

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Public Reason and Private Bias

Permalink

<https://escholarship.org/uc/item/488295nh>

Author

Jayaram, Athmeya

Publication Date

2018

Peer reviewed|Thesis/dissertation

Public Reason and Private Bias

By

Athmeya Jayaram

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Political Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in Charge:

Professor Mark Bevir, Chair

Professor Joshua Cohen

Professor Kinch Hoekstra

Professor Niko Kolodny

Professor Sarah Song

Fall 2018

Abstract

Public Reason and Private Bias

by

Athmeya Jayaram

Doctor of Philosophy in Political Science

University of California, Berkeley

Professor Mark Bevir, Chair

Public reason theorists argue that it is permissible for the state to enforce political norms, such as laws or constitutional principles, when those norms are acceptable to “reasonable people.” Reasonable people are neither actual people, with all their flaws, nor are they perfect people; they are rather a partially idealized group – realistic in some ways and idealized in others. Each of the major public reason theorists – John Rawls, Gerald Gaus, Jonathan Quong, Joshua Cohen – idealizes reasonable people to a different degree, but they all share two claims: 1) Reasonable people hold diverse views of the good life. Nevertheless, 2) Reasonable people can all accept basic liberal political norms grounded in freedom and equality.

My dissertation begins by arguing that theorists are not free to choose any level of idealization, but are constrained in this choice by the justifications of their theories. In particular, idealization is constrained by one essential part of public reason’s justification, which I call the “diversity argument.” The diversity argument explains the first element of reasonable people: why do they disagree about the good? The answers, I argue, attributes certain realistic qualities and tendencies to reasonable people, which therefore constrains how much we can idealize them.

In chapters on the major public reason theorists, I argue that they all offer a diversity argument that does not match the level of idealization that they employ. As a result, they are unable to show that liberal norms are acceptable to reasonable people, appropriately idealized. In the final chapter, I argue that the mismatch in these theories goes even deeper, which we can see when we ask *why* we must accommodate disagreement at all. The answers that philosophers have given us – reasonable disagreement is the inevitable result of human reasoning, human psychology, or free conditions – also apply to *irrational* disagreement. Irrational influences such as implicit bias and motivated reasoning are also inevitable results of who we are and how we live, which means we must accommodate these realistic tendencies in political justification.

So, if public reason theories must now accommodate disagreement among reasonable-but-sometimes-irrational people, what could be acceptable to all such people? I conclude by suggesting a new direction for public reason theories. People who disagree about the good life, but recognize their common biases, can still justify their views to each other by supporting institutions that mitigate those biases, such as non-discrimination laws and deliberative institutions. This requires a new kind of social contract theory – one that is grounded in the shared recognition of our limitations, rather than our shared reasons.

This is dedicated to my parents, for their dedication to me.
I could not imagine a more loving and supportive pair.

Table of Contents

| | |
|---|----|
| Introduction: How to Justify a Public Reason Theory..... | 1 |
| Chapter One – Government For the People, By the Viewpoints? Realism and Idealism in Public Reason..... | 16 |
| Chapter Two – Judgments versus Viewpoints: Minimal Idealization in John Rawls..... | 26 |
| Chapter Three – Extraordinary People: Moderate Idealization in Gerald Gaus..... | 36 |
| Chapter Four – A Reasonable Hope: Maximal Idealization in Jonathan Quong..... | 55 |
| Chapter Five – Liberal Democrats: Procedural Idealization in Joshua Cohen..... | 78 |
| Chapter Six – Public Reason and Private Bias..... | 98 |

Acknowledgements

I would like to thank my dissertation committee members Professors Mark Bevir, Joshua Cohen, Kinch Hoekstra, Niko Kolodny, and Sarah Song for their rigorous but constructive comments throughout. I also want to thank Richard Ashcroft and Andrius Gališanka, who always read my first drafts and still encouraged a second. Finally – which is a sentiment they would share – thank you to my family, Minny, Jai, Amshu and Niru, for putting up with the blank stares, closed doors, and staycations. I am grateful for their patience and their kindness.

Introduction: How to Justify a Public Reason Theory

In this introductory chapter, I describe the virtue of a public reason theory as its ability to claim that appealing political norms could be accepted by a certain subset of citizens – “reasonable” citizens. The combination of independent appeal and acceptability is what makes political norms legitimate. However, I argue that what makes a norm “acceptable” is vague enough for theorists to claim that just about any group of citizens “could accept” just about any political norms. So, to justify a public reason theory, we must be able to justify each of three elements of “acceptability.” And this can only be done in reference to an underlying justification for taking the public’s views into account in the first place. I then summarize how each of the theories that I consider in this dissertation suffers from an inconsistency between their justification and their interpretation of acceptability.

What is a Public Reason Theory?

What makes political norms – our political principles, laws, or social rules – legitimate?¹ That is, what gives the government the right to enforce those laws or a citizen the moral duty to obey them?² Broadly speaking, the answers to these questions have focused on two sources of value, consent and correctness, which both have an intuitive appeal. When a citizen consents to a law, there is an intuitive sense in which he has bound himself to obey it or to have it enforced, just as a promisor binds himself to a promisee. The appeal of a law that is correct – by some objective standard, independent of consent – is even more straightforward. A morally correct law or political principle may, on its own, imply the right to enforce it or a duty to obey it.

However, there is an equally intuitive doubt that either of these answers would be sufficient without the other. It does not seem as though just any law, no matter how incorrect, could be legitimate, just because most citizens consented to it. Nazi principles, however popular in 1930’s Germany, cannot be the basis of legitimate law and any theory that says otherwise seems wrong on its face. On the other hand, it does not seem as though a law grounded in true moral principles could be legitimate, if many citizens rejected it. David Estlund offers an example that supports this intuition: Imagine that the Pope is, in fact, infallible in his pronouncements.³ Even if we, as theorists, were certain that Papal law would be correct law, there seems to be something wrong with enforcing those pronouncements on citizens, the majority of whom would not consent to the Pope’s authority or accept his infallibility. If these two intuitions are right, then neither consent nor correctness, on its own, is a sufficient basis for legitimate law.

¹ As I will discuss, public reason theorists have different views on what must be legitimated, but I will use the term “political norms” to cover the various options.

² Theorists such as A. John Simmons distinguish between the right to enforce and the duty to obey, but public reason theorists rarely make this distinction, so I will also avoid this debate and discuss both together.

³ David Estlund, “Epistemic Proceduralism and Democratic Authority” *Does Truth Matter* ed. by Raf Geenens and Ronald Tinnevelt (Dordrecht: Springer, 2009) p.17.

