

# Moral uncertainty and public justification

Jacob Barrett (Global Priorities Institute, University of Oxford) and Andreas T Schmidt (University of Groningen)

Global Priorities Institute | September 2021

*GPI Working Paper No. 15-2021*

# Moral Uncertainty and Public Justification

Work-in-progress draft by Jacob Barrett and Andreas T Schmidt

**Abstract:** Moral uncertainty and disagreement pervade our lives. Yet we still need to make decisions and act, both in individual and political contexts. So, what should we do? The moral uncertainty approach provides a theory of what individuals morally ought to do when they are uncertain about morality. Public reason liberals, in contrast, provide a theory of how societies should deal with reasonable disagreements about morality. They defend the public justification principle: state action is permissible only if it can be justified to all reasonable people. In this article, we bring these two approaches together. Specifically, we investigate whether the moral uncertainty approach supports public reason liberalism: given our own moral uncertainty, should we favor public justification? We argue that while the moral uncertainty approach cannot vindicate an exceptionless public justification principle, it gives us reason to adopt public justification as a *pro tanto* institutional commitment. Furthermore, it provides new answers to some intramural debates among public reason liberals and new responses to some common objections.

**Key words:** public justification; moral uncertainty; public reason liberalism; moral disagreement

# 1 Introduction

Moral disagreement pervades our lives. We disagree about the rightness or wrongness of actions, the goodness or badness of outcomes, and the justice or injustice of institutions. These disagreements often seem quite reasonable – and equally intractable. Moral reasoning is hard, requiring us to navigate complex concepts and their intricate and often surprising implications. We come to this task with different life experiences, educations, and social networks, and so with different biases, priors, and evidence bases. And even when we agree about the considerations at issue in some case, we often disagree about their weights. Moral thinking, in other words, is subject to the “burdens of judgment” ((Rawls 2005, 55–57); compare (MacAskill, Bykvist, and Ord 2020, 11–14)). And it is a predictable consequence of these burdens that intelligent people reasoning in good faith will come to different conclusions about morality.

Given the many plausible moral views available to us, and their many capable and eager champions, it is difficult to know how to proceed. We must reckon both with the fact of our own uncertainty about morality, and with the fact that others will inevitably come to different conclusions than we do. These two facts, though closely related, have spawned two very different research programs in contemporary analytic philosophy: public reason liberalism in political philosophy and the moral uncertainty approach in ethics. Public reason liberals ask what laws are justified among individuals who reasonably disagree about morality. They argue that we should take all reasonable positions into account – holding that a law is justified not when the moral view we find most plausible says it is, but when the law can be justified to all reasonable people. Moral uncertainty theorists are concerned with what we morally ought to do in the face of our own uncertainty about morality. They argue that we should take all plausible moral positions into account – holding that what we morally ought to do depends not only on the moral theory we find most plausible, but also on the verdicts of all other moral theories in which we place some positive credence.

Our goal in this article is to bring these research programs into contact. To frame our

discussion, we investigate the hypothesis that the moral uncertainty approach lends support to public reason liberalism. Our tentative conclusion is that while the moral uncertainty approach cannot vindicate the stringent principle that all laws be publicly justified, it nevertheless provides several reasons to take public justification seriously. Specifically, from the perspective of moral uncertainty, a good case can be made for treating public justification as a weighty *pro tanto* institutional commitment – albeit one that can be overridden when the moral stakes are high.

Along the way, we also highlight some attractive features of our novel defense of public justification. For example, critics often argue that existing defenses of the public justification principle fail to cohere with the principle itself, because they assume controversial first-order views about morality or justice, either explicitly or in the way they narrowly delineate the class of “reasonable” people. The moral uncertainty approach sidesteps this issue, because it permits uncertainty about morality and justice all the way down and relies on a thin and independently motivated notion of reasonableness. Moreover, the moral uncertainty approach offers a fresh perspective from which to resolve some contested intramural debates among public reason liberals, not only about when to count someone as reasonable, but also about what it takes to justify a law to a reasonable person, and about the role of “shared reasons” in public justification.

We proceed as follows. In section 2, we outline public reason liberalism and the moral uncertainty approach and introduce our hypothesis. In sections 3, 4 and 5, we discuss arguments in support of this hypothesis. We comment on intramural debates on public justification in section 6 and conclude in section 7.

## **2 Two Second-Order Approaches to Justification**

### **2.1 Public Reason Liberalism**

How should we justify institutions and state action? In this article, we focus on coercive *laws*. But much of what we say applies, *mutatis mutandis*, to other institutional questions and forms of state

action – and perhaps even to the justification of informal institutions or social norms.<sup>1</sup>

On what we call the *first-order moral approach*, laws are justified if and only if they are justified according to the correct moral theory. Political philosophy, on this view, is about finding the correct moral theory and applying it in the political domain. To determine which laws are justified, we must figure out, first, if communitarianism, libertarianism, utilitarianism or some other theory is the correct view, and, second, what laws the correct view supports. This does not mean that utilitarians, for example, must ignore moral disagreement. Instead, disagreement is factored in as an empirical constraint: what laws do the most good when not everyone is a utilitarian?

Many political philosophers think that moral disagreement and the pluralism characteristic of modern societies make the first-order moral approach unsuitable for political justification. Instead, we need a mode of justification that takes disagreement seriously. The most common proposal is the *public justification principle*: governments should only exercise their power in a way that can be justified to every reasonable person in the population.<sup>2</sup> Call a law publicly justified when it is justified to all reasonable people, and a law publicly unjustified when it is not. The public justification principle imposes a prohibition against enforcing publicly unjustified laws, not a requirement to enforce publicly justified ones.<sup>3</sup> It says that governments are prohibited from

---

<sup>1</sup> Public reason liberals disagree about the object of public justification. For example, (Rawls 2005, 215) holds that only issues of basic justice and constitutional essentials must be publicly justified, but that “it is usually highly desirable to settle political questions by invoking the value of public reason.” (Quong 2011, chap. 9) extends the requirement of public justification to all laws; (Gaus 2011, 490–97) extends it to social norms; (Waldron 1993, 36–37) applies it to “all aspects of the social world.”

<sup>2</sup> See, for example, (D’Agostino 1996; Gaus 2011; 2011; Larmore 1990; Lister 2013; Nagel 1987; Rawls 2005; Vallier 2019; Waldron 1993). (Vallier 2018) provides additional citations.

<sup>3</sup> The public justification principle is often seen as a requirement on *legitimate* state action, which on some but not all conceptions of legitimacy is equivalent to permissible state action. We avoid that contested concept here and speak directly about what governments are morally permitted to do or prohibited from doing.

enforcing laws that some reasonable people reject.

Theorists that accept the public justification principle are known as “public reason liberals”. We use the same label, though it is a bit confusing since not all public reason liberals think that justification must proceed via “public reasons”. Roughly, we can distinguish between *consensus* and *convergence* liberals (D’Agostino 1996; Vallier 2011). According to consensus liberals, in determining whether a law is publicly justified, we must bracket people’s private or non-shared reasons and ask whether the law could be justified to all reasonable people factoring in only the reasons they share – that is, the reasons that all see as carrying some justificatory weight (even if they disagree about this weight). Convergence liberals, in contrast, hold that whether a law is justified to a reasonable person depends on all their reasons (or at least those meeting a minimal standard of “intelligibility”) – such that a law might be justified to different people for different reasons, or publicly unjustified because some reject it for non-shared reasons.

