Agreement-Based Practical Justification: A Comment on Wolff

BY

FABIENNE PETER

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.

<div align="center">

I

Wolff on Public Policy, Philosophy, and Disagreements

</div>

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