All of these intuitions present us with a problem. Both consent and correctness seem important to legitimacy and neither one seems sufficient without the other. The problem is that there appears to be a trade-off between the two. Political norms that can gain everyone's consent, if there are any, are not likely to be the correct ones, except in the luckiest of societies.⁴ And, norms that are correct, by some standard that is independent of what those citizens happen to think, are not likely to be consented to by all, or even most, citizens. If a high degree of consent and correctness are both necessary for legitimacy, and there is a trade-off between the two, it seems as if legitimacy is an impossible goal.

For public reason theorists, however, this dichotomy of who decides versus what is decided is not as stark as it appears. Even if we cannot maximize both consent and correctness, they argue, we can achieve the relevant value behind both approaches to a sufficient degree if, instead of looking for norms that all citizens agree on or laws that are objectively correct, we aim at norms that "reasonable" people "could accept."

How does this help with our problem? First, by only requiring that norms be acceptable to "reasonable" people, and not all citizens, we can potentially improve the quality (or level of correctness) of the norms they can accept. For example, based solely on how we use the term "reasonable," we can see several ways in which a group of reasonable people could accept higher-quality political norms. We can think of being reasonable as not being selfish, as recognizing that others' needs are also legitimate and therefore being disposed to give and take. We can think of being reasonable as not being overly emotional and therefore willing to listen and act on reasons. And, we can think of being reasonable as not insisting on the overwhelming importance of any particular desire or interest, but balancing considerations against each other. So, just based on an intuitive sense of the term, limiting the group to "reasonable" people can improve the moral quality of political norms in important ways. Reasonable people are more likely to broadly distribute the benefits and burdens of society according to reasons, rather than self-interest or power.

Second, by only requiring that reasonable citizens *could* accept the norms, and not necessarily *do* accept them, we can increase the number of citizens that "consent" to the norms. Of course, this is now a peculiar version of consent. Some would argue that saying that someone could accept a norm is not consent at all, since, in reality, she neither has chosen the norm nor do we know if she would actually choose the norm. Consent, they might argue, must involve an act of will by the individual in choosing the norm.

Public reason theorists would acknowledge this, but claim that there is still something valuable about acceptability. John Rawls might say that the value is that you *could* consent; that we are not asking you to change anything essential about yourself in order to consent. Gerald Gaus sees the value as a form of moral autonomy; that you would consent if you were acting on your own reasons.⁵ I will explore the details of each

⁴ This is true, provided that correctness is independent of consent.

⁵ Gerald Gaus, *The Order of Public Reason* (Cambridge: Cambridge University Press, 2011) pp.220-224.

account in later chapters, but for now, I will capture the general value as “responsiveness.” When a norm is responsive to me, it may not be one I choose, but it still bears a significant relationship to me. Given my beliefs, motivations, and capacities, that norm is accessible to me; it is something I could choose. This may not get us the full value of consent, but it does seem to answer our worry about being obligated to obey a truth that is completely alien to us. In the Pope case, for instance, we can account for our objection to enacting his pronouncements by saying that his pronouncements do not reflect the views and motivations of many citizens, and there are no (non-oppressive) conditions under which most citizens could come to accept Papal authority. By substituting responsiveness for consent, therefore, public reason theories are able to claim that political norms are *acceptable* to citizens, even if citizens do not, in fact, consent to them.

And third, asking which norms reasonable people could accept mitigates some of the trade-off between consent and correctness. Having higher quality norms does not mean that only a small subset of citizens can consent to them since we now only care about the consent of *reasonable* people who, by definition, can accept higher-quality norms. At the same time, having more people “consent” to the norms does not necessarily threaten their quality since we now only care about responsiveness, which can be more widely distributed without giving more people veto power over the norms.

So, by saying that legitimate laws are those that reasonable citizens can accept, public reason theories can claim a higher level of both responsiveness and correctness, while mitigating the trade-off between the two. Now, we can say that a law is not legitimate simply because it is agreed upon by most citizens; it must meet a certain level of quality, which is ensured by taking into account only reasonable citizens.⁶ Nor is a law legitimate simply because it is correct; it must be responsive to all reasonable people, not just the ones with the correct moral view. In this way, public reason theories achieve something of the values behind both consent and correctness.

How Not to Justify a Public Reason Theory: Balancing and Reflective Equilibrium

If the value of public reason theories comes from the levels of responsiveness and quality that the theories support, then the most obvious way to justify a public reason theory is to see whether it achieves a balance of the two values. But, there are many possible balances of responsiveness and quality. We could propose a very inclusive group of reasonable people (citizens who want to cooperate with each other, for instance) who could therefore only accept a lower level of quality, since the group is not defined by many epistemic or moral capabilities. On the other end, we could propose a narrow reasonable group (the group of political philosophers, maybe) that can therefore accept a high level of correctness. Both proposals seem equally valid because they both offer us some balance of correctness and responsiveness. There is, so far, no reason to prefer one to another.

⁶ I should note that this is not a theory of who deserves actual political power such as voting rights. It is a theory of how we should represent people in a hypothetical model that will determine which political norms are legitimate. These legitimate norms then determine the nature of actual political institutions, such as actual democratic processes.

To answer these charges, public reason theories will have to argue that the levels of responsiveness and correctness that their theories support are *sufficient*. Many do so using a method that Rawls calls “reflective equilibrium.” Theorists begin with a set of intuitions regarding the group that deserves responsiveness and the kinds of political norms that we must have. As public reason theorists use them, these are often intuitions on the minimally sufficient levels of both. These intuitions tell us who *must* belong in the reasonable group, if anyone should, or which laws *must* be legitimate, if any are. We then tweak the group and the laws, back and forth, until the former can accept the latter – until the two are in equilibrium.

For example, Gerald Gaus uses this method as part of the justification of his public reason theory.⁷ He starts with what he believes should be the reasonable group: those with views that are “mutually intelligible,”⁸ by which he means⁹ sharing a set of basic values: life, liberty, happiness, and so on. We have an intuitive sense of what it means for views to be intelligible to us, he says, but we don’t know the exact content. At the other end, Gaus believes we need a set of norms (or, his preferred goal, ‘social-moral rules’) that are specific and extensive enough for effectively “structuring our social life.” Again, he does not specify in advance exactly what it means to structure our social life, but this is the guiding intuition. Gaus then goes through the process of reflective equilibrium: he stretches the meaning of intelligibility and sees whether this larger qualified group can still accept social rules that structure social life. If not, he might contract the meaning of intelligibility or the need for structure – back and forth – until the two are in equilibrium. In the end, he arrives at a pairing of reasonable people (which he calls “Members of the Public”) and political norms that, he argues, is uniquely justified because it accounts for both sets of intuitions together. This pairing shows, better than any other, how we can satisfy our intuitions about both who deserves public justification *and* the quality of the laws we must have.