Public reason liberalism, then, is really a family of theories. We will have more to say about some intramural disputes later, but for now we merely note that for public reason liberalism to remain a viable and distinctive position, it must navigate between two poles. First, since public reason liberals are not anarchists, they must avoid claiming that no laws are justified to all reasonable people. Here, there are two basic strategies to respond to this “empty set” worry: one can restrict the class of people who count as “reasonable,” say, to those who embrace core liberal commitments (Quong 2011, chap. 5), or one can lower the standard of justification by holding, say, that a law is justified to a reasonable person when they see it as “better than nothing” even if highly suboptimal (Gaus 2011, 321–25). Second, since public reason liberals reject the first-order moral approach, they must avoid claiming that laws are only publicly justified when they are justified according to the true moral theory, for example, because the class of “reasonable” people is effectively restricted to those who accept the correct first-order moral theory (or, in practice, the theory favored by the public reason liberal). To avoid this “collapse” worry (Raz 1990, 46), they must allow that different laws can be justified to different reasonable people – so that people don’t

only count as reasonable, say, if they accept the true moral theory, and that a law doesn't only count as justified to someone, say, if that law is justified by the true moral theory.

In this sense, public reason liberalism is a “second-order” theory of justification: a theory of how to justify laws in societies whose populations do not agree about first-order morality. And it claims that, in such societies, governments are prohibited from enforcing publicly unjustified laws.

## 2.2 The Moral Uncertainty Approach

In political philosophy, public justification is arguably the most prominent approach to accommodating moral disagreement. In moral philosophy, a recently popular approach to moral decision-making under disagreement and uncertainty is the *moral uncertainty approach*.

The moral uncertainty approach asks: what morally ought I to do when I don't know what I morally ought to do (Sepielli 2009)? It therefore investigates the moral analogue to decision-making under empirical risk and uncertainty. Suppose you are sympathetic to utilitarianism but not fully certain. You also have some credence in other moral views, such as Kantianism and virtue ethics. Suppose you have three options to choose from – *A*, *B*, and *C* – and that different moral views give you different prescriptions about what to do. The moral uncertainty approach aims to tell you what to do given your uncertainty about the correct moral view.

Consider different ways of filling in the moral uncertainty approach. *My Favorite Theory* (MFT) says that you should simply choose the option implied by the moral view in which you have the highest credence (Gracely 1996; Gustafsson and Torpman 2014). So, if you are uncertain whether utilitarianism is the correct view but you still think it is the moral view most likely to be right, you ought to follow utilitarianism. Although many seem implicitly to adopt this approach, moral uncertainty theorists mostly reject it. One reason is the problem of theory individuation (MacAskill, Bykvist, and Ord 2020, 41–44). Imagine you have 60% credence in consequentialism, 30% credence in Kantianism, and 10% credence in virtue ethics. But there are many forms of consequentialism: for example, you could be a hedonic utilitarian or a prioritarian or a

consequentialist with a richer theory of the good. Say you have equal credence in three different types of consequentialism such that your credence in each of them is 20%. Instead of consequentialism, your favorite theory now turns out to be Kantianism!

An alternative that avoids this problem is *My Favorite Option* (MFO) (Lockhart 2000, 26). MFO says that under moral uncertainty you ought to choose the option that is most likely to be correct. So, to continue the above example, imagine all three versions of consequentialism say you ought to do *A*, Kantianism picks *B*, and virtue ethics picks *C*. MFO tells you to choose *A*, as *A* has a 60% chance of being the right option and is thus the option most likely to be right.

However, MFO still faces the problem that it is stakes-insensitive: the stakes implied by different moral theories sometimes vary greatly, and in such cases, the intuitively correct response to moral uncertainty can be to hedge against moral risk even if that means choosing an option that is not most likely to be right (MacAskill, Bykvist, and Ord 2020, 44–47). For example, imagine again that you have 60% credence in consequentialism but 40% credence in a deontological theory on which killing – even to promote moderately good consequences – is very wrong. This should lead you to refrain from killing when killing has only marginally better consequences than not killing, since a small chance of doing something very wrong (violating a serious deontic constraint) can outweigh a large chance of doing something slightly wrong (producing marginally suboptimal consequences). Since both MFT and MFO consider only the probability of different theories or options being correct but ignore the stakes between different theories, they are unable to accommodate such moral hedging.

For the purposes of this article, we thus assume, along with most moral uncertainty theorists, a stakes-sensitive version of the moral uncertainty approach. Specifically, where greater precision is needed, we assume:

“*Maximize Expected Choiceworthiness (MEC)*: When we can determine the expected choiceworthiness of different options, *A* is an appropriate option iff *A* has the maximal expected choiceworthiness” (MacAskill, Bykvist, and Ord 2020, 48); see also (Lockhart



2000; Ross 2006; Sepielli 2009).

MEC is the moral equivalent to expected utility theory: it sees moral theories as assigning choiceworthiness scores to options, multiplies the choiceworthiness score each theory assigns to an option by our credence in that theory, sums together the ensuing products to determine an option's *expected* choiceworthiness, and instructs us to choose the option with the greatest expected choiceworthiness. Much like expected utility theory, MEC can be adjusted in various ways to account for issues to do with attitudes toward risk and so on (MacAskill, Bykvist, and Ord 2020, 48). More to the point, it is stakes-sensitive and so allows for moral hedging, like in our example above where MEC would likely recommend not killing because of your credence in deontology.

Although MEC is the most popular approach to moral uncertainty, some object that it relies on overly strong assumptions about intertheoretic comparability: it presupposes that we can make intertheoretic “unit” comparisons of the form “the difference in choiceworthiness between *A* and *B* according to deontology is *n* times as great as the difference in choiceworthiness between *A* and *B* according to consequentialism” (Gracely 1996; Gustafsson and Torpman 2014) – for example, because *A* involves killing to bring about slightly better consequences.<sup>4</sup> However, we don't find this objection especially worrying, as we believe proponents of MEC have provided compelling responses (MacAskill, Bykvist, and Ord 2020, chap. 5 is a helpful survey). Moreover, some versions of the moral uncertainty approach capture intuitions about stakes and moral hedging without relying on intertheoretic comparisons, including bargaining approaches which treat moral theories as if they can bargain over what action to perform (Cotton-Barratt and Greaves 2019) and social choice approaches that treat moral theories as if they can cast votes on which action to

---

<sup>4</sup> Valentini and List raise this worry specifically about attempts to connect moral uncertainty to public reason liberalism in the only other discussion of the topic of which we are aware. But they do not develop the criticism, writing only: “We suspect that, contrary to what some recent literature on moral uncertainty suggests, [intertheoretic comparisons] pose insurmountable challenges” (Valentini and List 2020, 203–4).

perform (MacAskill 2016; Tarsney 2019).

In further work, it would be interesting to consider whether the case for public justification is stronger or weaker given different approaches to moral uncertainty or assumptions about intertheoretic comparability. But to keep things manageable, we stick with MEC here.

### **2.3 From Moral Uncertainty to Public Justification?**

The moral uncertainty approach has so far been primarily employed in applied ethics. Recently, however, there have been a few attempts to apply it to political philosophy (Bukoski 2015; C. Barry and Tomlin 2019). We welcome these developments and believe there is room for additional interesting work at the intersection of moral uncertainty and political philosophy. Here, we focus on public reason liberalism. There is good reason for public reason liberals to take an interest in, and draw on, the moral uncertainty approach.

First, as we have noted, the moral uncertainty approach and public reason liberalism share various concerns and points of contact: in moral matters, it is hard to know – and so we disagree about – which theory is correct, but we must at the end of the day still act. Public reason liberalism says that simply appealing to our favorite first-order moral or political theory won't do. Similarly, the moral uncertainty approach says that appealing to one's Favorite Theory (or Option) in ethics won't do. Moreover, both approaches suggest that we should take seriously other reasonable people's moral views. Specifically, public reason liberals are explicitly concerned with accommodating moral disagreement, and moral uncertainty theorists identify disagreement as one central reason to be morally uncertain (MacAskill, Bykvist, and Ord 2020, 12–13).

