, we can say:

(5) In virtue of his attack on *Y*, *X* is liable to be attacked in self-defence (by *Y*) or in other-defence (by *Z*), subject to the necessity and proportionality conditions on self-defence and other-defence.

Of course, the world that was described by (1) and (2)—the 'punishment world'—is not the 'defensive world' jointly characterized by (3), (4) and (5). In the punishment world, the imminent threat to *Y* has been and gone. It is therefore too late to apply self-defensive or other-defensive violent action to avert *X*'s attack on *Y*.[7] The point is, by now, a fairly familiar one in the literature, but it bears blunt restatement:

> When Moe harms Larry in self-defense, he harms Larry in order to *prevent* Larry from wrongfully harming him. But when the state punishes an offender, it punishes him precisely because he has already *succeeded* in wrongfully harming someone. It is easy to see how the notion of self-defense can justify harm to prevent a particular wrong from taking place. But it is far more difficult to see how an appeal to self-defense could be used to justify inflicting harm in response to a particular wrong when it is already too late to prevent that wrong from taking place and thus too late to provide a defense against it.[8]

Can the moral content of (5) nonetheless cast useful light on the permissibility of punishment? I believe it can.

Imagine that *X*'s attack on *Y* begins at t_0, and is successfully completed at t_1, as (1) stipulates. Between t_0 and t_1, *X* is liable to

---

[7] To save words from now on, the phrase 'self-defensive or other-defensive' will be compressed into, simply, 'defensive'.

[8] Boonin, *The Problem of Punishment*, 193-4; original emphases.

defensive attack. From t_1 onwards, X can be no longer be liable to defensive attack, for the simple reason that the conditions applying to defensive attack no longer obtain. But it is intuitively odd to think that X's moral standing is no longer impaired, just because X has now succeeded in an attempt which he was morally forbidden from making in the first place.[9] Plausibly, X's overall moral standing is even worse after t_1, since the successful completion of his criminal attempt to F is morally worse than his mere attempt to F. That is:

(6) X's successfully completed offence against Y is a morally worse action than X's mere attempt to commit an offence against Y.

But what should we conclude from (6)? Even if it is true of X, after t_1, that X was liable to defensive violence between t_0 and t_1, what further relevance does that fact have? On one view, that fact has no further relevance: X's liability to hard treatment is exhausted by his liability to defensive attack, which obtains only in the interval between t_0 and t_1. After t_1, in the world in which X's attempt is successfully completed, X is simply not liable to hard treatment. Call this the *No Further Implications View*.

The No Further Implications View strikes me as intuitively implausible. But it will not do, in this context, to rest the argument squarely on confidence in that verdict. Friends of the No Further Implications View may reply that, even though X is not liable to any hard treatment upon expiry of the conditions for defensive attack, X's moral standing is still impaired, *simply because*

---

[9] See Gerald Lang, "Why Not Forfeiture?," in *How We Fight: Ethics in War* (Oxford: Oxford University Press, 2014), ed. H. Frowe and G. Lang, at 43. My particular concern, in that discussion, was with the prospects for a forfeiture account of self-defence; now I am trying, on the assumption that there is some or other secure basis for the liability of an individual to defensive attack, to enlarge the basic story to encompass the individual's liability to punishment.

(6) is true. If (6) is true, then *X* can still be *blamed* for what he did, even if he cannot, on that basis, be *punished* for what he did.

I am not persuaded by this counter-response, for reasons that are partly connected with something that is laboured by Wolff: the communicative significance of punishment, or (to put it more neutrally) the communicative significance of having *some* effective form of social condemnation of criminal wrongdoing. On the No Further Implications View, *X* can be blamed, but not punished. However 'hot' blame can be—however ardently or forcefully it can be conveyed—it cannot take the same form as the hard treatment (in this case, defensive violence) that would be appropriate had *X* not successfully completed his attempt to wrong *Y*. But then it will seem that, by comparison with the sanctions available for dealing with the normative situation between t_0 and t_1, *X* is actually being rewarded for the successful completion of his criminal attempt against *Y*. If this 'sends out a message', then the message it sends out is that *X*'s position is improved if he does something that is morally worse.

To recapitulate, there are at least three problems to worry about if the No Further Implications View is taken to be true. First, public condemnation of *X*'s offence will be largely drained of force, conviction, or efficacy. The condemnation will seem relatively toothless. Second, the relative toothlessness of the public condemnation of *X*'s offence may contribute to, and help to sustain, *Y*'s feelings of psychological and social fragility. Third, this feature also seems apt to generate a perverse personal incentive for *X* to do something that is morally worse.

To escape these implications, I propose that we embrace the following:

(7) The facts which make *X* liable to defensive attack between t_0 and t_1 also make *X* liable to punishment after t_1.

We can refer to the account I have sketched, in admittedly embryonic form, as the *Extended Liability Account*. The Extended Liability Account plainly differs from the Rebalancing Account. But, as I have already hinted, I think there may be underlying similarities between them. I try to identify what they are in the concluding section.

IV
Conclusion: Rebalancing the Rebalancing Account

I have argued that Wolff's Rebalancing Account fails to deliver what he is after: namely, a secure backward-looking account of punishment. My view is that, if an offender is liable to the hard treatment constituted by defensive force, then he is also liable to the hard treatment constituted by punishment. As I see it, punishment does not rebalance the respective statuses of the offender and the victim. That is an essential element of Wolff's story, and it is no real part of mine.

However, the Extended Liability Account also leans, at certain points, on similar intuitions as Wolff's about the communicative or expressive significance of punishment; and it aims to respect and uphold the victim's status by helping to confirm that the offender's completed offence leaves him in a worse moral and personal position than he was in when he had not yet completed that offence. As different in crucial aspects as they may be, then, the Rebalancing Account and the Extended Liability Account still draw water from the same well.[10]
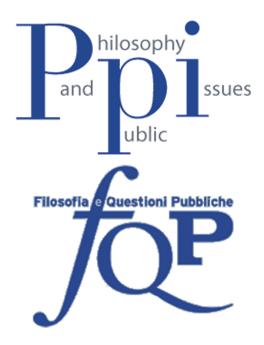

University of Leeds

Meta-Ethics in Debates on Public Safety:
A Critical Appraisal of Jonathan Wolff's
Bottom-Up Approach to Applied Ethics

BY

STEVEN R. SMITH

Meta-Ethics in Debates on Public Safety
A Critical Appraisal of Jonathan Wolff's Bottom-Up Approach to Applied Ethics

Steven R. Smith

J onathan Wolff's book *Ethics and Public Policy: A Philosophical Enquiry* is a welcome addition to a growing literature exploring the relationship between ethics and public policy-making.[1] The book's strengths are three-fold. First, EPP benefits from the lucidity of Professor Wolff's philosophical writing, while rarely losing the profundity and nuance of the issues at stake. Second, drawing from Professor Wolff's experience of contributing to Government advisory committees, EPP is replete with detailed accounts of public debates giving a refreshing richness to the philosophical arguments made. Third, a very good case is made for what Professor Wolff calls, a bottom-up approach to applied ethics. He demonstrates, with skill and authority, the importance of philosophers engaging with the specifics of public debate, including a proper consideration of well-researched empirical studies relevant to the policy area under review—rather than relying on philosophers' half-baked 'common-sense' assumptions which have frequently got philosophers into embarrassing holes, when positions are defended for this or that principle or ethical theory.