⁷ “Without accepting the details of Rawls’ account, we can conceive of the development of the Deliberative Model as an iterative procedure. In constructing the Deliberative Model we start out with a very broad range of moral persons who disagree, and construct Members of the Public as idealized representations. We then see what sorts of moral rules can be justified to them. We then may go back and consider those value systems that stretch intelligibility, which we have not yet modeled. We shall want to inquire whether these persons are committed to the moral enterprise and treating others as free and equal; if they are, we must see whether we can model any remotely intelligible Member of the Public to represent them. If we can see how some such persons might be described, we may iterate our deliberative model to determine whether the results radically change. Do we find that our justified social morality is fundamentally transformed in a way that is manifestly unacceptable (say, it shrinks to a very small core or even disappears)? If we find this, we may conclude that we have reached the limits of the reconciliation of reasonable pluralism with a morality and free and equal persons, and must confront the conclusion that, in the end, moral relations with such persons is not possible. To include them would radically, and unacceptably, reduce the scope of social morality. And this, of course, would be a great cost to us. It would impair social morality’s ability to perform the core function of structuring our social life. On the other hand, to exclude moral persons is also a terrific cost: our moral relations with such people would be transformed into the relations that obtain between us and those who are not capable of moral autonomy (§12.3b) such as the psychopath.” Ibid., p.282.

⁸ Ibid., p.282.

⁹ In the chapter on Gaus, I will note that what he means by “intelligible” changes over the course of his account.

While reflective equilibrium is an improvement on the balancing justification, however, it still fails to uniquely justify one public reason theory over all the others for two reasons. First, and most obviously, intuitions differ, on both the reasonable group and the minimal level of correctness. As we have seen, Gaus starts with intuitions supporting a broad group and a lower level of correctness. Jonathan Quong, another prominent public reason theorist, does not share these intuitions. He considers Gaus' group too broad and his rules to be insufficiently correct.¹⁰ For Quong, the laws must be more robustly liberal, so the minimum we can accept are norms that ensure and prioritize certain liberal rights that Gaus' theory does not guarantee. The difference in their intuitions on the reasonable group and the sufficient norms would produce very different public reason theories, both apparently justified by reflective equilibrium.

But, there is a deeper problem with the 'reflective equilibrium' defense. Let us imagine that we all shared the same intuitions on the requisite group and quality of the laws. And, at least after some equilibrating, the intuitively justified group "could accept" the intuitively justified laws. At this point, it seems like we have found a winner. For one thing, this public reason theory achieves a level of responsiveness – because the qualified group can accept the laws – *and* a level of correctness – because the laws represent some level of correctness. Second, instead of a mere balance of the two values, we can give some *justification* for the levels of responsiveness and correctness that the theory achieves: they accord with our (we are now assuming) shared intuitions.

However, while we have now (potentially) justified the reasonable group and the political norms, we have not yet justified the third key element of a public reason theory: the connection between the two. We must justify the claim that the qualified group *could accept* the laws. This claim of acceptability is crucial to public reason. It is the basis of the ultimate claim that the norms are responsive to the qualified group, which is, in turn, the reason that we should choose those norms over ones that are "fully correct."

Public reason theorists pay insufficient attention to this third element, possibly because it sometimes seems obvious that a particular reasonable group could accept particular norms. For instance, if our intuitions tell us that the minimally sufficient group is the

¹⁰ This is how Kevin Vallier describes the difference between his Gaussian theory and Quong's theory, in terms of the level of responsiveness: "Quong argues that for the sectarianism objection to succeed, 'it would have to show that my account of political liberalism is sectarian in this objectionable sense—that it fails to be justifiable to some people whom we should identify as reasonable, and thus people to whom our political principles ought to be justifiable.'" I contend that the coherence and attractiveness of the convergence view shows that coercion ratified by the internal conception cannot be justified to some people that we should identify as reasonable—All Things Considered Reasoners"

In terms of the difference in the intuitive level of correctness: "Quong might reply by insisting that the internal conception is still more liberal than the convergence conception because convergence cannot show that the justification of illiberal conclusions is impossible. After all, a moderately idealized constituency with different cultural and moral values might reject liberal institutions, so the convergence view could count an illiberal regime as justifiable. But we should reject the condition that a liberal political theory is liberal to the extent that it makes justifying illiberal laws impossible." Kevin Vallier, "On Jonathan Quong's Sectarian Political Liberalism," *Criminal Law and Philosophy* 11(1), pp.15-16.

group of liberal-democratic citizens and the minimally sufficient norms are liberal ones, then it seems tautological to say that the former could accept the latter.¹¹

Justifying this claim of acceptability, however, is much harder than it seems. The claim that a group “could accept” the laws, unlike the claim that they *do* accept the laws, does not have a definite meaning. We cannot justify its meaning intuitively because, unlike (perhaps) the “reasonable group” and the “minimally sufficient norms,” we do not have definite intuitions about what “could accept” means. For example, if I said that you could accept a piece of fruit instead of dessert, I might mean that you would not spit it out or that you would welcome it, or something in between. And, I might mean “you,” just as you are, or I might mean “you,” if you understood the relative nutritional values, or something in between.

The indeterminacy and flexibility in “could accept” has led public reason theorists to interpret the concept in many different ways. So many ways, in fact, that I will argue it is possible to claim that just about any reasonable group “could accept” just about any political norms by manipulating three variables in the definition: the level of idealization, the object of acceptance, and the intensity of acceptance. If this is right, then the claim of acceptability carries no justificatory value on its own for the same reason that the balancing justification did not; any public reason theory can make this claim. To see how this is possible, let me describe the three variables in more detail.

The Level of Idealization

Let us start with the level of idealization. As I said, public reason theories all claim that a certain group of reasonable citizens can accept certain political norms. So, the first difference among the theories is who they claim is reasonable. Quong, for instance, starts with liberals and Gaus with what he calls “moral people.” I will return to these differences later. For now, I want to discuss how these citizens’ views are represented through the different levels of idealization a theorist can employ.

Imagine we start with a very broad group of reasonable citizens such as those who accept democracy. If we find some political norms that all democrats can accept, it would seem like we have done something important; we have found norms that are responsive to a large majority of citizens. At first glance, however, this does not appear possible. People that agreed with some version of the democratic ideal – agreed that, let us say, we should settle political differences through fair voting procedures – are not likely to agree on much else. They will not only differ in nearly all of their other interests and values, but they may even have different ideas about what democracy means, apply its values inconsistently, or simply be confused.

Idealization helps to smooth out some of these edges in ways that, individually at least, seem fairly innocent. For instance, let us say that I won some money at cards and I decided to spend it on one of my children. The first question I would ask would be which of my kids deserves a gift – whose wishes should I take into account? And If I answered

¹¹ This is a simplified version of Quong’s theory, which I will argue is not, in fact, tautological.

that my eldest son was most deserving, then I would know whom to consider and that might be the end of it.