Second, public reason liberals are often confronted with a challenge: why should we accept the public justification principle? A range of plausible answers are on offer. However, critics argue that such answers seem to rely on, or smuggle in, controversial first-order moral theories themselves (Enoch 2013; Wall 2002). For example, some argue that respect or the ideal of political community can ground the public justification principle. But is there not reasonable disagreement

about these very ideals? And, if so, why isn't public reason liberalism "but another sectarian doctrine" (Rawls 1985, 246) – a first-order theory in competition with other first-order theories like utilitarianism and Kantianism, rather than a second-order theory that purports to stand above the fray? Others take a different tack and draw a stark epistemological distinction between conceptions of the good and morality on the one hand and principles of justice and institutional justification on the other (B. Barry 1996, 169–71; Nagel 1987). The former but not the latter, they argue, lack sufficient epistemic status to be used in justifications of state coercion. However, it is far from obvious that theories of justice or institutional justification are indeed epistemically stronger or more robust than views on morality and the good life (Clarke 1999; Enoch 2017). They, too, are subject to widespread disagreement and uncertainty.

Of course, public reason liberals respond to these worries (Bajaj 2017; Beshpalov 2021). Be that as it may. Our point is simply that moral uncertainty might provide a welcome and underexplored route to public reason liberalism: the moral uncertainty approach allows for reasonable disagreement and uncertainty all the way down.<sup>5</sup> So, it does not, for example, require a firm commitment to any first-order theory of morality or justice. Moreover, it does not hold that political principles have a stronger or more robust epistemic status than ethical views and conceptions of the good. Instead, it includes moral uncertainty about first-order political philosophy too.

Finally, we will see that the moral uncertainty approach provides new answers to some intramural debates among public reason liberals about how best to formulate the public justification principle. That such details can be filled out in many different ways and that public reason liberals disagree about them is itself sometimes turned into an objection to public reason

---

<sup>5</sup> Well, maybe not quite *all* the way down: there is still the question of how to deal with uncertainty about which approach to moral uncertainty to use (Weatherson 2014). For responses to this "regress" problem see (MacAskill, Bykvist, and Ord 2020, 33; Trammell 2019).

liberalism: the public justification principle purports to tell us how to proceed in the face of reasonable disagreement, but then how are we to handle reasonable disagreement about the proper formulation of the public justification principle (compare (D’Agostino 1996))? The moral uncertainty approach’s ability to shed new light on intramural disputes in question is therefore not only independently interesting, but may also help public reason liberals to ward off this concern.

It is therefore worth investigating whether moral uncertainty offers support for public reason liberalism. To this end, we consider:

*Hypothesis*: we MU-ought to design political institutions such that governments adhere to the public justification principle.

Note two features of this hypothesis.

First, we use “MU-ought” to refer to the “ought of moral uncertainty” or “what we ought to do according to the moral uncertainty approach”. Distinguish this from the “objective ought” of what we ought to do according to the correct moral theory. So, in our earlier example, the moral uncertainty approach implies that a person with credences across consequentialism, Kantianism, and virtue ethics sometimes MU-ought to pick a different option from what they ought to pick on the objectively true theory (when they engage in moral hedging for example).<sup>6</sup>

Second, our hypothesis is subtly different from:

*Hypothesis\**: governments MU-ought to adhere to the public justification principle in every instance.

An analogy might help. Suppose you are a utilitarian and think that under specific and unlikely circumstances, it can be morally right to torture someone (say, to save lots of lives). Still, you might think that we are better off designing institutions that never legally countenance torture, because

---

<sup>6</sup> Some are skeptical that the MU-ought exists, e.g. (Weatherson 2014). For responses, see (MacAskill, Bykvist, and Ord 2020, 18–21; Sepielli 2013).

you think the probability of justified torture are so minute, and the relative risk of unjustified torture so high, that a strict legal prohibition on torture far outweighs its downsides. Imagine now that instead of a committed utilitarian, you are morally uncertain about the correct first-order theory with credences distributed over utilitarianism, deontology, and so on. Even if many theories say that torture is always wrong, since you put some credence in utilitarianism, you might hold that in rare circumstances, we MU-ought to torture someone (say, to save an enormous number of lives). However, you might again think these circumstances are so rare and the relative risk of unjustified torture so high that, on balance, we MU-ought to adopt an institutional principle prohibiting torture. So, we MU-ought to design political institutions that make no legal room for torture, even though there are some cases where governments MU-ought to torture.

Now, coming back to public justification, the moral uncertainty approach might similarly give you reason to design institutions that live up to public justification even if you think that, in some instances, governments MU-ought to depart from public justification. Put in terms of MEC: even if public justification sometimes leads to decisions that do not maximize expected choiceworthiness, it might still maximize expected choiceworthiness to design institutions that adhere to public justification. So Hypothesis might be true even if Hypothesis\* is not.

Before discussing arguments for Hypothesis, we must dispense with one more ambiguity. Public justification is about institutional justification. The moral uncertainty approach, in contrast, is about what an *individual* ought to do. So, the moral uncertainty approach might not be a natural framework for political philosophy. We see, broadly, three ways of using moral uncertainty in political philosophy.

First, we might address collective agents. So, when a government or society needs to decide on a law, we treat the relevant collective as having credences in different moral views. Here we could treat people's beliefs and how they are represented in a population as analogous to an individuals' subjective probabilities. However, such a view encounters a serious challenge: how should we aggregate individuals' beliefs into collective credences? Furthermore, whose beliefs

should we aggregate: all people in society, all reasonable people, or only parties to the decision (say, government officials)? Perhaps good answers are available, but we will not pursue them here.

Second, we could use the moral uncertainty approach to justify institutions *to a particular person* who is morally uncertain. So, imagine you are somewhat uncertain about what the correct moral theory is. We could then use the moral uncertainty approach to convince you that, given your own credences in different moral theories, you should accept public justification as an institutional principle. We might then repeat this process and seek to justify public justification to others who are also morally uncertain in different ways.

Third, instead of going for *de facto* distributions of moral beliefs in a population or trying to justify public justification to any one individual, we might claim that there are credences over different moral views that are rational to have. We need not assume that there is one rational probability distribution over moral views, but could assume a range: for example, perhaps it is irrational to have a very high credence in a deeply racist view but rational to have a wide range of non-negligible credences in deontology, libertarianism, or consequentialism. It is then based on these rational probability distributions that we construct arguments in political philosophy.

The second and third approach above track a disagreement among moral uncertainty theorists about whether what agents MU-ought to do depends on their actual credences or their “epistemic” credences (the credences it would be epistemically rational to have), which we cannot hope to resolve here (MacAskill, Bykvist, and Ord 2020, 4). If one prefers the “actual credence” view then our ensuing arguments can be interpreted as claiming that agents who are in fact substantially uncertain about morality should favor public justification, at least if they allot their credences among the sort of views that are generally seen as contenders in moral and political philosophy. If one prefers the “epistemic credence” view then our arguments can be interpreted more broadly as suggesting that rational agents should favor public justification.

Note that this latter view does *not* commit us to thinking that reasonable people (of the sort referenced by the public justification principle) have credences in line with the set of rational

credences. Indeed, as we will explain, we do not assume that all reasonable people are themselves morally uncertain nor that they endorse MEC. Our hypothesis is therefore not that all reasonable people must accept the moral uncertainty approach, such that all should accept that we MU-ought to design political institutions that adhere to public justification. It is rather that the moral uncertainty approach implies that we MU-ought to accept public justification as an institutional principle, regardless of whether reasonable people endorse the approach.

Let us now explore arguments for our hypothesis.

### **3 The Proto-Moral Uncertainty Argument**

We begin with an argument briefly suggested, in incipient form, by two of public reason liberalism's best-known proponents: Rawls and Gaus (Gaus 2015, 1085; Rawls 2005, 125). The argument – which we call the “proto-moral uncertainty argument” – goes like this. In general, even when our favored moral view suggests that we ought to enforce a law, there is a considerable chance that this is wrong. A government that enforces laws according to our (or its) favored moral theory, therefore, runs a serious moral risk. Adopting the public justification principle is an excellent way of reducing this risk. After all, if a government is prohibited from enforcing laws that some reasonable person rejects, then if, for any given law under consideration, at least one reasonable person gets the objectively correct answer about whether this law is justified, the government will only enforce laws that are justified according to the correct moral theory. Publicly justified laws are therefore “safe” options; publicly unjustified laws are not. To the extent that it is plausible that at least one reasonable person will have the correct moral opinion on any given law, it is plausible that requiring public justification will ensure we never enforce laws that are objectively unjustified – that is, unjustified according to the correct moral theory.