Subsequently, my brief comments here should be read with these observations in mind. I have considerable sympathy with EPP's general direction and approach, and acknowledge that

[1] J. Wolff, *Ethics and Public Policy: A Philosophical Enquiry* (London: Routledge, 2011), hereafter referred to as E

Professor Wolff has broken new ground applying ethics and philosophy to public policy analysis and debate. It is also important to recognise that the book was written, in part, for a non-philosophical readership, and so if his target-audience was only philosophers then my concerns below may well have been addressed. Nevertheless, taking EPP at face-value, I focus on its lack of attention to meta-ethics and questions concerning the nature of value pluralism. Certainly, philosophy has important things to say about meta-ethical issues and how different versions of value pluralism affect decision-making. I have argued elsewhere that these considerations are highly pertinent for policy-making, as public debate is often conducted via implicit non-monist interpretations of value commitment.[2] My contention here is that EPP could have engaged with these issues more explicitly, shedding greater light on the dilemmas and conflicts in public debate and the bottom-up approach to applied ethics. Throughout, I will refer to the chapter on 'safety' in EPP (83-108), as this chapter illustrates well how the book may have benefitted from this kind of meta-ethical analysis.

II

What does Professor Wolff mean by a bottom-up approach to applied ethics? In the absence of a chapter explaining this methodology we can elicit his understanding of this approach by

---

[2] For example, see Steven R. Smith, *Equality and Diversity: Value Incommensurability and the Politics of Recognition* (Bristol: Policy Press, 2011); Steven R. Smith, "Applying Theory to Policy and Practice: Methodological Problems and Issues," In S. R. Smith (ed.), *Applying Theory to Policy and Practice: Issues for Critical Reflection* (Aldershot: Ashgate Publications, 2007), 1-18; Steven R. Smith, *The Centre-Left and New Right Divide: Political Philosophy and Aspects of UK Social Policy in the Era of the Welfare State* (Aldershot: Ashgate Publications, 1998), 214-262.

piecing together a range of quotes from EPP, drawing out four general rules underpinning this methodology:

> Ethics is, in my view, more like the science of medicine than physics. Theoretical and technical knowledge is important, but there is no reason in advance to think that it can all be stated in terms of clear, simple principles, or made to fit into a single, complete coherent framework, or that every problem can be solved with a complex algorithm (6).

> It is becoming increasingly common to distinguish 'top-down' (or theory-driven) from 'bottom-up' (problem-driven) approaches to applied ethics. This book, naturally enough, aims to be a contribution to 'bottom-up' theorizing, where the first task is to try to understand enough about the policy area to be able to comprehend why it generates moral difficulties, and then to connect those difficulties or dilemmas with patterns of philosophical reason or reflection (9).

> The methodology implicitly recommended here suggests that when thinking about a practical issue, we should start at the other end: not at the philosophical theories but current disagreements in the public policy area. We need to ask: what do people think they disagree about? And is that the best way of understanding their disagreement? Is there a better way? ... But to do this one has first to become immersed in the debate in which one wishes to intervene (36).

> It is implied in the methodology I am suggesting that participants in public debates do not always fully comprehend or perfectly articulate what they disagree about. A simple slogan or principle, while helpful for campaigning, can have a distorting effect on argument (36).

> Understanding real-life examples, and working out the basis of the dilemmas they create, is an essential part of the attempt to come to a resolution of the issues (107).

> Accordingly, there is no shortage of philosophers who hope to change the world. But what they sometimes have failed to do is to interpret the world they live in. Often they fail to investigate why it is society does the things it does [...] Of course no one thinks that somehow the world will miraculously conform itself to the intellectual ideal, but philosophers

sometimes fall short of taking up the challenge of thinking hard about questions of the process and, even more importantly, consequences of implementation (191-192).

A philosophical approach to public policy analysis, I believe, will be very likely to go wrong if it starts from the announcement of a set of principles or values claimed to be true. Where should it start then? ... Philosophers arrive on the scene like a (very slow moving) emergency service. There is a problem that needs resolution, if possible. Obviously, therefore, no progress can be made unless the problem is understood (196).

Following the above, four general rules of the bottom-up approach emerge:

(1) *The anti-abstraction rule*: Ethical principles should not be articulated before considering the specific policy issues at stake, as if these principles could, in their abstraction from specific issues, offer complete and watertight solutions to problems and dilemmas in policy-making. Applied ethics, then, is less of an abstract mathematical calculation providing once-and-for-all answers to problems and questions, and more of a trial-and-testing process worked-out in 'the field' of policy-making.

(2) *The full description of policy/public debate uncovering moral dilemmas rule:* The policy area and associated public debate should be described in detail first, anticipating that this description will uncover various moral dilemmas and disagreements. Only after this full description occurs can relevant philosophical questions and arguments be raised and applied to illuminate better these dilemmas and disagreements.

(3) *The full description of empirical realities ensuring accurate interpretation rule:* Part of describing the relevant policy area is not only exploring detailed examples of the moral issues at stake to uncover dilemmas and disagreements as identified in

(2), but also to ensure philosophers are able to accurately interpret the world they are theorising about, as they consider these dilemmas and disagreements. Accurate interpretation requires that philosophers accommodate empirical investigations and studies pertinent to the policy area examined, properly informing the premises used when constructing philosophical argument.

(4) *The over-arching applied philosophical analysis rule:* Only after (1)—(3) is implemented can philosophical analysis be usefully applied to public policy, to expose more clearly disagreements in public debate, and even to provide some solutions to the dilemmas and problems policy-makers encounter.

Certainly, a lot could be said about these rules—their more precise definition, how they relate, whether sub-rules are implied, and so on. However, my job here is not to unravel these wider debates concerning bottom-up theorising. Rather, I will assume their coherence as articulated by Professor Wolff, from which I will raise questions about how these rules are specifically implemented in EPP—most notably concerning those meta-ethical considerations outlined earlier.

III

Much of EPP explores directly moral dilemmas and conflicts as these relate to public debates—covering topical issues in, scientific experiments on animals, gambling, drugs, safety, crime and punishment, health, disability, and the free-market. However, there is surprisingly little meta-ethical discussion in the book about the relationship *between* values as these are applied to policy, and especially when the values committed to are thought to be many, and in conflict.

Certainly, public debate is filled with examples of competing ethical positions carrying weight in political argument, leading to conflicts between these positions, and reflecting ethical and political compromises in decision-making. These compromises are often uncomfortable to live with but, for Professor Wolff, this heightens the importance of engaging in philosophical arguments concerning abstract theorising, and how these arguments are applied to policy-making. For example, the chapter on 'safety' explores the conflict between utilitarian considerations of cost and maximizing happiness for all/the many, versus absolutist considerations of the sanctity of life for one/the few. Professor Wolff concludes that: "In the end the issue seems to come down to this cold-blooded question: is it worth spending that amount of money to save each life?" (95). However, once this question is addressed, it can be seen how the conflict between utilitarianism and absolutism is a feature of practical decision-making, even if philosophers are uncomfortable making decisions in this context:

> Philosophers can retreat, shaking their heads, and refusing to take further part in the discussion. Or they can grapple with the question either of what price we should put on life, or how we can make safety decisions without valuing life. None of these options is likely to be comfortable (108).

Nevertheless, I contend here that not only should difficult questions concerning the conflict between ethical positions be addressed (as illustrated in EPP) we should also examine the conflict *within* ethical systems when values are assumed to be many and conflicting. My main claim is that examining the latter helps our understanding better the character of bottom-up theorising, and how conflict is handled on philosophical grounds when value pluralism is assumed. With risk of over-simplifying, there are three major versions of value pluralism: incommensurability, intuitionism, and lexical orderings.

Incommensurabilists argue that when values conflict, whether these are moral or non-moral values, they are often incomparable.[3] Incomparability implies that one value is no greater or lesser than the other value, but neither are these values equal.[4] For the incommensurabilist, philosophical resolutions cannot be found which 'solve' the conflict between these values, because values cannot be weighed against each other, as if there is an appropriate trade-off which can be justified through philosophical reasoning. This kind of justification would imply that values have been compared in the weighing-up of each value, where X amount of P value is defended as equivalent to Y amount of Q value, and so on. Instead, for the incommensurabilist, trade-offs derive from political conflict, where philosophical reasoning at best provides undefeated reasons for particular trade-offs, but cannot supply reasons which defeat all other alternatives.[5] Certainly, practical difficulties emerge when two options cannot be chosen at the same time, and each alternative is justified for coherent philosophical reasons, but according to the incommensurabilist when a particular option is chosen for one set of reasons, the other set is not invalidated as a result.