But, there might also be a second question, which is more awkwardly stated: which aspect of my eldest son am I rewarding? How should I think of him when I consider what to give him? For instance, I could just ask myself what he likes, which would represent him by his desires. But, I could also ask what his current tastes say about what else he might like, and then get him one of those. Or, what does he *value*? Or, what would he value if he were exposed to a wider range of experiences? In all these cases, I am representing my son according to some aspect of him (his desires or his values) and under some conditions (if he were exposed to more objects or experiences). Instead of just taking into account what he currently wants, I am considering him as if he has been represented in particular ways. My son is still deserving of a gift, but deciding what would constitute a gift for him might mean considering this idealized son.

In the public reason context, however, idealizations can go a long way toward transforming any initially contentious group into one that can accept the same set of norms. For instance, we might start the idealization by saying that, since we are trying to justify laws to those who believe in democracy, we should consider what could be justified to them *as democrats* and not in their other identities as Muslims or mathematicians or capitalists. So, we can begin by *abstracting* to only the values associated with democracy, whatever we deem them to be. Then, we might say that, since we want to justify the norms to democrats, we must idealize away any confusion that they have about what democracy requires. So, we might attribute to them the *correct interpretation* of democracy, ensure that they apply this interpretation *consistently* and follow the right rules of inference and evidence. Since we would not want to justify laws to democrats just based on the fact that they are poorly informed, we then idealize them to possess all the *relevant information*. Finally, we might want to represent the group as not only holding democratic views but as being *motivated* to uphold them, so we idealize their motivations.

At the end of this process, we have transformed a real group of democrats into a viewpoint that is abstracted, interpreted, rational, informed, and motivated to adhere to these enhanced views under all conditions. It seems likely that, after all of these steps, we can claim that this group of democrats “could accept” a much wider set of norms than they could prior to the idealization. Certainly, we can now say that this initially broad group of democrats can accept a very high level of correctness, in contrast to our initial impression.

The Object of Acceptance

While public reason theorists have debated the proper level of idealization, much less attention has been paid to the other two elements of “could accept”: the object and intensity of acceptance. In the language of public reason, what we mean when we say that a (now reasonable and idealized) group “could accept” political norms can vary along two lines. It can vary according to the *object* of acceptance – that is, *what* must be

acceptable for the laws to be legitimate. And it can vary by the *intensity* of acceptance – how *strongly* the idealized citizens must endorse the laws.

Public reason theories are commonly divided into two types according to the first category: the object of acceptance. Some theories argue that the object(s) of acceptance – what must be acceptable for norms to be legitimate – are the laws, principles, or social rules themselves; these are called “convergence theories” because the reasonable citizens “converge” directly on the norms from their different beliefs and values. Other theories argue that norms are legitimate when the *reasons* supporting those principles are acceptable; these are called “consensus theories” because they first require idealized citizens to reach a “consensus” on a fund of acceptable (“public”) reasons. Idealized citizens then support their favored norms based only on these acceptable reasons.

For example, take a political norm like “all citizens should have equal political rights.” One idealized citizen, Anu, might find that norm acceptable directly. In this case, let us say that Anu finds it to be a true or optimal principle. Another idealized citizen, Babu, might deny that this is a morally true principle but still find it acceptable because it is based on what he considers genuine moral reasons. Babu might agree, for example, that the ‘equal political rights’ principle is based on moral reasons such as the importance of widespread political participation and the value of just political decisions. He simply thinks that these reasons better support a different principle: ‘all citizens should have some political rights, but educated citizens should have more.’ So, in some sense of the term, both Anu and Babu “could accept” the principle, but public reason theories will vary in the object of that acceptance; whether what must be acceptable is the norm itself (convergence) or the reasons that support the norm (consensus).

How does changing the object of acceptance make it easier to connect any group with any norm? In two ways: First, by arguing for a consensus theory rather than a convergence theory, it becomes significantly easier to claim that an idealized group can accept the same norms. Let us say again that we are trying to find principles that all democrats can accept, a seemingly impossible task. So, we idealize this reasonable group in all the ways listed above and thereby narrow the range of disagreement, though we cannot eliminate it. Since they still cannot accept the same principles, we would say that the views of the idealized democrats do not “converge.” If convergence were required, therefore, our public reason theory would fail to find principles that all reasonable citizens could accept.

However, if we adopt a consensus theory, then it is still possible to claim that these idealized democrats could accept the same principles because now we mean something different by “could accept.” A consensus theory does not ask whether the idealized democrats agree on the principles directly, but merely that they see the reasons that support the principles as genuine reasons, which, in this example, means “true moral” reasons. This meaning of “could accept” makes it much easier to claim that the idealized democrats could accept the same set of principles because they no longer have to see those principles as supported by their “all things considered” reasons – that would be a convergence theory – but just as supported by *some* genuine reasons. The claim that all

democrats could see a set of norms as based on genuine – even if not the best – reasons is a much easier claim to defend. We only have to find some reasons that (now idealized) democrats share and see which norms those reasons would support and which they would rule out. So, the move to a consensus theory makes it easier to claim that an initially contentious group could accept the same laws.

There is a second way that varying the object of consensus can help to connect a group to norms, and this one is available to both convergence and consensus theories. Both objects of acceptance – reasons and norms – can vary on a scale from abstract to concrete. And, whether the object is reasons or norms, it is much easier to gain agreement among a group if the object is very abstract. For example, it is reasonable to say that most American citizens agree with (at least most of) the political principles embodied in the Bill of Rights. It is reasonable to say that, however, only because they are phrased so abstractly: the right to free speech; no cruel and unusual punishment; not to be deprived of life, liberty and property without due process. Once these phrases are given a concrete interpretation that can be enforced, our agreement disappears.

To take an example from public reason, consider Gaus' convergence theory. As a convergence theorist, he argues that we must justify the laws, principles, or social rules to citizens and not just the reasons behind them. But, there is a great difference between setting the object of acceptance at laws and setting it at principles – a difference of abstraction. At the very abstract end, a theorist might require the acceptance of political principles such as the constitutional rights just mentioned. At the very concrete end, a theorist might require the acceptance of particular laws like the Affordable Care Act, which would be much harder to achieve. Gaus falls in the middle.¹² He rejects very abstract objects of acceptance, like constitutional norms, because the acceptance would disappear once they were interpreted. On the other hand, he also rejects a requirement to accept individual laws, which are too specific and complex for individuals to plausibly have views on. So, he argues for an object of acceptance that he calls “social rules,” which are more concrete than principles but more abstract than laws. They are broad enough to cover a range of cases, but specific enough to actually determine a conclusion in these cases. Setting the object of acceptance at this intermediate level allows him to claim more acceptance than a theorist that requires a more concrete object.

So, whether the object of acceptance is reasons or norms, it is much easier to claim acceptance if the object of acceptance is abstract. It is easier still if the object of acceptance is reasons not norms. Varying the object of acceptance in these two ways goes a long way toward claiming that a very diverse group “could accept” the same laws. If it is not yet clear that the claim that a group “could accept” certain norms is an empty claim (without further justification), there is yet another way to posit a connection: changing the intensity of acceptance.