The proto-moral uncertainty argument shows that adopting the public justification principle is in one respect desirable from the perspective of moral uncertainty: it reduces the

probability that we adopt laws that are unjustified according to whatever moral theory turns out to be true. Of course, objectively unjustified laws may slip through the cracks of the public justification test if no reasonable person has the correct moral view on some issue. But we can at least remain content knowing that, barring simply imposing a blanket prohibition on enforcing laws, we could not have realistically reduced this risk any further.

However, this argument has two problems.

First, the proto-moral uncertainty argument assumes a strong asymmetry between wrongly enforcing laws and wrongly failing to enforce them. Let us say that when a government enforces a law that is not justified according to the true moral theory, this is a “moral false positive”. When a government fails to enforce a law that is justified according to the true moral theory, call this a “moral false negative”. Instituting the public justification principle reduces the risk of moral false positives, but it also increases the risk of moral false negatives: some laws that fail a public justification test will be objectively justified since, for at least some controversial laws, the side believing we ought to enforce the law will be correct. So, the proto-moral uncertainty argument gives far greater weight to avoiding false positives than avoiding false negatives, and we need a justification for this asymmetry.

Second, the proto-moral uncertainty argument foregrounds the probability of avoiding laws that are objectively unjustified but fails to be stake-sensitive. Its reasoning is reminiscent of MFO where stress is laid on minimizing the probability that we choose an option that is objectively wrong (or maximizing the probability we choose what is objectively right). However, we saw above that plausible approaches to moral uncertainty focus not only on probabilities but also on stakes. As our example of killing to promote better consequences showed, we sometimes MU-ought to hedge and prefer an option likely to be slightly wrong over an option with a lesser chance of being very wrong. So, arguments for our hypothesis cannot be based only on the probability of moral false positives or negatives. We must also consider the severity or stakes of such errors.

The Proto-Moral Uncertainty Argument therefore will not do. We require reasons to think



that the public justification principle not merely reduces the risk of moral false positives but appropriately balances the risks of false positives against the risks of false negatives while also being sensitive to the stakes of each. We now propose four such reasons.

#### **4 Four Considerations Favoring Public Justification**

As we have seen, when balancing the risks of moral false positive against the risk of moral false negatives, we must consider two basic factors. First, we must consider the *probability* of either risk accruing. As the likelihood that a law is objectively justified decreases, the risk of a moral false positive goes up and the risk of a moral false negative goes down. Second, we must consider the *severity* of each sort of moral error. The morally worse it would be to enforce some law that is objectively unjustified, the more severe is the risk of a false positive; the morally worse it would be to fail to enforce some law that is objectively justified, the more severe the risk of a false negative.

Let us now say that a “publicly unjustified false positive” occurs when a publicly unjustified law is enforced when it objectively should not be (it is a moral false positive), and that a “publicly unjustified false negative” occurs when a publicly unjustified law is not enforced but objectively should be (it is a moral false negative). To defend our hypothesis we must come up with some considerations that either drive up the *severity* of publicly unjustified false positives relative to the severity of publicly unjustified false negatives, or that increase the relative *probability* of publicly unjustified false positives. Specifically, these two types of considerations must be strong enough to show that refraining from enforcing a publicly unjustified law typically has greater expected choiceworthiness than enforcing it. Once again, some exceptions may occur, as we are concerned with public justification as an institutional rather than an exceptionless principle. Still, such exceptions should be rare and not so severe that they significantly undermine the expected choiceworthiness of implementing the public justification principle.

Here is a consideration of the second type:

*Public Justification Tracks Objective Justification.* When a law is publicly justified, it has a high chance of being objectively justified. When a law is publicly unjustified, there is at least a significant chance that it is not objectively justified.

The first half of this claim is from the proto-moral uncertainty argument above. The second half does not follow immediately from the first, but we think it is quite plausible. If some reasonable people believe that a law is not justified, then we should take the possibility that the law is not justified seriously. In fact, we can guarantee this result by defining a “reasonable” person in a way that makes it true: when a reasonable person believes a law is not justified, then there is a significant chance they are right. This will hold if we define a reasonable person with respect to some law either as someone who holds a credible view with respect to that law or as someone whose testimony that a law is unjustified we should take seriously, in the sense that if they deny a law is justified, we should think there is a significant chance the law is objectively wrong. Alternatively, rather than defining a reasonable person this way, we might simply note that our intuitive notion of a reasonable person overlaps quite considerably with these categories: when someone disagrees with us about something, and we think this disagreement is reasonable, we believe there is at least a significant chance they are right. So *Public Justification Tracks Objective Justification* may hold either definitionally, or at least as a general tendency. And it implies that, when a law is publicly unjustified, we should think that there is a significant chance the law is objectively wrong.

By itself, this consideration doesn’t get us very far. It gets our foot in the door by suggesting that the probability of publicly unjustified false positives is worth taking seriously. But it implies nothing about either the severity of publicly unjustified false positives or even about how probable they are in relation to publicly unjustified false negatives. A second consideration helps us further open the door to public justification:

*Public Reason Liberals Might be Right.* Public reason liberals offer several reasons why laws, in virtue of being publicly unjustified, are objectively unjustified. Under moral uncertainty, we

should give some weight to these reasons.

Public reason liberals defend the public justification principle in several ways. They propose various deontic reasons suggesting that publicly unjustified laws are wrong, for example, because they are disrespectful (Larmore 1990) or authoritarian (Gaus 2011). And they propose various axiological reasons, suggesting that publicly unjustified laws undermine social trust (Vallier 2019), political community (Leland and Wietmarschen 2017; Lister 2013), or a morally attractive notion of stability (Rawls 2005). Relatedly, they argue that publicly unjustified laws are less instrumentally effective at securing whatever they aim to achieve, since, all else being equal, they are less stable, more likely to generate resistance and backlash, and so less predictable (Barrett and Gaus 2020). If we are morally uncertain, we presumably have some credence in these views, and so give some credence to the idea that, simply in virtue of being publicly unjustified, a law is wrong.

This consideration has two effects. First, much like the first consideration, it drives up the probability of publicly unjustified false positives. To the extent that we think public reason liberalism might give us the right first-order theory of when we should implement laws, we should think it less likely that we should objectively enforce any given publicly unjustified law. Second, and more subtly, it drives down the severity of false negatives. For even if a law is all-things-considered objectively justified despite being publicly unjustified, to the extent that we put credence in the various considerations public reason liberals marshal against laws that are publicly unjustified, we should think that the law is less choiceworthy (in expectation) than it would otherwise be, because it carries some moral risk – for example, of being disrespectful or having bad effects.

*Public Reason Liberals Might be Right* gets a grip because of the role asymmetries play under moral uncertainty (MacAskill, Bykvist, and Ord 2020, 183-187). Some theories claim that public justification is in itself valuable; others hold that public justification doesn't matter. But no plausible theory claims that public justification is in itself disvaluable. So, under moral uncertainty, we should treat it as *somenwhat* valuable. Our third consideration concerns a similar, more general asymmetry:

*There May be a Presumption Against Coercive Laws.* There are many plausible moral theories on which there is a strong presumption against coercive laws, and no plausible moral theory on which there is the reverse presumption. So under moral uncertainty, there is a weak presumption against coercive laws.