What are the implications for public safety and, say, the decision to spend large amounts of money to save relatively few lives (EPP, pp. 87-102)? As previously outlined, according to

---

[3] For one of the earliest and clearest expositions of value incommensurability see Joseph Raz, *The Morality of Freedom* (Oxford: Clarendon Press, 1988), 321-368; also see Ruth Chang (ed.), *Incommensurability, Incompatibility, and Practical Reason* (Cambridge MA: Harvard University Press, 1997).

[4] Joseph Raz, *The Morality of Freedom*, 332.

[5] Also see Joseph Raz, "Incommensurability and Agency," in R. Chang, *Incommensurability, Incompatibility, and Practical Reason*; and see Joseph Raz, *Value, Respect, and Attachment* (Cambridge: Cambridge University Press, 2001), and my arguments in S. R. Smith, *Equality and Diversity*, 31-58.

EPP, arguments for and against this policy are coherent, reflecting absolutist appeals to saving life as much as possible, and utilitarian appeals to maximising welfare for all/the many. But, if either argument holds weight in public debate and political compromises are made, at first blush, there is a philosophical problem, because both cannot be committed to without inconsistency. Subsequently, for Professor Wolff, there is a clear distinction between philosophers' and policy-makers' standards of argument, given this appeal to inconsistency:

> The philosopher's favourite weapon is rather blunted in public policy. This is the appeal to inconsistency ... But in public policy this does not work so well. While one can hope for consistency it would be foolish to expect it, and pointing out inconsistency is not a decisive argument. Here, once more, philosophy and public policy have different standards of argument. (82).

However, my point here is that this distinction can be exaggerated if value pluralism is under-theorised. For example, as just stated, the incommensurabilist tolerates certain forms of inconsistency regarding value commitment, as conflicts can demonstrate value incomparability. Therefore, when two values conflict we need not abandon both *or* either one, to supposedly 'solve' the inconsistency. So, using the incommensurabilist formulation, preserving life at a high cost could be seen as neither better than, worse than, or equal to a utilitarian decision calculating the opposite. It matters that whatever option chosen is for coherent undefeated reasons, but not for reasons which defeats all others. This allows the latter reasons to remain in the 'wings' anticipating that decision-makers might favour alternative options at other times.[6]

---

[6] Also see my arguments in S. R. Smith, *Equality and Diversity*, 31-58.

Intuitionists take a different pluralist turn. Trade-offs between moral values are not simply the result of political 'pushes', but also the ethical 'pulls' of our intuitions concerning the conflict between values. When values conflict we balance them consistent with our intuitions as we weigh-up what is morally acceptable.[7] This balance implies value pluralism; increases in one value at the expense of another is ethically tolerable, but provided this increase does not extinguish the other value. Similar to the incommensurabilist, philosophical argument cannot explain decision-making 'all the way down' and so cannot provide thorough knock-down reasons for certain trade-offs over others. Nevertheless, unlike the incommensurabilist, comparisons between values are made by the intuitionist, as trade-offs assume that the quantity of one value is equivalent to the quantity of another, where X amount of P value is weighed *against* Y amount of Q value, and so on.[8] Indifference curves are produced indicating more precisely which judgements are acceptable, with shapes and gradients reflecting different intuitions concerning how these values are compared and balanced.[9]

---

[7] For seminal expositions of this position, see Brian Barry, *Political Argument – With a New Introduction* (London: Harvester Wheatsheaf, 1993); and James Urmson, "A Defence of Intuitionism," *Proceedings of the Aristotelian Society*, 1974-1975, 111-119.

[8] Also see Robert Sugden, "On Modelling Vagueness," in *Aristotelian Society Supplementary Volume*, Vol. 105, no. 1, 95-113.

[9] B. Barry, *Political Argument*. Also see John Rawls, *A Theory of Justice* (Cambridge (MA): Harvard University Press, 1971), 34-40. It is important to note that the slope of the curve descends from left to right as one value on the vertical Y axis is traded-off against the other value on the horizontal X axis. Moreover, the curve produces a relatively steep descent on the left side with it levelling out on the right, reflecting that a large decrease in an already heavily-weighted value can be justified, even if there is only a relatively small increase in the lesser-weighted value.

Again, concerning public safety, intuitionists could argue that a number of acceptable balances can be found along indifference curves reflecting, say, the conflict between utilitarian and absolutist values. Moreover, these variances help explain the contrasting judgements made at different times concerning the acceptable cost of public safety. Certainly, other factors may be relevant, such as who is seen as responsible for accidents and safety which also vary (EPP, pp. 102-17). Nevertheless, although these factors may change the shape of indifference curves, the character of a trade-off whatever its shape, is understandable in intuitionist terms, relating to what is understood as the right comparative balance between competing values.[10]

Finally, lexical orderings of moral values is famously defended by John Rawls in *Theory of Justice*.[11] Briefly put, his argument is that competing values can be placed in a hierarchy with a first value being fulfilled as a matter of priority (liberty in Rawls' case), after which another value can be fulfilled, without referring to the first (equality in Rawls' case). The key to lexical orderings is ensuring that each value is satiable, so limits are put on its fulfilment, enabling other values to come into play even if these are lower down the priority listing. With lexical rankings, values are certainly compared (unlike the incommensurabilists), as one value has a higher priority over another, but these values are not being traded-off *against* each other as with the intuitionists. For the intuitionist, indifference curves produce a variety of right balances between conflicting values, so tolerably changing the priority of one value over another, most obviously at the two

---

[10] Indeed, it is precisely at these philosophical junctures that philosophers can most usefully "grapple with the question either of what price we should put on life, or how we can make safety decisions without valuing life." (108). And see my arguments in, for example, R. S. Smith, *Equality and Diversity*, 31-58.

[11] J. Rawls, *A Theory of Justice*, for example, 3-53.

ends of the curve.[12] Alternatively, lexical orderings place values in an unambiguous line, with a first-ranked-first-served rule applying, blocking a number of answers or 'balances' being made right simultaneously. Further philosophical argument is required to justify one *type* of ordering over another. But what is common in all orderings is the constraining of values, with a clear understanding of which value should be fulfilled first before moving to the second.[13]

Concerning public safety, lexical orderings may constrain both the utilitarian and absolutist claims so that a delimited version of each can be fulfilled, allowing both values to be prioritised accordingly.[14] For example, a lexical orderer who puts a constrained absolutist claim first could coherently argue for a greater increase in public spending to protect a relatively few number of people, even if this aggregate benefit is less than the aggregate burden placed on a larger number of people. Nevertheless, the constrained character of this absolutist claim still means that a price is put on a life saved, but with the figure tending to be higher than if the lexical ordering was reversed. So,

---

[12] At one end of the curve P value has the greatest weighting compared with Q value; whereas at the other end the weighting is reversed, revealing most starkly the changing relative importance of the one value *over* the other, and especially as the curve extends. This notion of *relative* importance highlights again the point of departure with the incommensurabilists, as relativity implies a comparative weighting, given that one value is viewed as more/less important, *compared with* the other value.

[13] This constraint of principles is necessary for any lexical ordering as J. Rawls explains, *A Theory of Justice*, 40-53.

[14] Putting constraints on either absolutist or utilitarian claims initially may look odd, as both are often presented as maximising principles, and so are unconstrained. However, I argue below that a commitment to either in a certain lexical ordering still allows the moral premise at the heart of each ethical position to be expressed, even if a lexical ordering requires their constraint.

a lexical orderer who puts a constrained utilitarian claim first could coherently argue for a decrease in public spending, if the aggregate benefit is less than the aggregate welfare burden on a larger number. Nevertheless, the constrained character of this utilitarian claim still means that a limit would need to be set on the gains of the larger number of people (relating, for example, to the quality and significance of these gains for individuals in the larger group). The tendency to lower the costs of safety will though be justified from this lexical ordering on the grounds that constrained utilitarian calculations should hold sway first, *before* other conflicting principles come into play.