¹² Gerald Gaus, *The Order of Public Reason* (Cambridge: Cambridge University Press, 2011) p.271.

The Intensity of Acceptance

When we say that certain objects of acceptance (laws, principles, reasons or social rules) are acceptable to an idealized group of citizens, the word ‘acceptable’ contains a range of intensities. At one extreme, the idealized citizens may find the laws barely tolerable or “not rejectable”; that is, they are just decent enough that they don’t cause the citizens to secede from society or want to overthrow the regime. At the other extreme, idealized citizens may find the laws acceptable as in optimal or ideal, according to whatever standard of correctness that they (or the theorist) employs, such as truth or justice. In the latter case, the connection between reasonable people and the norms that they could accept would be much tighter.

In between, the citizens could accept the principles at many intensities, including, as both Rawls and Gaus argue, seeing them as within a range of possibilities that are, in some sense, good enough.¹³ Rawls, for instance, argues that, as long as citizens find the reasons to be genuine moral reasons, they can disagree about the weight of those reasons. Some might find them to be the most weighty (optimal) reasons, while others consider them barely reasonable¹⁴ – one minor consideration among others. Gaus has a more specific interpretation of ‘good enough’ social rules. He argues that, for a social rule to be in the eligible range, idealized citizens must see adopting that rule as better than having no rule at all.¹⁵ In both cases, however, the citizens “could accept” the object as more than merely tolerable but as less than optimal.

Clearly, reducing the intensity of acceptance required to claim that laws are acceptable greatly increases the number of citizens that “could accept” those laws. For instance, not everyone sees freedom of religion as optimal since some believe that they have the one true religion and others should be suppressed, while others believe that all religions are false and they should be discouraged. But, if the intensity of acceptance only required toleration, then a lot of those people can be said to tolerate freedom of religion – to not be so opposed that they could not abide by it. So, reducing the intensity of acceptance required makes it much easier to claim that group “could accept” a set of norms.

How to Justify “Could Accept”

Between idealization and the object and intensity of acceptance, public reason theories have many ways to claim that just about any reasonable group can accept just about any norms. If this is right, then, without more, the claim of acceptance is any empty one. We could simply begin with the most correct norms – most just, most true – and show that those norms could be accepted by the broadest group. For instance, we can plausibly argue that Kantian liberalism could be accepted by all reasonable citizens, if we idealize

¹³ “...the best response is to give up on the hope that we can construct a compelling description of the deliberations of members of the “realm of ends” that will lead them to agree on the same rule. The most we can achieve, I shall argue, is a compelling description that selects, as Rawls put it in his later work, a set of reasonable rules...the disagreement in our private judgment is extensive but is bounded within a set. Within this set, Members of the Public will differ in their ranking of proposed rules.” Ibid., p.43

¹⁴ John Rawls, *Political liberalism* (New York: Columbia University Press, 1993) p.446.

¹⁵ Gerald Gaus, *The Order of Public Reason* (Cambridge: Cambridge University Press, 2011) p.322.

them to have only relevant reasons, full information, and flawless reasoning, and they only have to find very abstract reasons behind the laws to be barely tolerable.

So, what do we need to make the claim of acceptance meaningful? Since the problem is our ability to manipulate these variables, we need to provide a justification for choosing one or another level of idealization, object of acceptance, and intensity of acceptance. David Enoch has argued that the justification for the level of idealization is constrained by the “underlying motivations” (or justification) of the theory¹⁶ – why do we want political norms to be responsive at all?

To return to the earlier example, if I wanted to get a gift for my idealized son, I still cannot idealize him in any way I choose; it must be a normatively acceptable way, in two senses: 1) I must idealize him such that, at the end of the process, I can still plausibly claim to be giving a justification *to* him. If I idealize my son by considering what he might enjoy if he were exposed to more things, then I am still giving *him* a gift. However, if I transform my son into a philosopher, then nothing about him, as an individual, is influencing my choice. I might be giving him a gift that *I* could appreciate, but not one that he would. In public reason terms, I might be giving him the correct norms, but he is not included in deciding their correctness.

Any idealization must also be normatively acceptable in a second way. It must be consistent with the justification for selecting the reasonable group in the first place. For instance, if the driving value of my theory is respect for people’s reasoning, then I might include every adult in the reasonable group, since every adult reasons, to some degree. It would be perfectly consistent with this justification for me to then represent the views of reasonable people by their reasoning and not by their desires or phobias. The idealized views would then be the reasoning of all adults, at whatever levels of sophistication they reason. However, it would not be consistent with this justification for me to represent reasonable people as more sophisticated reasoners or attribute only true beliefs to them, because the purpose of including them in the first place was to respect their reasoning as it was, not to respect sophistication or truth. So, the validity of the two moves in public reason – the selection of the reasonable group and their idealization – depends on the common justification of both: why must we take into account people’s views in the first place?

However, I would go further and argue that not just the level of idealization and selection of the reasonable group, but every element of the public reason theory is constrained by its underlying justification: who is reasonable, how we can idealize them, the object they must accept, and the intensity with which they accept it. For example, let us say that we proposed a public reason theory that was justified by the need for a stable political

¹⁶ “...for the idealizing move to be legitimate, it has to be shown that the idealization is consistent *with the underlying motivations for the view*, the considerations that led us down the path of tying the relevant normative phenomena to people and their responses. If such idealization is in fact inconsistent with the underlying motivations for the view, it is objectionably ad hoc. In such a case, going for idealization would be analogous to the hungry – seeing that food is hard to find – settling for hypothetical food.” David Enoch, “The Disorder of Public Reason,” *Ethics* 124 (1): p.165.

system. Unless the public “could accept” the political norms, we argue, they might revolt or reject the system, which would make it unstable. This justification places clear constraints on all the elements of the theory: 1) The reasonable group must be broad since we are aiming for widespread obedience. 2) The object of acceptance cannot be too abstract because that would allow the public to agree in the abstract but disagree in the particular, which would not alleviate the worry of instability. 3) The intensity of acceptance, however, could be fairly low, since all that is required is mere obedience and toleration. If these elements were otherwise, then they would not explain how the public “could accept” the political norms in a way that would contribute to stability. And, when these elements are ad-hoc or inconsistent with the underlying justification, we have reason to suspect that their purpose is merely to connect the reasonable group with desirable norms and make the theory work.

Most public reason theorists recognize the need for, and provide, an underlying justification for their theories. But, in this dissertation I will argue that none are fully consistent in aligning all the elements of their theory with its underlying justification. Each of the major public reason theories suffers from a mismatch between its underlying justification and one or more of the necessary elements of its theory.