This consideration picks up on a common refrain of public reason liberals – namely, that coercion stands in special need of justification, in a way that failures to coerce don't – that serves as a premise in many arguments for the public justification principle (Feinberg 1984, 9; Rawls 2001, 44; Gaus 2011, 319–21). In the current context, its relevance is the following. We are after reasons to believe that, when it comes to publicly unjustified laws, the risk of a moral false positive at least typically outweighs the risk of a moral false negative (and that exceptions to this rule tend to involve only small losses in expected choiceworthiness). *There May be a Presumption Against Coercive Laws* supports this claim because it drives up the severity of the risk of moral false positives relative to the risk of moral false negatives in general – and not just when it comes to publicly unjustified laws. Specifically, on many moral theories, coercion is in itself always *pro tanto* wrong or unjustified coercion is in itself very wrong. Some theories deny this (e.g. (Wall 2010)), but no plausible moral theories claim that failures of coercion are, in virtue of being failures to coerce, wrong. And similarly, on many moral theories, wrongful acts are, everything else being equal, worse than wrongful omissions. Some theorists, utilitarians for example, deny this, but none claim that wrongful omissions are worse than wrongful acts. Since enforcing a law is a coercive act, and failing to enforce a law is a non-coercive omission, these asymmetries suggest that, under moral uncertainty, there are additional factors contributing to the severity of false positives but not false negatives. So this gives us further reasons to err on the side of not enforcing laws in general, and especially on the side of not enforcing publicly unjustified laws which, by *Public Justification Tracks Objective Justification* and *Public Reason Liberals May Be Right* have a good chance of being wrong.

The above three considerations are of the type we were looking for: they drive up the

probability and severity of publicly unjustified false positives relative to the probability and severity of publicly unjustified false negatives. However, we doubt that these considerations are strong enough to confirm our hypothesis unless one happens to assign quite high credences to public reason liberalism (or the values it relies on), the presumption against coercion, or the act-omission distinction. We therefore raise a fourth consideration which applies only to some conceptions of public justification and gives them a significant boost under moral uncertainty:

*Public Justification Is a Low Bar.* On many conceptions of public justification, a law is justified to a reasonable person even if they see it as highly suboptimal. It must only meet some low threshold of being “better than nothing” or “something they can live with.”

Public reason liberals typically view public justification as a standard short of optimality. A law doesn’t count as publicly justified to reasonable people only if everyone sees it as a first choice. Rather, all must see law as merely “good enough” (Gaus 2011; Vallier 2019).

Now, the lower the bar one sets here – is a law justified to a reasonable person when they see it as pretty good, better than nothing, or merely non-disastrous? – the more severe the risk of publicly unjustified false positives becomes. For the lower the bar of public justification, the more severe the risk of enforcing a publicly unjustified law: if a law fails to meet a high bar of public justification because some reasonable people see the law as slightly wrong, there is a moral risk of enforcing a slightly wrong law, but if a law fails a low bar of public justification because some reasonable people see it as extremely wrong, then there is a moral risk of enforcing an extremely wrong law. Furthermore, the lower we set the bar of public justification, the fewer laws will fail the public justification test in general – so the lower the probability of publicly unjustified false negatives. If we set the bar quite low, many laws will be publicly justified, and the remaining laws that aren’t will (by *Public Justification Tracks Objective Justification*) have a serious chance of being severely wrong.

Lowering the bar also interacts with *Public Reason Liberals Might be Right*, since most of the

substantive considerations that public reason liberals raise against publicly unjustified laws seem to scale with where we set the bar. For example, if it is disrespectful or authoritarian to enforce a law when people object to this law, then this is presumably more disrespectful or authoritarian the stronger their objection. Similarly, if enforcing publicly unjustified laws undermines social trust or stability since individuals are less willing to comply with laws they do not endorse, this effect is presumably stronger when people more strongly object to the law in question. Lowering the bar therefore magnifies the two effects we earlier mentioned relating to *Public Reason Liberals Might be Right*, further increasing the probability and severity of publicly unjustified false positives relative to the probability and severity of publicly unjustified false negatives.

Once we lower the bar, however, we need to slightly reinterpret *Public Justification Tracks Objective Justification*. Rather than holding that publicly justified laws are very likely to be objectively justified in the sense that they are morally optimal or permissible, we must now hold that publicly justified laws are very likely to meet whatever the relevant bar of justification is (for example, being better than nothing). Yet this is all to the good since it makes room for the sort of moral hedging that plausible versions of the moral uncertainty approach recommend. To illustrate this, consider two reasonable people with a choice between different laws on a particular subject:

	Moral assessment of reasonable person 1	Moral assessment of reasonable person 2
No law	Disastrous	Disastrous
Law 1	OK	OK
Law 2	Great	Disastrous
Law 3	Disastrous	Great

If a law is justified to reasonable people only when that law is their first choice, then the public justification principle would generate a disastrous outcome: the government being required to select “no law.” However, with a lower bar, Law 1 would count as justified to both reasonable people, such that the government would be permitted to enforce Law 1 (but not Laws 2 or 3).

Now, assume – for simplicity – that both people’s moral views represent views in which we have (rational) credences of roughly 50%. MEC would likely tell us to hedge and choose Law 1, even though it is not the morally best option. And, more generally, a lower bar permits governments to engage in moral hedging in a way a higher bar rules out. This is an interesting connection between moral hedging and public justification. Moreover, it helps to illuminate the way that *Public Justification is a Low Bar* reduces the probability of publicly unjustified false negatives and the severity of publicly unjustified false positives: fewer laws will be ruled out by a lower bar and those that are ruled out will carry more severe risks.

Altogether, then, this leaves us with the following picture. If a law is publicly unjustified, then by *Public Justification Tracks Objective Justification*, the probability that it would be a moral error to implement it is significant: when reasonable people think a law is wrong there is a significant chance they are right. *Public Reason Liberals Might be Right* has the same effect of driving up the probability of publicly unjustified false positives, but also decreases the severity of publicly unjustified false negatives: to the extent that we put credence in public reason liberalism, or the considerations public reason liberals adduce, we should think it less likely that publicly unjustified laws are objectively justified and should see publicly unjustified laws as carrying important moral costs. As a general matter, *There May be a Presumption Against Coercive Laws* tips the scales against enforcing laws by increasing the severity of the risk of false positives without similarly raising the severity of false negatives, since, under moral uncertainty, there is an asymmetry between the (greater) risk of coercive actions and the (lesser) risk of non-coercive omissions. Finally, by *Public Justification is a Low Bar*, the lower the bar we set for a law counting as justified to reasonable people, the greater these various considerations are magnified: the probability and severity of false positives goes significantly up, and the probability and severity of false negatives goes significantly down, if we think of a law needing to meet a less stringent bar of, say, being something that everyone can live with, rather than a law that everyone sees as optimal.

## 5 Are These Considerations Enough?

The above four considerations support our hypothesis by significantly increasing the risk of publicly unjustified false positives relative to publicly unjustified false negatives. But are they enough to ground public justification as an institutional principle? Do they suggest that refraining from enforcing a publicly unjustified law generally has greater expected choiceworthiness than enforcing it, and that exceptions to this involve only mild losses of expected choiceworthiness?

On balance, it seems plausible that if we set the bar of public justification low enough, governments will typically do what they MU-ought when they refrain from enforcing publicly justified laws. Under moral uncertainty, it is generally a good idea to avoid options that carry significant chances of very bad outcomes, which publicly unjustified laws always carry (by *Public Justification Tracks Objective Justification*, *Public Reason Liberals Might be Right*, and *Public Justification is a Low Bar*), and this is especially so when comparing option-types that carry asymmetrical moral risks – in this case the asymmetry between coercive actions and non-coercive omissions (by *There May Be a Presumption Against Coercive Laws*). However, exceptions to this general tendency may occur when failing to enforce a publicly unjustified law also carries a significant risk of producing a very bad outcome. For example, imagine some reasonable people believe a law unjustified, but others think *failing* to enforce it would be a moral disaster. In such cases, if we assign greater credence to the view that failing to enforce the law would be a disaster, or if we think failing to enforce it would lead to a graver moral error if it turned out to be one, it may maximize expected choiceworthiness to enforce the publicly unjustified law.

But how often will such cases arise? At first glance, they might seem rare. For consider:

*Laws that Prevent Moral Disasters are Typically Publicly Justified.* At least typically, if failing to enforce a law would be very morally bad, then reasonable people will recognize this, and agree that the law is justified (at least given a low bar of justification).