IV

Where do these meta-ethical considerations lead us regarding the bottom-up approach to applied ethics defended by Professor Wolff? In conclusion, I will briefly outline how these considerations relate to this methodology and the wider lessons about its application to public debate and policy-making. As identified, there are four main rules to bottom-up theorising: (1) The anti-abstraction rule (2) the full description of policy/public debate uncovering moral dilemmas rule (3) the full description of empirical realities ensuring accurate interpretation rule, and (4) the over-arching applied philosophical analysis rule.

Regarding (1) the anti-abstraction rule, certainly getting to grips first with the detail of policy debate can expose well the central moral dilemmas and conflicts relating to public debate and policy formation. However, the claim here is that meta-ethical considerations articulated in their abstract form offer additional 'critical tools' for analysing policy—that is, concerning the philosophical character of these dilemmas and conflicts and how these may be variously interpreted. This leads to (2), the full

description of policy/public debate uncovering moral dilemmas rule, where policy description helps us understand better disagreements in public debate. Again, the claim here is that a more abstract description of these dilemmas and conflicts can operate from the 'philosophical end' as and when additional meta-ethical arguments are exposed. This further philosophical exposition can help explain how and why the same decisions can be differently justified according to each meta-ethical perspective; where different decisions are differently justified according to each meta-ethical perspective; or, where different decisions are justified according to the same meta-ethical perspective. Regarding (3), the full description of empirical realities ensuring accurate interpretation rule, certainly describing empirical realities grounded in good quality empirical research is a necessary component of properly structured philosophical interpretation, as assumptions made in philosophical argument are reliably informed. Nevertheless, factual description cannot reveal how the *same* facts (for example, the fact of value pluralism)—however accurately described or presented—are often variously interpreted by contrasting meta-ethical perspectives. So, while reporting in detail the fact of value pluralism is essential when describing matters relating to public opinion and debate (as EPP so vividly demonstrates), this reporting does not necessarily tell us that much about how value pluralism can be variously interpreted *philosophically speaking*.[15]

Recognising that the same facts can be interpreted differently gets us to (4), the over-arching applied philosophical analysis rule, and to more general lessons about bottom-up theorising. Certainly, when rules (1) to (3) are adhered to, philosophical analysis becomes more pertinent and meaningful for real-world

---

[15] Also see my arguments in R. S. Smith, "Applying Theory to Policy and Practice," 1-18.

policy-making. However, abstract meta-ethical debates concerning the relationship between values allow top-down theorising to come back into play, albeit in a reconstituted form. It will be in a form that does not naively assume the possibility of 'off-the-shelf' ready-made philosophical solutions to dilemmas and conflicts of the kind Professor Wolff quite rightly criticises. Nevertheless, it does provide further philosophically abstract analysis concerning the over-arching theoretical structure of public debate. This latter type of abstraction is firmly connected with real politics as it assesses directly specific issues concerning policy-making. However, it also helps to map more precisely the philosophical terrain as applied to these issues, so not only uncovering why the tensions and conflicts within public debate occur (as identified in EPP), but also how these tensions and conflict can be differently interpreted by both policy-makers and philosophers alike (as identified here).
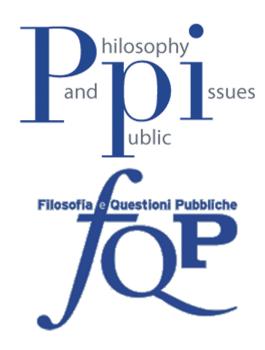
Finally, given this possibility of having various interpretations of the fact of value pluralism exposed through abstract philosophical argument, perhaps it is more accurate to view applied ethics as neither top-down or bottom-up. Rather, it might be seen as a symbiotic *process* between the two approaches— where the role of abstract theorising and the full description of real world policy-making are not two ends of an either/or pole as implied by the terminology of bottom-up and top-down, but are always best viewed as operating simultaneously.

The University of South Wales

If you need to cite this article, please use the following format:

Smith, Steven R., "Meta-Ethics in Debates on Public Safety: A Critical Appraisal of Jonathan Wolff's Bottom-Up Approach to Applied Ethics," *Philosophy and Public Issues (New Series)*, Vol. 4, No. 3 (2014), xx-xx, edited by S. Maffettone, G. Pellegrino and M. Bocchiola

The Policy Maker: On Ethics and Public Policy

BY

ELVIO BACCARINI

The Policy Maker
On *Ethics and Public Policy*

Elvio Baccarini

According to one of the interpretations of *The Prince*, Machiavelli's intention was to write a handbook for the statesman, to teach him how to realize the duties of his political role. An analogous intention may be read in Jonathan Wolff's book *Ethics and Public Policy. A Philosophical Inquiry*. The book, among else, appears as an excellent handbook for moral and political philosophers interested in participating to the creation of public policy. So, what would be wise to do for a political and moral philosopher (in the following part of the text referred to as 'the philosopher') who wants to participate in public policy making?

Wolff's book is rich of insights, sophisticated arguments and enlightening examples. I focus on only some of them. I take for granted that the reader of this article is familiar with Wolff's book, and, therefore, I reduce the descriptive part to the minimum.

I
Pluralism and Principles

I agree with Wolff's thesis that the dedication to too abstract and idealized theories, as well as to the research of the true theory, is not a suitable path for the solution of issues in public policy. I appreciate the intention to make the creation of public

policy a process hospitable for a wide range of different views, as well as Wolff's recommendation of the attitude to compromise and half-meeting. Such goals generally deserve approval.

My question is whether Wolff's recommendations concede too much in the approach to compromise and half-meeting. The risk of such generally laudable goals is to transform the philosopher into a neutral figure and a mediator. On the contrary, he or she can legitimately appeal to and defend a conception of justice in debates on public policy.

This is very important in cases when views and practices that contradict equality or basic liberties and rights are involved in the debate. In such cases to carve the edges of the normative requirements is not sufficient and the wish to look for hospitality for various views has dangerous consequence. Let's think about a case that took place in Seattle. I found the case described by Jacob Levy[1]. Communities that immigrated there practice infibulations, which, as we all know, is an extremely cruel ritual. "Those who do not die of blood loss or infection face a life of great pain during sexual intercourse and great danger during childbirth".[2] As Levy indicates, there was a debate among the committee of the Medical Center and representatives of the community in order to look for a compromise. A base of compromise was found, because at least some representatives of the communities agreed that sunna circumcision (judged by medical experts as analogous to male circumcision) in appropriate hygienic conditions would be sufficient to meet the cultural and religious requirements met by infibulations. The sense was to have a symbolic representation of male supremacy. The policy was not endorsed, and the reason is that there was opposition to

---

[1] Jacob Levy, *The Multiculturalism of Fear* (Oxford: Oxford University Press, 2000).
[2] *Ibid.,* 54.

endorsing a policy that accepts a manifestation of gender discrimination. It seems to me that the refusal of the policy was a case of cruelty, and that the decision was catastrophically wrong. The girls were sent to the countries of their origin with tragic results.

Based on my reading of *Ethics and Public Policy* I believe that Wolff would support the (unrealized) Seattle compromise. This compromise is a case of reasonable policy.

But in my view it is important to remark that the philosopher has to accept this compromise only as a tactical and temporary allowance from justice,[3] and he or she, even in bodies that debate on public policy, has to be not acquiescent with views that affirm inequality as a matter of justice. The leading inspiring intention must not be that "people can get much of what they want without taking too much away from others".[4] The leading intention must be to protect the discriminated as much as possible, and, therefore, that the repressive groups receive the less possible of what they want.