In Chapter One, “Government For the People, By the Viewpoints?” I argue that one major inconsistency is inherent in the public reason approach, which is why it cannot succeed in its present form. Part of the underlying justification for every public reason theory is what I call “the diversity argument.” The diversity argument tells us why political norms ought to be acceptable to people with diverse views of the good, which explains why we should turn to “public reasons” instead of right reason. In the process of performing this essential function, however, the diversity argument *itself* causes a major problem for public reason theories. In explaining why reasonable people will inevitably hold different views of the good, diversity arguments attribute certain realistic tendencies and limitations to reasonable people, therefore committing the theories to finding norms that realistic people could accept.

However, instead of showing liberal norms to be acceptable to realistic people, public reason theories show them to be acceptable to fixed and sophisticated “viewpoints”: Christian viewpoints, Utilitarian viewpoints, citizen-viewpoints, and so on. If public reason theories are committed to finding norms that are acceptable to realistic people, but can only find norms that are acceptable to viewpoints, their conclusions fail their own test of legitimacy.

Over the next four chapters, I argue that each of the major public reason theorists – John Rawls, Gerald Gaus, Jonathan Quong, and Joshua Cohen – offers a diversity argument that does not match the level of idealization that they employ. The mismatch is different in each case. Each offers a different diversity argument that, in turn, commits his theory to a different level of realism. Each also idealizes reasonable people to different levels: minimal, moderate, substantive and procedural idealizations, respectively. In each theory, however, there is a mismatch between the commitment to realism and the idealization

needed to ground their conclusions. This is not surprising given that, as I argue in Chapter One, the inconsistency is built into the public reason approach.

Chapter Two, “Judgments vs. Viewpoints: Minimal Idealization in John Rawls,” argues that even the least idealized of public reason theories – Rawls’ “political liberalism” – faces a conflict between a commitment to realism and the need for viewpoints. The commitment to realism comes from Rawls’ diversity argument, “the burdens of judgment,” which explains why reasonable people will inevitably disagree about the good life. After evaluating several interpretations of the burdens of judgment, I argue that Rawls must see reasonable people as burdened by the influences of their particular life experiences, and by their difficulties in consistently making the right decision on difficult moral questions. If this is right, however, then his solution to the problem of pluralism cannot work. If reasonable people are realistic and limited in those ways, then they cannot consistently prioritize public over private reasons and reliably arrive at liberal conclusions. In other words, on pain of inconsistency, Rawls cannot imagine reasonable people as unchanging and consistent viewpoints.

Similarly, in Chapter Three, “Extraordinary People: Moderate Idealization in Gerald Gaus,” I argue that even a theory that accommodates more realistic human qualities, Gerald Gaus’ “moderate idealization,” must abandon this realism to claim that diverse people can accept basic liberal norms. Gaus argues that social norms must accommodate diverse views because those norms must be internalized by ordinary people, and ordinary people have limited time, information, and reasoning ability, so we cannot expect them to arrive at the same truth. Instead, we must only enforce norms that people have sufficient reason to accept, given their diverse starting points and cognitive limitations. He then argues that diverse viewpoints would all accept certain basic liberal norms by showing how, if we abstract to each person’s core reasons, order them correctly and extrapolate them accurately, we can get convergence on rights such as freedom of conscience. However, to imagine that ordinary people can perform such reasoning – can identify and reason from a coherent and sophisticated viewpoint – is precisely what “moderate idealization” stands against.

In Chapter Four, “A Reasonable Hope: Maximal Idealization in Jonathan Quong,” I consider Quong’s attempt to solve the inability to justify liberal norms by further idealizing reasonable people – this time into liberal people. We do not have to justify liberalism to non-liberals, Quong argues, but merely to justify political liberalism to perfectionist liberals. Since we are assuming the value of liberalism, we can assume that reasonable people all hold basic liberal values. What we cannot do, he argues, is assume that reasonable people hold perfectionist liberal values because that level of agreement could not come about in any possible liberal society, no matter how ideal. As long as humans live under conditions of freedom, they will disagree on perfectionist values, but they may not disagree on “political” liberal values.

However, while this maximal idealization does solve the problem of how reasonable people could all accept liberal norms, it comes at the cost of even the minimal realism that Quong requires. Once we understand what reasonable people must accept to endorse

political liberalism, we see that it is just as much an ideal theory as perfectionist liberalism. To retain the value of a public reason theory, we must idealize the public in ways that are consistent with at least a minimal realism requirement.

In Chapter Five, “Liberal Democrats: Procedural Idealization in Joshua Cohen,” I examine a view drawn from Cohen’s work, which does not imagine reasonable people to begin with liberal values or sophisticated viewpoints, but instead attributes only an initial acceptance of the deliberative-democratic ideal and its attendant procedures. On this view, political norms must accommodate diverse views because even ideal democrats under ideal deliberative conditions would retain their diverse and deep commitments, such as their fundamental religious views. Cohen further argues that, when faced with conflicting fundamental commitments, ideal democrats would always side with those whose freedom is restricted over those who would want to restrict freedom, which is why ideal democrats would endorse substantive liberal norms such as freedom of religion and expression. However, I argue that this presumption of liberty is neither contained in the deliberative-democratic ideal nor a product of ideal deliberation. So, to draw these substantive liberal norms from diverse democrats, Cohen must assume more than a belief in democracy.

In Chapter Five, “Public Reason and Private Bias,” I show how deep the public reason challenge truly goes and offer a new direction to meet it. As I have argued, public reason theories claim that we must accommodate religious disagreement in part because of our realistic human tendencies and limitations, which lead us to disagree. However, if we must accommodate one product of our tendencies and limitations, then we may have to accommodate other such products. And however you characterize the tendencies that produce religious diversity – for example, you might say that even the most competent reasoners are influenced by their upbringing – there will be other, *irrational*, influences that fit that characterization and therefore must be accommodated. Consider these possibilities: motivated reasoning, cognitive illusions, identity politics, stereotyping, and implicit bias. If any of these fit the justification for accommodating religious reasoning, then public reason must also accommodate private bias.

Public reason theories are therefore committed to answering a more challenging question than they acknowledge: “which norms would be acceptable to diverse and, in some ways, irrational reasoners?” If this is the question, then the answer can no longer just be “the norms that rely only on public reasons,” because if the reasons are public but the reasoning is irrational, the conclusion would still not be acceptable to all.

Instead, I will argue that reasonable-but-irrational people can make their views acceptable to each other, not just by using public reasons, but also by supporting a system that will correct for or mitigate their irrationality. And this system can contain many of the features and safeguards of a liberal-democratic system. For instance, being aware of our irrational tendency to favor our own group may require us to support non-discrimination rights. Or, being aware of the stereotypes that result from our limited exposure may require us to support broad democratic deliberation. In this way, we may be able to acknowledge the human limitations that public reason theories are committed to

accommodating, while still justifying elements of a liberal-democratic system. This new direction for public reason theories promises to justify appealing, though more modest, political norms, while remaining consistent with the theories' aim: to find norms that reasonable *people* could accept.