This consideration is *prima facie* intuitive. It follows from the basic idea that when failing to enforce a law would be very bad, reasonable people will tend to converge on this perspective – or at least,



will tend not to get things so backward that they think enforcing the law would be bad enough that it fails to meet a low bar. When something is very bad, it tends to be bad for many reasons, and so from any reasonable perspective. For example, no reasonable person rejects laws needed to maintain basic order in society and to avoid widespread carnage, death, and destruction. Thus, to the extent that this consideration holds, it drives down the probability of publicly unjustified false negatives that are severely bad. And this may be enough to save public justification as an institutional principle, since it suggests that, even if there are some cases where it would maximize expected choiceworthiness to enforce a publicly unjustified law, failing to enforce such laws will not decrease expected choiceworthiness by a large margin.

Unfortunately, this consideration does not seem to hold with much generality. There are obvious counterexamples where a law is publicly unjustified but there is a serious chance that failing to enforce it would be morally disastrous. Suppose you are confident that animal suffering matters and that imposing a tax on meat consumption would reduce animal suffering by a huge amount, and that even if you are wrong about this and the law turns out to be unjustified, there is relatively little moral cost to enforcing it. Yet some reasonable people reject this tax because they believe that animals have no moral status and that the imposition of such a tax involves unbearable moralizing. In this case, the risk of a publicly unjustified false negative may outweigh the risk of a publicly unjustified false positive. A government that adheres to the public justification principle might err, quite severely, in failing to enforce the law.

Can we avoid this result if we make justification a low enough bar that a tax on meat consumption will clear it even for those who deny animals have moral status? Perhaps – but this will not get us out of the problem. Imagine a series of progressively more restrictive laws that each serve to reduce animal suffering. At some point, the law will get restrictive enough that reasonable people who deny the moral importance of animal suffering will reject the law as passing even a minimal bar of justification. But if the law reduces animal suffering by a huge amount, and we have a high enough credence that animal suffering matters, we will still see the law as much more

choiceworthy than its absence – the objections of some reasonable people notwithstanding.

In cases like this, we may simply have to take a stand and insist that we MU-ought to enforce a law that is publicly unjustified, because the moral stakes of failing to do so are too high. Now, if such cases were few and far between, this might still leave intact the hypothesis that we should treat public justification as an institutional principle. But we worry that such examples are not isolated and instead represent a structural or systematic problem. Specifically, there appears to be at least one predictable sort of case where *Laws that Prevent Disasters are Typically Publicly Justified* fails, such that the public justification principle may lead governments to err by quite large margins. These are cases where a reasonable disagreement comes down not to a difference in how reasonable people weigh conflicting considerations but rather to a fundamental and high-stakes disagreement about whether some consideration matters at all.

For example, consider that there are also deep disagreements about whether laws should be justified by reference to their effects on future generations, including those in the far future, or whether states have obligations towards faraway citizens in other countries. Such disagreements are particularly troubling for our hypothesis, because such problems are *high-stakes* problems: if non-human animals, humans in other countries, or far-future people matter, failing to take them into account would create a grave moral risk, given the gigantic number of them and given that many moral theories are scale-sensitive. Accordingly, if we apportion significant credence to scale-sensitive theories that assign moral status to such beings, these theories will be very influential in moral uncertainty calculations. So, if there are reasonable people who deny that such considerations matter to the justification of laws, then insisting on public justification will sometimes require us to incur particularly large moral risks.

This makes us skeptical: even as an institutional principle, the moral uncertainty approach might not support public justification. There is a predictable sort of case where governments that adhere to public justification will act in ways that are very low in expected choiceworthiness.

Let us explore two responses.

## 5.1 Constraining Reasonableness

A public reason liberal might respond by stipulating that someone who denies, say, that animals or far-future people have moral status is unreasonable. By considering such people unreasonable, and then continuing to narrow our understanding of reasonableness as the case demands, we could eliminate all the high-stakes counterexamples to public justification.

However, we do not advocate this solution.

First, such an understanding would not chime with the moral uncertainty approach, which, we have suggested, should interpret people who disagree with us as “reasonable” when we think there is at least a significant chance they are right – at least if the all-important consideration *Public Justification Tracks Objective Justification* is to be preserved. Again, we do not insist that reasonable people be defined any particular way, but merely note that, if public reason liberals want to invoke moral uncertainty in their favor, there must be a large overlap between people we consider reasonable and those we think have a significant chance of being right. But this still means the set of reasonable people will be quite wide. Importantly, it will include those who believe that animals lack moral status (or at least that minor losses to human welfare trump much larger losses in animal welfare) or who believe versions of the person-affecting view that entail that we do not have obligations towards far-future people. Indeed, such views are not only held by regular citizens but are serious positions in academic debates.

A potentially more promising response might be to hold that “reasonable” can allow for people who believe animals do not have moral status but cannot allow for people who believe this *with certainty*. That is, reasonable people must themselves be morally uncertain to some extent and laws need only be justified to people who themselves are somewhat morally uncertain. Of course, we could then only require that reasonable people be “somewhat uncertain” not that they are reasonable only if they fall within some range of rational credences over different moral views. This move could be used in defense of our hypothesis: any reasonable person will have some moral uncertainty over whether animals have moral standing. This implies they should largely be in

support of animal welfare legislation, seeing that the vast scale of non-human animals drives up the moral risk that comes about if we continue to allow animal suffering in factory farms. For example, imagine that you only have a 0.05 credence that it is bad to raise and kill non-human animals in factory farms. You should still find factory farming disastrous in expected moral choiceworthiness, as it kills more than 100 billion animals a year.

This is a possible move. But we doubt that public reason liberals should make it. That reasonable people should be somewhat morally uncertain is a controversial but perhaps plausible interpretation of Rawls's idea that "reasonable persons recognize and accept the consequences of the burdens of judgment" (Rawls 2005, 488). However, the problem is that this move furthermore presupposes a particular picture of moral uncertainty—namely, that reasonable people must employ MEC (or at least another stakes-sensitive approach). Imagine someone instead adopts MFT as their response to moral uncertainty, and so maintains that, since they think animals most likely don't matter, we MU-ought not to take their interests into account. Would such a person be unreasonable? To us, it again seems more plausible to say they are reasonable if their favorite theory has some significant chance of being true. Of course, we could stipulatively define reasonableness so that only people who endorse MEC are included, but this would rule out so many people as to render the public justification principle toothless. As we have noted, many people seem implicitly to employ something like MFT and, indeed, even moral philosophers have only very recently begun to consider alternatives. We do not want our principle of public justification to include in the "public" only a small minority of people, primarily composed of academic philosophers (compare (Valentini and List 2020, 204)).

These arguments are inconclusive, and perhaps someone else can find good reasons for defining reasonableness in a way that implies some level of moral uncertainty and that one employs MEC (or at least a stake-sensitive approach to moral uncertainty). This notion of reasonableness could then be used to avoid problematic cases. However, we find another route more promising.

## 5.2 Going Pro Tanto

The above proposal sought to avoid problematic cases by narrowing the set of reasonable people. The alternative is to keep this set quite broad but weaken the public justification principle. When formulating our hypothesis, we left this somewhat open:

*Hypothesis:* we MU-ought to design political institutions such that governments adhere to the public justification principle.

Here, the notion of “adherence” is ambiguous, and perhaps we could rescue the hypothesis by interpreting it less stringently:

*The Pro Tanto Interpretation.* Governments adhere to the public justification principle when they treat it as a weighty *pro tanto* institutional commitment that can be overridden by other weightier concerns (compare (Wendt 2019)).

We could then construe this interpretation in one of two ways: we could view the criterion as a binary criterion such that a law is either publicly justified or not and when it is not that counts *pro tanto* against enforcing that law. Or we could view public justification as scalar such that a law can be publicly justified to different degrees.