The problem is how the philosopher can support this intention. I agree with Wolff that to put forward a full moral theory (what Rawls calls comprehensive doctrine) proclaiming it as the true theory is not a good way. Such structures are complicated and very sophisticated and probably not accessible to the wider public, like in the case of economic theories or theories in natural sciences. Philosophers who rely on comprehensive doctrines in a public debate will almost certainly be ignored immediately after they have spoken, or even while they are

[3] Even a very firm defender of egalitarian liberalism like Brian Barry is ready for such concessions. See Brian Barry, *Culture and Equality. An Egalitarian Critique of Multiculturalism* (Oxford: Oxford University Press, 2000), 19-54.
[4] Jonathan Wolff, *Ethics and Public Policy. A Philosophical Inquiry* (London: Routledge, 2011), 203.

speaking. This practical problem is relevant, but I indicate another problem, as well. Even if, in some way, one succeeds in passing the law that follows from his or her theory (perhaps, in virtue of some contingent political alliance), the decision is not legitimate. I say this in virtue of the liberal principle of legitimacy. John Rawls's formulation of it is the most relevant. He says: "Our exercise of political power is fully proper only when it is exercised in accordance with a constitution the essentials of which all citizens as free and equal may reasonably be expected to endorse in the light of principles and ideals acceptable to their common human reason."[5] Contrary to Rawls and in agreement with some public reason philosophers, I think that the public reason requirement, properly reformulated, applies to all decisions on public laws and policy and not to constitutional essentials only. To be sure, the principle does not look for the endorsement of all actual people in actual societies, but looks for a level of idealization. The endorsement of only reasonable citizens is all that is required. Such are citizens that appeal to the basic principles of the family of eligible conceptions of justice: (a) certain basic rights, liberties and opportunities; (b) the assignment of special priority to them; (c) measures to ensure to all citizens the use thereof.[6] They appeal to other principles, values and ideals that citizens who view each other as free and equal can endorse, as well.

In brief, I totally agree with Wolff's recommendation to abstain to appeal to the 'true moral theory' in public policy engagement. But I rely on an alternative possibility of philosophical background that Wolff does not defeat in his book, at least not explicitly. This is Rawlsian political liberalism that,

---

[5] John Rawls, *Political Liberalism* (New York: Columbia University Press, 1993/2005), 137.
[6] *Ibid.*, 6.

although it recommends not to appeal to comprehensive moral doctrines and to moral truth in public policy debates, it requires the philosopher not to renounce to the reasonable conception of justice. According to political liberalism, it is important to translate comprehensive doctrines in the form of public values, i.e. values that we can bona fide expect to be accessible to reasonable fellow citizens that see each other as free and equal and that they can reasonably endorse. In practice Wolff's and the political liberal's recommendations may converge frequently, as is visible, for example, in the chapter on scientific experiments on animals. There he indicates that the very general methodologies, i.e. those that, inside a comprehensive doctrine, look for a master-feature that gives credentials for entering in the realm of beings with moral status, have proved unsuccessful. Wolff proposes a more nuanced methodology, one that consists in relying on features that we take as morally relevant, and making proper use of them. For example, because of the fact that we take pain as morally relevant, we have to attribute proper moral consideration to animals that feel pain, and, consequently, to protect them from pain.[7] But, importantly, the philosopher who is a political liberal defends, among else, his or her position by remarking the supreme relevance of principles that protect equality and basic rights and liberties. We can carve them on the edges and regulate them, but not neglect them, and only in extreme cases and in temporary limits they can be overridden at their core meaning. For the political liberal it is important to be engaged to obtain that everybody "can get much of what they want without taking too much away from others,"[8] but only as far as the view of each citizen as free and equal is not threatened. When there is collision, principles that protect freedom and equality must be privileged over intersubjective agreement. My question is whether this is

---

[7] Wolff, *Ethics and Public Policy*, 11-36.
[8] *Ibid.*, 203.

acceptable for Wolff. There is indication that it is,[9] but I think that the question is legitimate, because Wolff's discussions where he gives detailed explanation of his view regard cases when only edges, and not the core of liberty and equality principles, are threatened, and I have doubts about what his position would be when the core would be menaced. My doubts also regard the reasons that for Wolff permit changes that are highly contentious or unpopular, and have no consensus behind them.

In any event, it seems to me that having a strong commitment for the protection of equality and basic rights and liberties is a virtue of political liberalism.

Wolff can object, here, by relying on the experience about public decisions that regard the regulation of gambling, where it appears that an irremovable appeal to liberty principles is not recommendable, because the commitment of the wider audience to them is not firm as it is among many philosophers. Most people appear ready to take distance from liberty principles when the regulation of gambling is at stake. But political liberalism does not suggest the rigid top down appeal to principles. It is related to the Rawlsian method of reflective equilibrium that gives the optimal resource to philosophers in their engagement in the public decision making process.[10] Reflective equilibrium recommends considering adjusting principles in relation to widely

---

[9] "This is not to say that radical, discontinuous, change is impossible. Slavery was abolished. Neither must there always be a consensus behind the change. Often change is highly contentious or unpopular" (35).

[10] In my opinion, the final, applicative, chapter of Mill's *On Liberty* is an exercise of reflective equilibrium, with the intention to develop understanding of his liberty principle. Unfortunately, I am not able to argue for this, here.

shared beliefs. At the same time, it does not recommend to surrender to such beliefs.[11]

To put it with Wolff's example of gambling, the mere fact that most people are ready to accept restrictions to gambling is not a decisive reason for the liberal philosopher to renounce to the firm appeal to a liberty principle in such a case. But reflective equilibrium suggests that wide and strong rejection is a reason for the philosopher to re-think about the principles that he or she endorses. For example, in front of wide and strong rejection of unlimited access to gambling, or of the decriminalization of drugs, the philosopher must reason about whether they really represent central cases covered by liberty principles. There is no guarantee that strict formulations of liberty principles win against limitations. But the reason is that we are not always dealing with instances of protection of the core meaning of liberty principles (it is difficult to imagine that, in ordinary situations, for someone gambling and consuming drugs are linked to a conception of good life or living in accordance to conscience, some of the core rationales for liberty principles). When this is the case, the wise application of the method of reflective equilibrium requires adjustment on the edges. But sometimes there may be an abuse over someone's conception of good life, or conscience (for example, in the case when consuming drugs is related to religious rituals), and in such circumstances the philosopher is right in insisting against the limitations, even if this means that he or she is not taken with consideration in a stage of the debate. It is important to defend principles, when they are concerned with basic rights and liberties. In practice there can be frequent convergence with what Wolff recommends, precisely in all cases when the core meaning of basic rights and liberties is not

---

[11] John Rawls, *A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971/1999), 18-19, 42-45.

threatened, and most probably, this is the case in the issue that Wolff debates, i.e. gambling. But the adjustments that, by relying on Rawls's political liberalism, I recommend are limited to cases of nested inconclusiveness. Such are situations where competing conclusions can be reasonably justified on the base of the same principle, and the principle is conclusively justified.[12] In other words, such are cases where there is reasonable pluralism about the understanding of the principle, but not about the principle itself. For political liberalism, the re-thinking of the principle is generally fruitful only in the limits of nested inconclusiveness. If a law or public policy is opposed to all reasonable interpretations of a principle that protects a basic liberty or right, then it requires full opposition. As I said above, I see this as a virtue of political liberalism.

The reason of my uncertainty about whether Wolff is committed to the protection of equality and basic rights and liberties with the same strength of political liberalism is that his detailed discussions are limited to cases of nested inconclusiveness of reasonable understanding of principles.[13] My question is what his attitude is in relation to the protection of the core meaning of liberty and equality principles and (if he is ready to accept it) which kind of compromise he would be ready to accept as far as the core meanings of these principles are concerned.

---

[12] Gerald Gaus, *Justificatory Liberalism* (Oxford: Oxford University Press, 1996), 157-158, 180-182.
[13] Specifically, of Mill's liberty principle, where Mill himself is aware of the fact that the core of the liberty principle is probably not sacrificed by some restrictions.