Chapter One – Government For the People, By the Viewpoints? Realism and Idealism in Public Reason

Since John Rawls, public reason theorists have attempted to show how liberal political norms could be acceptable to people with diverse religious and philosophical viewpoints. However, these theories overlook the importance of the distinction between acceptability to realistic people and acceptability to viewpoints, which matters because public reason theories are committed to the former, but only deliver the latter, thereby failing to justify liberal norms. These theories therefore face a dilemma: abandon realistic people and lose normative appeal, or retain realism and fail to justify liberalism.

Introduction

At a basic level, public reason theories seek to ensure that there is a significant relationship between the political norms – the laws, constitutional principles, or social norms – that rule us, and the people they rule. That is why these theories are often justified by values like respect for persons, anti-paternalism, and anti-authoritarianism. The idea is that, for political norms to be legitimate, they must be acceptable to realistic, though perhaps more “reasonable,” versions of actual citizens.

If we accept this requirement, public reason theorists argue, then we must also embrace religious and ethical diversity because reasonable-but-realistic citizens will inevitably disagree about the good life. These realistic citizens will live in different places, prioritize different values, and have different cognitive abilities, so even those that sincerely aim at the truth will end up with diverse views. If we want to respect people with these realistic tendencies and motivations, as public reason says we must, then we must find political norms that would be acceptable to the diverse perspectives they end up with. And, since the only norms that seem to be broadly acceptable are non-sectarian ones such as political and social equality, public reason theories ultimately claim to justify familiar liberal values.

The trouble is that the same realistic qualities that cause reasonable people to disagree also make it quite hard to figure out what they could all accept. For example, let us say that I was a sincere but not fundamentalist Hindu, who arrived at my views due to my realistic, though reasonable, human qualities: my best friend in high school was a devout Hindu, and I came to adopt her views over time. But, due to my lack of formal training, I know more about certain applications of Hindu thought than I do about its principles. Since I am reasonable, however, I am entitled to respect, and you must find political norms that I could accept. How would you do it? Given that you know me to sometimes misidentify my own principles, be willing to change them in conversation with others, and struggle to apply them to novel questions, it might be hard to determine what I could accept. Or, worse yet, what I could actually accept might be inconsistent with my Hindu beliefs, which is what you are trying to accommodate.

It is therefore very sensible to ask which norms would be acceptable to a version of me that lacks all these flaws, and that is precisely what public reason theorists end up doing. Rather than ask which political norms would be acceptable to people with realistic qualities, such as those who make good faith errors in reasoning, theorists ask which political norms would be acceptable to particular *viewpoints*, which are collections of beliefs that are precisely identified, rank-ordered, and have clear rules of application. Theorists then argue that norms are “acceptable” to a viewpoint when the norms follow from, or are at least consistent with, the set of beliefs. By transforming realistic people into these rationalized viewpoints, public reason theorists can then answer the question of what I could accept, as a Hindu, by seeing whether a political norm is consistent with Hinduism’s core principles, which is much easier to do than by considering my realistic understanding, and, in one sense, more true to what I believe.

We can see how useful this transformation is in the formulations of the two most prominent public reason theorists, John Rawls and Gerald Gaus. Instead of considering how to justify norms to realistic citizens, with their ill-considered and often changing beliefs, Rawls begins¹⁷ by asking whether political norms can be acceptable to each of us *as* rational, liberal-democratic, citizens. As liberal-democratic citizens, we have a set of stable and identifiable beliefs – those that are “implicit in a liberal-democratic culture.” And, as rational citizens, those beliefs are internally consistent and applied correctly in evaluating political norms.

Since we are seeking public justifications for political and social institutions...we think of persons as citizens. This assigns to each person the same basic political position. In giving reasons to all citizens we don’t view persons as socially situated or otherwise rooted, that is, as being in this or that social class, or in this or that property and income group, or as having this or that comprehensive doctrine...Rather we think of persons as reasonable and rational, as free and equal citizens, with the two moral powers and having, at any given moment, a determinate conception of the good, which may change over time.¹⁸

Gaus, similarly, seeks political norms that would be acceptable to people, but only once they are represented by their “sufficient reasons,” which are identified, arranged, and applied rationally.

We can, I think, make progress in seeing when the Basic Principle of Public Justification is met by translating the problem of what rules all normal moral agents have sufficient reason to internalize, endorse, and follow into a deliberative problem of idealized “Members of the Public”¹⁹

¹⁷ After determining what the viewpoint of a citizen could accept, Rawls then considers whether that result would be acceptable to another set of viewpoints – the various religious and ethical “comprehensive doctrines” in a pluralist society.

¹⁸ John Rawls, *Political liberalism* (New York: Columbia University Press, 1993), p.481.

¹⁹ “I will argue that a Member of the Public is an idealization of some actual individual; a Member of the Public deliberates well and judges only on the relevant and intelligible values, reasons, and concerns of the real agent she represents and always seeks to legislate impartially for all other Members of the Public. The

who deliberate only on their sufficient reasons. We can understand these Members of the Public as the *rationalized* counterparts of real moral agents.²⁰ (Emphasis added)

However, while the move from realistic people to viewpoints facilitates and captures an appealing notion of acceptability, it is ultimately inconsistent with public reason's purpose. Switching from people to viewpoints would be like trying to find rules that your son will accept when he is panicked by imagining what he would accept when perfectly calm. Of course it would be easier to find such rules if you thought of him as being able to appreciate the consequences of his actions and work to mitigate them. Moreover, your son may even have "good reason to accept" those rules. But, that is simply not the task. Thinking of him as calm completely changes the answers you get. On a realistic view of his state of mind, it is possible that the only rule he can currently accept is a time-out, which is therefore the only acceptable norm.

Similarly, I will argue, the task that public reason theories are committed to is finding out which political norms would be acceptable to reasonable *people*, conceived of as collections of realistic human tendencies, limitations, and beliefs. Theorists therefore cannot remove these tendencies and limitations when thinking about what such people could accept; doing so leads to completely different answers and fails to respect people in the way that the theorists themselves claim that we must.

To show this, I will consider public reason theories, first in the abstract, and then, in later chapters, in the theories of Rawls and Gaus. In all three chapters, I will argue that the justification of public reason theories commits them to finding norms that are acceptable to reasonable, but realistic, people. Briefly, that justification consists of an explanation of why we should care about ethical pluralism, which must make reference to the realistic tendencies and limitations of reasonable people. I will then show that, for public reason theorists to claim that particular liberal norms are acceptable, they must turn those realistic people into viewpoints, which removes these tendencies and limitations. In short, to get the desired result, public reason theories must change the question that they are committed to answering.