Although either approach could likely avoid the problem we have raised, we believe that the scalar approach is especially plausible. On this view, the extent to which a law is publicly justified depends on several variables. Most obviously, it depends on what bar of public justification a law achieves. For reasons that are now familiar, if a law fails a very low bar of public justification (say, because some reasonable people think that enforcing it would be a disaster) there is a stronger reason not to enforce that law than if it only fails to meet a relatively high bar (say, because some reasonable people think it is only slightly suboptimal). So, we may say that laws are publicly justified to a lower degree if reasonable people have stronger objections to them and that, in such cases, the *pro tanto* institutional commitment to public justification should be stronger.

We might also distinguish between laws that are *fully* publicly justified in the sense that they are justified to every reasonable person, and laws that *approximate* public justification in the sense that they are justified, say, to an overwhelming majority (Barrett and Gaus 2020, 224). We could then say that laws that are justified to more reasonable people are publicly justified to a greater degree, and that the pro tanto institutional prohibition on enforcing publicly unjustified laws should vary accordingly. The basic thought here is that, just as laws that fail a lower bar of public justification are less likely to be laws that governments MU-ought to enforce, laws that fail to hit this bar for a wider class of people are less likely to be laws that governments MU-ought to enforce. This is for two reasons.

The first relates to *Public Justification Tracks Objective Justification*. At least as a general tendency, it seems plausible that when more reasonable people object to a law, the law is less likely to be objectively justified (especially when their judgments are more independent). Finding out that most people think a law is unjustified should shake our confidence that the law is justified more than finding out that a small minority object, which should shake our confidence more than finding out that a single reasonable person objects. So if a law fails to even approximate public justification – if it is not justified to a large majority – the risk of publicly unjustified false positives is higher than if it only fails to be fully justified.

Second, most of the substantive considerations that public reason liberals raise against public justification seem to scale not only with the severity by which the law fails the public justification test, but also with the number of people who object to a law. For example, if it is disrespectful or authoritarian to enforce laws against people who don't accept them, or if this undermines social trust or stability, then presumably this is a larger problem if more people object to the law. The effects of *Public Reason Liberals Might be Right* are therefore greater for laws that are not even approximately publicly justified.

On the whole, then, we find that adopting public justification as an institutional principle will generally lead governments to make decisions that appropriately balance moral risks. However,

there are a predictable class of exceptions (such as those involving animal welfare or far-future people) where governments failing to enforce a publicly unjustified law carries a grave moral risk, and this calls into question the plausibility of our hypothesis that we MU-ought to design institutions that require governments to adhere to the public justification principle. To rescue our hypothesis, we have suggested weakening the public justification principle to a pro tanto institutional requirement that typically prohibits governments from enforcing publicly unjustified laws, but which can be overridden in the high-stakes cases that otherwise make trouble for our stronger hypothesis. This pro tanto requirement can be given either a binary interpretation or – more interestingly – a scalar interpretation. On the latter, the pro tanto institutional commitment to refrain from enforcing publicly unjustified laws should be stronger the more reasonable people who object to it and the stronger their objections.

## 6 Intramural Debates

We have focused on whether the moral uncertainty approach can vindicate public reason liberalism. In this section, we briefly survey how our discussion has shed light on some intramural disputes among public reason liberals. One such dispute, which we have already addressed at some length, is which people to count as “reasonable”. From the perspective of moral uncertainty, we have seen that the category of reasonable people must at least significantly overlap with the category of people who have a decent chance of being right about that law – either because we (rationally) assign significant credence to their moral view, or because we take them to be morally reliable and so take their testimony seriously.<sup>7</sup> A central advantage of this notion of reasonableness is that it is independently motivated and avoids smuggling in so much specific normative content that it collapses public reason liberalism into a variant of the first-order moral approach.

---

<sup>7</sup> Plausibly, we should also count someone as reasonable if they are themselves morally uncertain and make judgements approximating MEC, though we have argued against construing this as a necessary condition on being “reasonable.”

Moral uncertainty also has implications for the two other intramural disputes we mentioned earlier. First, one dispute concerns when a law qualifies as justified to a reasonable person. Here, we have suggested that the importance of public justification is higher the lower we set the bar. If a law counts as unjustified to some person when they merely see it as somewhat suboptimal, then the case for not enforcing a publicly unjustified law is weaker than in the case where a reasonable person sees the law as morally terrible. Relatedly, although we are less aware of discussion on this point, we have suggested that public justification might be more weighty, under moral uncertainty, if we interpret the principle as less than perfectly stringent – as requiring a law to be justified not to each and every reasonable person, but approximately, say, to an overwhelming majority. In cases where only one reasonable person objects to a law, the case against that law is weaker than in a case where many reasonable people object.

Finally, recall the debate between “consensus” liberals who hold that public justification should proceed by reference only to reasons that all reasonable people share and “convergence” liberals who hold that we should factor in reasons that some reasonable people deny carry justificatory weight. Although we have not said much about shared reasons in this article, it is worth noting that, from the perspective of moral uncertainty, there appears no reason to endorse the shared reason requirement. To factor in moral considerations only when all reasonable people share them is effectively to treat non-shared considerations as if we have zero credence in their moral importance, which seems antithetical to the moral uncertainty approach. Furthermore, in many cases, appealing to non-shared reasons may bring additional epistemic benefits, since judgments reflecting non-shared reasons may often be more independent.

In fact, our above discussion suggests that the public justification principle may be most likely to lead us astray when people don’t share reasons, since this is, plausibly, the most common sort of case where reasonable people will reject a law that we have high credence would be very bad not to enforce. For example, not enforcing a law might severely harm a class of beings (animals, far-future people, or so on) that some reasonable people see as lacking status such that their harm



does not create a shared or public reason. If this is so, then while consensus theorists rightly point to the importance of the distinction between shared and non-shared reasons, they do not locate its importance in the right place. From a moral uncertainty perspective, the distinction is significant not because we should only take shared reasons into account, but because cases where governments MU-ought to override the public justification principle tend to be those where there are strong non-shared reasons to enforce the law.

Note, however, that public justification by shared reasons may still be desirable from the perspective of moral uncertainty. After all, consensus liberals often suggest that the values grounding public justification (grouped under *Public Reason Liberals Might be Right*) become stronger under consensus: for example, laws that are publicly justified by shared reasons may be more respectful or may better promote political community (e.g. (Leland and Wietmarschen 2017; Lister 2013)). Convergence liberals of course deny this (e.g. (Van Schoelandt 2019)), but we can ignore such niceties here. Our point is only that, while the moral uncertainty approach does not endorse a shared reason requirement, it can hold that shared reasons play a valuable role in public justification – at least insofar as one puts credence in consensus liberals’ first-order arguments to this effect.

## 7 Conclusion

We have argued that while considerations of moral uncertainty cannot vindicate an exceptionless public justification principle, they should at least make us take public justification seriously. If we are uncertain about what moral or normative political theory is right, and thus uncertain about what laws are justified, we should design institutions such that governments at least operate under a pro tanto version of the public justification principle. However, this is only plausible given a fairly non-demanding interpretation of what public justification amounts to: public justification must be a low bar. Moreover, the prohibition against enforcing public unjustified laws may need to be overridden in high-stakes cases where we are confident that failing to enforce a law is quite bad even though

some reasonable people disagree – for example, because the law is needed to protect a class of beings that some reasonable people deny have status. We therefore have provided a preliminary defense of public justification, not as an exceptionless principle, but as a *pro tanto* one.

That said, our arguments here are only first attempts at making the case for public justification through moral uncertainty. We think it unlikely this will be the last word on the matter. Moreover, our hypothesis could be changed in various ways and several robustness checks could be performed. For example, in future work, it would be interesting to consider what more can be said about the conditions under which an insistence on public justification is likely to lead us astray under moral uncertainty, as well as how our conclusions might change given other approaches to moral uncertainty (especially those that reject intertheoretic comparability). It is also worth considering how our conclusions might change if we move from the public justification of laws to the public justification of other objects discussed in the public reason liberalism literature, such as principles of justice, constitutional essentials, or social norms. And, then again, we might compare the public justification principle to a range of other principles to see if we can find one that does even better from the perspective of moral uncertainty, rather than just asking whether instituting the principle has greater expected choiceworthiness than not (as we have implicitly done here). We think there is much fruitful work to be done in this area and, more generally, at the intersection of moral uncertainty and political philosophy.