II
Consistency and the Advantage of the *Status Quo*

Wolff indicates the issue of the legal status of consuming drugs as a case where it appears that a KO resource in philosophical discussions is not so conclusively powerful in public policy disputes. This is the appeal to (in)consistency. As Wolff indicates, one of the favourite philosophical moves appears not to be efficacious in such debate. Usually, the philosopher investigates whether there are relevant differences between consuming alcohol and consuming drugs, and requires an identical regulation if there are not. Wolff says that this is not sufficient in public policy controversies. The basic reason is that, contrary to what happens in philosophical debates, the question, here, is not what is just at an abstract level, but what we must do, starting from where we are: what would result from de-penalizing the use of drugs in our actual situation?

In my opinion, it is not so clear that the appeal to inconsistency is so weak an argument as Wolff wants to show. Precisely, I am not sure that Wolff has really offered a counter-example to the strength of the appeal to inconsistency in public policy debates. Consistency, here, corresponds to a valid application of the principle of universalizability. To put it simply, the principle requires that if we make a normative decision (express a judgment) on the base of a (set of) feature(s), we must make the same decision when the same relevant features are present (and there are no defeaters of them). So, in order to offer a counter-example to the relevance of consistency, Wolff has to show that we can plausibly expect that the alcohol and the drug case do not merit the same reaction, even if they are equivalent in the (set of) reason(s) we have found as relevant.

It does not seem to me that Wolff really indicates examples where the use of drugs has legitimately a different treatment from

the use of alcohol, while the two cases are truly equivalent. A point that Wolff expresses is crucial. When we judge about this, as well as in other issues, it is important to debate the issue not in abstract, but by focusing on the contextual elements. Exactly as Wolff says, the question is not simply whether we have good reasons to criminalize drugs, but whether we have good reasons to modify the actual legislation, when drugs are already criminalized. What would the decriminalization of drugs imply? It may happen, for example, that it implies the overall increase of criminality. This is a reason to avoid the decriminalization of drugs. What would the criminalization of alcohol imply? Perhaps, again, a catastrophic increase of criminality (as, in fact, happened in the age of prohibitionism). But these reasons differentiate substantially the case of drugs from the case of alcohol. There is no inconsistency between forbidding one of them, and not criminalizing the other one, or leaving the latter legal, while leaving the former criminalized. We cannot universalize from one case to the other, because there are relevant features they do not share.

An inconsistency might appear, nonetheless, in the treatment of people. The worry is that by criminalizing consuming drugs, we stigmatize the behaviour and the person in relation to drugs, while we do not do this in relation to alcohol, although they are equivalent in the relevant intrinsic moral features. Still, the inconsistency of criminalization is not proved. It is important to remember that consuming drugs is penalized not for its intrinsic features, but for its external contingent features. But the reference to stigmatization extends the discussion and opens a new question, where claims of consistency can reappear with authority. Perhaps, we are speaking of a case when criminal sanctions and stigmatization do not come together.

To be sure, if we stop speaking ex hypothesis and look at the drugs case concretely, we see that the possible bad consequences of the decriminalization of drugs are still unproved. Now, for Wolff it seems that the possibility is sufficient for not realizing decriminalization. There is an advantage for the status quo, and until reasonable doubts are removed the new policy must not be introduced. But is there really an advantage for the status quo in general? This is disputable.

Probably, it is reasonable to argue for the advantage of the status quo in the case of recreative drugs. The consequences of decriminalization can possibly be deleterious, while recreative drugs are not a central component of conceptions of the good life and of living in accordance to conscience for people. Not so clearly reasonable in the case of medical use, like the medical use of cannabis. Although perhaps there is still no clear evidence that it really helps, many people strongly believe that it does. Forbidding the use appears as an interference into a domain that is rather embarrassing to control by state interference, i.e. the domain of people's decisions in central matters of their health. This appears as a central domain of decisions in people's life where the person involved has a strong presumption of the right to have the last say. The point is, perhaps, even stronger in the case of ritual religious use of drugs.[14]

Another example where, in my opinion, there is no advantage for the status quo is mentioned by Wolff in his book: the decriminalization of homosexuality. I add a further example, the use of assisted procreative techniques, opposed by many in the past. In such cases, there are clear discriminations of some people in some central aspects of their lives. In the former case, the right

---

[14] Cfr. Samuel Freeman, "Liberalism and the Accommodation of Group Claims," in P. Kelly (ed.) *Multiculturalism Reconsidered*, (Cambridge: Polity Press, 2002), 24.

to have a full realization of emotional life, something that is taken for granted for heterosexual people. In the latter case, the right to have genetically related children (for at least one member of the couple). Again, this is something taken for granted for most people. Here we have clear evidence of the damage of the status quo. In both cases, central aspects of life aspirations and of components of part of what many people take as a full life realization, taken for granted for many people, are denied to other people.

A further case might be represented by enhancement policies. Some people strongly oppose them in virtue of precaution that regards innovations. But perhaps we urgently need enhancement (for example, cognitive enhancement) for finding solutions for many bad things that happen to humans in virtue of their fragility: threats to health, disease, natural catastrophes, etc.[15]

The clearest cases, however, are cases of discrimination. "Justice too long delayed is justice denied," said M.L. King.[16] King's thought is fully proper for the decriminalization of homosexuality, as well as for assisted procreation if delayed long time after the technique became available. On one side, we have clear cases of discrimination,[17] on the other side only conjectures

[15] John Harris, "Moral Enhancement and Freedom," *Bioethics* 25 (2011): 102-111. I find reasonable supporting the advantage of status quo in some of the cases remarked by Persson and Savulescu, precisely cases when we can lose a lot, for a possible marginal advantage, even when the probability of the former is much smaller than the probability of the latter. Ingmar Persson and Julian Savulescu, *Unfit for the Future. The Need for Moral Enhancement* (Oxford: Oxford University Press, 2012).

[16] Martin L. King, "A Letter from Birmingham Jail," in *Civil Disobedience in Focus*, ed. Hugo A. Bedau (London: Routledge, 1991), 72.

[17] I take the right to assisted procreation, here, as a negative right, and I accept as a reasonable question whether, in the world of scarcity where we live it is a positive right, as well. Cfr. Colin Farrelly, *Justice, Democracy and Reasonable Agreement* (Basingstoke: Palgrave MacMillan, 2007).

about possible future harms. Why is the status quo to be privileged? It seems to me that here we have a case of unreasonable attitude adverse to risk. The devil that we know is not always the better.

Now, there is certainly a sense in which the social condition must mature in order to implement rights in these, as well as in other fields. As a matter of fact, in democracies, typically, a right can be implemented only when a sufficient part of the demos is ready to accept it. So, Wolff is right that the decriminalization of homosexuality was available, as a matter of fact, only when there was a sufficiently numerous part of the demos around the idea that the state must not interfere in the moral life of people. What was required was, at least, a division in the demos. But this is no reason for the philosopher not to fully and uncompromisingly oppose criminalization even before the society is ready to accept it, and to do this not only in his or her separate, future oriented activities, but in his or her engagement with actual public policy, as well.

Does this imply the risk for the philosopher to be ignored, or marginalized in the creation of public policy, and that the debate will continue without him or her? Perhaps this can happen in specific committees, councils, etc., where he or she risks finding the incredulity, scepticism, interest-based rejection, or even the sense of moral disgust of the interlocutors. But taking part in such bodies is not the only form of engagement with public policy. The philosopher can, for example, try to find an institutional alliance in the Supreme Court, as the moral/political philosophical dream team has tried to do in relation to the right

to physician-assisted suicide.[18] Their attempt has failed, but they have indicated an important possible path.

Another one is participation in activities of civil disobedience. This practice can be very important for the vitality of democracies, sometimes even crucial, as Daniel Markovits shows from a republican point of view.[19] Here, there is an ample space for philosophers to influence policy making in the immediate and actual, even when the views that they defend are more enlightened than the common sense of the age, and, for this reason, they would be marginalized in the mainstream institutions.[20]

The University of Rijeka

Ethics and Public Policy. Response to Baccarini,
Kulenovic, Lang, Lever, Peter, and Smith

BY

JONATHAN WOLFF

Ethics and Public Policy
Response to Baccarini, Kulenovic, Lang, Lever, Peter, and Smith

Jonathan Wolff

I t is an extraordinary privilege to have the opportunity to respond to such a powerful array of responses to my book *Ethics and Public Policy* as faces me here. I'm honoured that the editors of this journal were prepared to devote so much space to this exchange, and to the contributors to spending such time and care in preparing their papers.

I

Understandably many of the contributions concentrate primarily on methodological issues concerning the relation between political philosophy and public policy. They raise questions about what I say I am doing in the book, what I really am doing, and what, perhaps, I ought to have been doing or should do in future work. Let me start with Enes Kulenovic's extremely helpful classification of six different models of the relation between political philosophy and public policy. First, there is the Syracuse model, with Leo Strauss as an exemplar, which 'wants public policy to be directed by a philosophical principle', on the grounds that the political philosopher has access to the truth. Second, the Rawlsian model, inspired by Rawls' later work, identifies the basic values that a society identifies with, drawing primarily on arguments that are acceptable in public debate. The third model is called the 'value-pluralist' model,

inspired by Isaiah Berlin, and recommends the strong advocacy of one's own position while at the same time recognising that other positions, which tragically will have to be excluded, may also have validity. The fourth model, the proceduarlist, is Stuart Hampshire's extension of Berlin, which advocates procedures to deal with value pluralism, in which all views are given a hearing. Next we come to the 'Wizard-of-Oz' model, where the philosopher reveals, and critiques, the assumptions hidden behind the curtain of ordinary debate. And then, finally, we have the bullshit model, in which the philosopher acknowledges that there are no decisive considerations, but acts as a 'gun-for-hire' in public debate.

Reading through the paper my anxiety was how I was going to be classified. With the exception of the two extreme cases, I saw elements of each of these models in my thinking, which is exactly what Kulenovic concludes. Does this show fatal equivocation on my part? Kulenovic is kind enough to suggest not, and that the different models can work together. I would also suggest that there may also be variation by problem. In some cases, for example, it may be essential that all views get a hearing, when fundamental questions of life and death are involved. In others, where decisions have to be made urgently but the issues are less serious, then arguably less elaborate procedures would be appropriate. But generally the message is that you have to take cases as you find them. If you set out deciding that in every case there is unmasking to be done, then your approach will come over as forced and dogmatic. But if you never consider whether it is time for a bit of unmasking then you risk naivety. Different cases may need different approaches, and you will not generally know in advance.

II

Steven Smith also concentrates on methodological questions, in effect constructing my 'missing chapter' on methodology, drawing out four 'rules' from my scattered comments. But I have to confess that I break out into a cold sweat when particular rules or principles are attributed to me. For I doubt that there are rules, in the sense of firm guides that must be followed, in moral and political philosophy. I have long been sympathetic at least to the title of Feyerabend's book *Against Method*. Unlike, for example, in accounting practices, or routine engineering, we are on very uncertain terrain in applied moral and political philosophy and have to make up the rules as we go. At best there are helpful starting points.

The rules attributed by Smith to me are:

1) The anti-abstraction rule.

2) The full description of policy/public debate uncovering moral dilemmas rule.

3) The full description of empirical realities ensuring accurate interpretation rule.

4) The over-arching applied philosophical analysis rule.

The first states that ethical principles should not be articulated before considering the specific policy issues at stake. The second and third are perhaps self-explanatory, ruling that full descriptions of both the ethical dilemmas and the empirical realities are needed. And the fourth states that rules 1-3 have to be followed before philosophical analysis can usefully be applied.

Do I accept the these rules? Regarding numbers 2 and 3 I recall Nancy Cartwright once observing in a fraught committee meeting that 'full discussion' is a 'bogus concept', and the claim that 'we shouldn't make a decision until we have had a full

discussion' is generally merely a delaying move. But if we replace 'full' with 'extensive' I have no quarrel with these two rules. What do I think about (1) and (4)? I can understand why I can be read as endorsing these rules, but I can also see why they are problematic. First, it is hard to see how we can articulate the dilemmas of public policy without using normative language, which may well include ethical principles. Hence rule 2 seems to assume the denial of rule 1. And, more subtly, if we accept, as I do, that observations are very often theory laden, then rule 3 also pre-supposes rule 1, in that descriptions of reality very often are as they are because of an individual's value commitment. Accordingly rule 4 is also false.

Reflecting on my position from this vantage point, I would want to replace rules 1 and 4 with warnings that although abstract principles can be appealed to at any point in the discussion, doing so is very likely to lead to problematic policy recommendations (empty, or out of date, or counter-productive) unless there has been extensive engagement with the policy dilemmas, empirical conditions, and history of the area under discussion.

Smith argues for what I think is a similar conclusion by a more substantial route. He very clearly sets out three approaches to value-pluralism: incommensurability, intuitionism (allowing trade-offs on the model of indifference curves) and lexical priority, and carefully explains how these different meta-ethical approaches will construct policy dilemmas in different ways. Hence, he argues for a symbiotic relationship between abstract philosophical reflection and empirical investigation. I think that this must be right. The last thing I would want to do is replace one one-sided model with its mirror image. Nevertheless, there is still to me something right in the idea that we should start with problems in need of solutions, rather than theories in need of problems to solve.

III

Annabelle Lever is also concerned with methodological issues, and after a very generous and insightful discussion of the main themes of the book—a discussion that in many ways improves on what I say myself and links it to some of my other work—gently accuses me of not pushing the methodology as far as I should. Specifically, she points out, I pay very little attention to issues of power, voice, and implementation of policy, and also to issues of standards of evidence. I am not going to try to defend myself against these points as they are very fair and important criticisms. These are significant gaps in the book, and they do need to be addressed. Lever hopes for a sequel in which I pay attention to the very difficult issues that take place 'before the philosopher enters the room': questions of who is asked for their input, what they are told about their role, why they are selected, and what is going to count as the evidence they are to assess and use to inform their recommendation. It is a reasonable request but a daunting task, going beyond philosophy and public policy to the sociology and politics of knowledge and power. Just to take one prominent example, drug regulation policy in the UK has been surprisingly resistant to evidence about the harms of drugs both relative to other drugs and relative to the other risks of life (alcohol, and, notoriously, horse riding). How to make progress in a policy area which is explicitly set up to receive, review and take account of scientific evidence but then routinely ignores it, requires a depth of understanding of politics, sociology and the media. Writing on these topics will be a challenge for me and in the spirit of co-operation my own hope is that someone else— perhaps Annabelle Lever herself?—writes the missing sequel before I have a chance.

IV

Fabienne Peter, again after providing a generous and insightful account of the main themes of the book focuses on the question 'why does agreement matter?', especially from the point of view of political legitimacy. Peter asks

> How can we make sense of the idea that agreement about a moral judgement matters more than the truth of that judgement in the sense that agreement is both necessary and sufficient to legitimise basing a policy choice on that judgement?

However, when stated in such terms, I wonder whether anyone really does hold the stated position that agreement is necessary and sufficient for legitimising a policy choice. This is one of many areas where, while the concepts of necessity and sufficiency are indispensible for clear thinking, they come to be something of a straightjacket when it comes to setting out positions to be defended. Agreement, surely, cannot be necessary for legitimacy, as a single hold-out would render a policy illegitimate. But actually what Peter means by agreement includes democratic procedures in which minorities can be outvoted, and her point is to distinguish procedural, agreement based approaches from 'truth-based' approaches to decision making, correctly noticing that for me agreement is more important than truth. Is agreement (unanimous or democratic majority) sufficient for legitimacy? Yes, if it is well-informed and non-coerced, and perhaps meets some other conditions. But then that is a way of saying it is isn't sufficient after all. Or to put it closer to Peter's own terms, we must distinguish a descriptive and a normative sense of legitimacy. Even if, descriptively, everyone does in fact agree, still from a normative point of view something may nevertheless have gone wrong. On this view, the connection between agreement and normative legitimacy must be looser than the quoted position assumes. I think I would be prepared to

endorse the view that the greater the agreement the greater the probability that the outcome will be legitimate, but even the limit case of full agreement is not enough to guarantee legitimacy, if, for example, there are adaptive preferences to oppressive policies.