## Bibliography

- Babaj, Sameer. 2017. 'Self-Defeat and the Foundations of Public Reason'. *Philosophical Studies* 174 (12): 3133–51. <https://doi.org/10.1007/s11098-016-0850-9>.
- Barrett, Jacob, and Gerald F. Gaus. 2020. 'Laws, Norms, and Public Justification: The Limits of Law as an Instrument of Reform.' In *Public Reason and Courts*, edited by Silje A. Langvatn, Matthias Kumm, and Wojciech Sadurski, 201–28. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108766579>.
- Barry, Brian. 1996. *Justice as Impartiality*. Oxford: Clarendon Press.
- Barry, Christian, and Patrick Tomlin. 2019. 'Moral Uncertainty and the Criminal Law'. In *Handbook of Applied Ethics and the Criminal Law*, edited by Kimberly Ferzan and Larry Alexander. New York: Palgrave.
- Bespalov, Andrei. 2021. 'Against Public Reason's Alleged Self-Defeat'. *Law and Philosophy*, July. <https://doi.org/10.1007/s10982-021-09418-6>.
- Bukoski, Michael. 2015. 'Moral Uncertainty and Distributive Sufficiency'. Manuscript.
- Clarke, Simon. 1999. 'Contractarianism, Liberal Neutrality, and Epistemology'. *Political Studies* 47 (4): 627–42. <https://doi.org/10.1111/1467-9248.00221>.
- Cotton-Barratt, Owen, and Hilary Greaves. 2019. 'A Bargaining-Theoretic Approach to Moral Uncertainty'. Global Priorities Institute Working Paper. [https://globalprioritiesinstitute.org/wp-content/uploads/2020/CottonBarratt\\_%20Greaves\\_bargaining\\_theoretic.pdf](https://globalprioritiesinstitute.org/wp-content/uploads/2020/CottonBarratt_%20Greaves_bargaining_theoretic.pdf).
- D'Agostino, Fred. 1996. *Free Public Reason: Making It Up As We Go*. New York; Oxford: Oxford University Press.
- Enoch, David. 2013. 'The Disorder of Public Reason'. *Ethics* 124 (1): 141–76. <https://doi.org/10.1086/671386>.
- . 2017. 'Political Philosophy and Epistemology: The Case of Public Reason'. In *Oxford Studies in Political Philosophy*, edited by David Sobel, Steven Wall, and Peter Vallentyne, 3:132–65. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198801221.003.0007>.
- Feinberg, Joel. 1984. *Harm to Others*. Oxford: Oxford University Press.
- Gaus, Gerald. 2011. *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*. Cambridge: Cambridge University Press.
- . 2015. 'On Dissing Public Reason: A Reply to Enoch'. *Ethics* 125 (4): 1078–95. <https://doi.org/10.1086/680904>.
- Gracely, Edward J. 1996. 'On the Noncomparability of Judgments Made by Different Ethical Theories'. *Metaphilosophy* 27 (3): 327–32. <https://doi.org/10.1111/j.1467-9973.1996.tb00212.x>.
- Gustafsson, Johan E., and Olle Torpman. 2014. 'In Defence of My Favourite Theory'. *Pacific Philosophical Quarterly* 95 (2): 159–74. <https://doi.org/10.1111/papq.12022>.
- Larmore, Charles. 1990. 'Political Liberalism'. *Political Theory* 18 (3): 339–60. <https://doi.org/10.1177/0090591790018003001>.
- Leland, R. J., and Han van Wietmarschen. 2017. 'Political Liberalism and Political Community'. *Journal of Moral Philosophy* 14 (2): 142–67. <https://doi.org/10.1163/17455243-46810052>.
- Lister, Andrew. 2013. *Public Reason and Political Community*. London; New York: Bloomsbury Publishing.
- Lockhart, Ted. 2000. *Moral Uncertainty and Its Consequences*. Oxford: Oxford University Press.
- MacAskill, William. 2016. 'Normative Uncertainty as a Voting Problem'. *Mind* 125 (500): 967–1004. <https://doi.org/10.1093/mind/fzv169>.
- MacAskill, William, Krister Bykvist, and Toby Ord. 2020. *Moral Uncertainty*. Oxford: Oxford University Press.

- Nagel, Thomas. 1987. 'Moral Conflict and Political Legitimacy'. *Philosophy and Public Affairs* 16 (3): 215–40.
- Quong, Jonathan. 2011. *Liberalism Without Perfection*. Oxford: Oxford University Press.
- Rawls, John. 1985. 'Justice as Fairness: Political Not Metaphysical'. *Philosophy and Public Affairs* 14 (3): 223–51.
- . 2001. *Justice as Fairness: A Restatement*. Cambridge, Massachusetts: Harvard University Press.
- . 2005. *Political Liberalism*. New York: Columbia University Press.
- Raz, Joseph. 1990. 'Facing Diversity: The Case of Epistemic Abstinence'. *Philosophy and Public Affairs* 19 (1): 3–46.
- Ross, Jacob. 2006. 'Rejecting Ethical Deflationism'. *Ethics* 116 (4): 742–68.  
<https://doi.org/10.1086/505234>.
- Sepielli, Andrew. 2009. 'What to Do When You Don't Know What to Do'. Edited by Russ Shafer-Landau. *Oxford Studies in Metaethics* 4: 5–28.
- . 2013. 'What to Do When You Don't Know What to Do When You Don't Know What to Do...'. *Noûs* 47 (1): 521–44. <https://doi.org/10.1111/nous.12010>.
- Tarsney, Christian. 2019. 'Normative Uncertainty and Social Choice'. *Mind* 128 (512): 1285–1308.  
<https://doi.org/10.1093/mind/fzy051>.
- Trammell, Philip. 2019. 'Fixed-Point Solutions to the Regress Problem in Normative Uncertainty'. *Synthese* 198 (2): 1177–99. <https://doi.org/10.1007/s11229-019-02098-9>.
- Valentini, Laura, and Christian List. 2020. 'What Normative Facts Should Political Theory Be About? Philosophy of Science Meets Political Liberalism'. In *Oxford Studies in Political Philosophy*, edited by David Sobel, Steven Wall, and Peter Vallentyne, 185–220.
- Vallier, Kevin. 2011. 'Convergence and Consensus in Public Reason'. *Public Affairs Quarterly* 25 (4): 261–80.
- . 2018. 'Public Justification'. In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2018. <https://plato.stanford.edu/archives/spr2018/entries/justification-public/>.
- . 2019. *Must Politics Be War?: Restoring Our Trust in the Open Society*. Oxford: Oxford University Press.
- Van Schoelandt, Chad. 2019. 'Convergence in the Political Liberal Community'. *Public Reason* 11 (2): 3–18.
- Waldron, Jeremy. 1993. *Liberal Rights: Collected Papers 1981-1991*. Cambridge: Cambridge University Press.
- Wall, Steven. 2002. 'Is Public Justification Self-Defeating?' *American Philosophical Quarterly* 39 (4): 385–94.
- . 2010. 'On Justificatory Liberalism'. *Politics, Philosophy & Economics* 9 (2): 123–49.  
<https://doi.org/10.1177/1470594X09345677>.
- Weatherston, Brian. 2014. 'Running Risks Morally'. *Philosophical Studies* 167 (1): 141–63.  
<https://doi.org/10.1007/s11098-013-0227-2>.
- Wendt, Fabian. 2019. 'Rescuing Public Justification From Public Reason Liberalism'. In *Oxford Studies in Political Philosophy Vol. 5*, edited by Steven Wall, Peter Vallentyne, and David Sobel, 39–64. Oxford: Oxford University Press.  
<https://doi.org/10.1093/oso/9780198841425.003.0002>.