I'm not, though, completely sure that the concept of legitimacy is as key to my thinking as Peter would like it to be. I deliberately avoided engaging in the high political philosophy of legitimacy as developed by Larmore, Gaus and Peter herself. I'm largely concerned about a more prosaic question that has taken something of a backseat in philosophy. I'm interested in what are now often called 'second-round' effects. How can we best ensure that any changes we introduce will be accepted and followed, rather than being seen as obstacles to be ignored or worked round, which, given infinite human ingenuity, will happen if the changes are not accepted? To take a very topical example, I was on a committee that recommended that the number of 'fixed odds betting terminals' per bookmaker outlet in the UK would be limited to a certain number. We also, for other reasons, recommended the relaxation of some planning and zoning restrictions. We did not, however, consider it a serious possibility that these particular betting terminals would be so profitable that it would be worthwhile for the same company to open more than one bookmaking outlet on the same street in order to have more terminals, and that the changes in zoning, for the first time, would make this possible. As a result of bookmakers working out how to 'get round' regulations they clearly did not agree with, the number of bookmaker outlets in the UK has risen, against our intentions, at least in prominent locations, to great public disquiet.

Clearly there are connections between legitimacy and the 'bindingness' of laws and regulations, although there will be many factors that determine whether such laws will tend to have

unwelcome unintended consequences. Of course I am interested in the normative questions of legitimacy that Peter raises, and discusses with insight and originality, but not to the exclusion of the descriptive questions of how regulations and policies will actually be received. Achieving normative legitimacy will be little consolation if new policies badly backfire because they have not achieved the approval in practice of enough of those who have to follow, implement, or enforce them.

V

Elvio Baccarini raises a number of important questions, but here I will focus on just one: his worry that the methodology of seeking compromise as a way out of disagreement concedes too much to obnoxious moral positions. He discusses the example of 'infibulations', also known as genital cutting, focusing on a case in Seattle described by Jacob Levy in which a hospital offered to substitute a much less intrusive, safe and largely symbolic version of the operation in hygienic conditions in place of the dangerous and brutal versions of the operation that were currently taking place. This compromise, however, was refused on the grounds that it was wrong to collude with policies of gender discrimination. As a result young girls were sent back to have the operations in their countries of descent, with 'tragic results'.

Baccarini is right that, on the basis of the limited information given in his description, and the small amount of other material on the topic I have read, I would be inclined to support the hospital policy, in the hope it will prevent much worse practices, although I would want to look at a wider range of views and concerns before making a final recommendation. But nevertheless, so far, as I take it, we agree. But Baccarini believes we differ in that, given that genital cutting, however minor and

symbolic, is an exercise in male supremacy, any compromise should be regarded as a 'tactical and temporary allowance from justice', for, as I read his argument, Baccarini regards the practice of genital cutting as an expression of an 'unreasonable' conception of the good in Rawls's sense, offending against freedom and equality. I certainly agree with Baccarini that there can be policy proposals that so offend against freedom and equality that they should not be tolerated, even in a tactical sense. For example, consider a religion that decided that all first born children should be male (or female) and practiced infanticide for children of the wrong sex. There is every reason for prohibiting this practice with severe penalties, and a hospital that offered to collude with the practice, even to prevent more painful forms of infanticide would rightly be condemned. However the lines are hard to draw. Baccarini, in effect, offers a three way distinction. First, there is a compromise we can be happy with in the long term. Second, there is a 'temporary and tactical' compromise. And third, areas where no compromise is acceptable. Here we are concerned with the distinction between the first two.

Consider the practice of abortion. For many people abortion is simply a former of murder and hence in conflict with freedom and equality. Some of these people regard abortion as others would treat infanticide and actively, and sometimes violently, campaign for its complete, or near complete, prohibition. But there are others who oppose abortion but treat it in the same spirit that the Seattle hospital offered to treat genital cutting. Knowing that desperate women, especially those of low income, will turn to very dangerous forms of abortion, they will accept hospital based abortions under restrictive conditions as the lesser of two evils. Such opponents of abortion have Baccarini's attitude, that this is a 'tactical and temporary' retreat and would hope to convert others to their view. Yet many others consider current policy on abortion as a sound compromise that should be

preserved unless facts change or knowledge improves in relevant ways. In other words, how to draw the distinction between reasonable and unreasonable conceptions of the good, and therefore permanent and temporary compromises, already seems to depend on one's conception of the good, and there is no truly neutral ground. Accordingly I am unsure how to draw the distinction between 'tactical and temporary' compromises and some sort of valuable permanent compromise. While there may be a distinction to be drawn, different people will draw the line in different places and put different compromises on different sides. The key lesson, though, from Baccarini is that even if you accept a compromise for the purposes of public policy, this should not stop you from arguing and trying to convince others that further changes are needed, if that is what you think.

VII

Finally Gerald Lang discusses my chapter on crime and punishment. Lang correctly observes that this is in a way less empirically well grounded than other discussions in the book, as my work on crime and punishment has been less extended than on some other topics. My motivation for this chapter was the observation that the philosophical discussion of punishment has often taken place with only minimal reference to the crime for which punishment is said to be deserved. I wanted to explore the consequences for a theory of punishment if one starts from the intuition that punishment should in some way be responsive to what it is that we find objectionable about crime. After all, if we were not especially worried about crime then it is unlikely that we would have developed substantial practices of punishment. And in particular I wanted to consider what it is like to be the victim of a crime, believing that it would be fruitful to consider the

difference between being the victim of crime and being the victim of an unsuccessful attempt. I conjectured that when a crime is disrupted then although the attempt is disturbing nevertheless one may emerge with a feeling of triumph. This would not be the case when an attempt succeeds. I further conjectured that the difference is that a successful attempt in some way reduces the victim's status, at least in their own eyes, and that punishment is an attempt to rebalance status.

Lang, while generously showing a degree of sympathy, provides a sophisticated argument designed to show that my account cannot explain or justify our practices of punishment, and substitutes his own theory in which punishment seems to be derived from the right to self-defence. To simplify a complex line of thought, one question is that, even if rebalancing of status is needed, why is 'hard treatment' in the form of punishment the correct way of doing this? This objection forces me to reflect on the ambition of my account, and my feeling now is that it should not have been offered, if it was, as a sole account of punishment, rather than one factor, among others that explain and justify our practices. And it is probable that Lang's account is another element in the patchwork of justifications to which we must appeal. For Lang's theory, if we take it as a singular account, also leaves us with some puzzles. If I am a shopkeeper and you attempt to rob my shop, in self-defence I can probably lock you in my store room for an hour or two until the police arrive. But I can't keep you there for three months, even if the crime would normally lead to a three month sentence. Or for another example, many minor offences lead to a fine, but as a victim I have no right to take your money in 'self-defence'. We seem to draw on different sources to calibrate justified self-defence and justified punishment. Lang, no doubt, has answers but considerable elaboration and refinement, which may well include appealing to

other considerations, will be necessary to generate a complete account.

VIII

In brief conclusion I would like to repeat my thanks to my commentators for their excellent contributions. All of them have led to changes in my view, or clarifications, or the recognition that more is needed. They are the sort of comments I would have liked to have received while still working on the text, and undoubtedly would have improved the final version. One day, perhaps, there will be a revised edition where I can make proper use of these excellent thoughts.

University College London