Since the problem is the inconsistency between public reason's justification (which refers to realistic people) and its methods (which rely on viewpoints), the solution must be to either abandon the viewpoint method or to change the justification so it does not refer to realistic people. This presents a dilemma for public reason theories. If they abandon the viewpoint method, then they must show us how realistic people, with all their flaws, could still accept liberal political norms, which is a significantly more difficult task than they currently attempt.

moral rules (giving rise to imperatives) that a Member of the Public endorses are the ones that as a moral person she has sufficient reason to endorse." Gerald Gaus, *The Order of Public Reason* (Cambridge: Cambridge University Press, 2011) p.26.

²⁰ *Ibid.*, p.25-26.

The second option is to keep the viewpoint method but to argue that respecting a viewpoint is the right way to respect realistic people. In this paper, I consider and reject two strategies to bridge that gap between people and viewpoints, but there may be others. What is clear, however, is that public reason theories cannot continue to employ the viewpoint method without giving us a reason to care about what viewpoints, belief-systems, or doctrines “think.” Without such a reason, viewpoint theories lack the normative force of theories that aim to respect people, (mostly) as they are.

People, Not Viewpoints

All public reason theories begin with some version of the claim that legitimate political norms – laws, principles, or social rules – should be acceptable to “reasonable people.” However they define the term, the concept is meant to refer to an idealized version of actual people. The idea is that political norms do not have to be acceptable to actual citizens just as they are – some are crazy or immoral or simply cannot belong to a political community. But, political norms ought to be acceptable to some version of those citizens, even if we have to conceive of them with better beliefs, values, motivations, or cognitive abilities.

There is, however, a limit to how much we can idealize “reasonable people” and that limit is set by a part of the justification of public reason theories, which I will call the “diversity argument.” The diversity argument explains why we must find norms that are acceptable to people with diverse religious and ethical beliefs. In doing so, I will argue, the diversity argument also commits public reason theories to finding norms that are acceptable to people who are somewhat idealized, but nevertheless realistic in crucial ways, which I will describe.

To see how the diversity argument commits public reason theories to accommodating these reasonable-but-realistic people, consider the general form of the public reason argument:

1. To be legitimate, political norms must be acceptable to reasonable people
2. Reasonable people have diverse religious and ethical views of the good
3. So, political norms must be acceptable to diverse religious and ethical viewpoints
4. Certain basic liberal norms are acceptable to diverse religious and ethical viewpoints
5. So, these basic liberal norms fulfill at least one, and possibly the only, condition of legitimacy

The diversity argument comes in step two, which explains why even reasonable people will²¹ arrive at diverse religious and ethical views. It serves an essential purpose in any public reason argument. If we had no reason to think that reasonable people would disagree, then we are free to assume that reasonable people could all accept the singular

²¹ There are many possible formulations of this phrasing, including that reasonable people “will inevitably arrive”; “are very likely to arrive”; “will arrive without oppressive state intrusion”; “should be reasonably expected to arrive,” and so on. I do not believe any variant changes my argument.

truth. In that case, true norms would be the same as acceptable norms, and public reason would be the same as right reason. So, the diversity argument motivates the turn to publicly acceptable norms, rather than simply true ones.

In serving this essential purpose, however, the diversity argument *itself* causes a major problem for public reason theories. To explain why reasonable people will disagree about the good, public reason theories must attribute certain realistic human tendencies to reasonable people. Why is this required? Consider the alternative: let us say that theorists did not attribute such human tendencies to reasonable people and, instead, conceived of them as “correct” reasoners, in the way that a computer can “reason” correctly. That is, reasonable-and-correct people would be able to accurately identify their priorities, order them according to weight, and draw logical conclusions from that starting point. Like a computer, reasonable-and-correct people would not be omniscient – they would only have the evidence they can acquire – but they would always reason correctly using that evidence. If reasonable people reasoned correctly in these ways, how could we explain disagreement on fundamental ethical questions, as the diversity argument requires?

Broadly speaking, there are three kinds of claims that would explain disagreement among correct reasoners: 1) If there were multiple correct answers to fundamental ethical questions, then it would be no surprise if correct reasoners did not converge on one. 2) If reasoning alone could not lead you to the whole truth – if reason were indeterminate – then even correct reasoners will sometimes come up short and extend their reasoning in different ways. Or, 3) If correct reasoners were influenced by their different backgrounds or experiences, then reasonable people may diverge without reasoning incorrectly.²²

The first two claims seem like the kinds that would suit public reason theories and their concern for pluralism. And, in one sense, they do; they justify a concern for multiple views of the good. However, they do so by rejecting a fundamental element of many views of the good, namely, their insistence that there is a singular truth and human reason can determine it. Since this is the basis of at least some religious views²³ and many

²² There are two other explanations for disagreement that I will not discuss here. I leave out a skeptical interpretation because the claim that reasonable people are reasoning correctly about morality implies that there is, at least, something to be correct about. The other possible explanation is an interpretation of Rawls that one can draw from quotations such as this: “...there are many conflicting reasonable comprehensive doctrines with their conceptions of the good, each compatible with the full rationality of human persons, *so far as that can be ascertained with the resources of a political conception of justice.*”[Emphasis added] (Rawls 1993, p.135) Here, Rawls seems to argue that, if one *already accepts* the public reason requirement to make arguments that others can accept, then one will think that reasonable people would inevitably disagree about the good. This is because, to claim that all would converge on a single conception of the good, one would have to appeal to controversial claims, such as the truth of the Bible. However, since we start by accepting the public reason requirement, such controversial claims are not available and we must acknowledge that it is not reasonable to expect agreement. While this is a very interesting defense of public reason, it relies on assuming that the public reason requirement is correct from the beginning, so it is difficult to evaluate. In this paper, I will assume that public reason theories must justify that requirement to those who would deny it, which requires a diversity argument that, in turn, creates the inconsistency between people and viewpoints. As such, I will leave this defense of public reason aside. I thank Joshua Cohen for suggesting it.

²³ Such as, perhaps, a Catholic natural law theory.

ethical theories, a public reason theory cannot be grounded in a claim that denies this basis. If it were, then, rather than accommodating pluralism, a public reason theory would be insisting that its own foundations were true while certain religious and ethical foundations were false.²⁴

So, that leaves the third possibility: reasonable people are influenced by different backgrounds or experiences, which cause their views to diverge, despite their reasoning correctly. Most public reason theories, including the two major ones I will describe in this paper, cite some factor among these to explain the diversity of views among reasonable people. For example, public reason theorists often assume that reasonable people would disagree about the good because they grow up in different traditions, even if their reasoning proceeded correctly from those different starting points. Or, as in my earlier example, reasonable people may be influenced by some people or experiences more than by others, which leads to different views, even from the same starting point.

However, it is an open question whether these explanations of disagreement are actually consistent with reasoning correctly. One way in which they may be consistent is to say that reasonable people are not exposed to sufficient evidence for alternative traditions²⁵, so they are correct to retain their own. This interpretation is unsuitable to public reason theories for three reasons: 1) We are trying to explain disagreement on fundamental religious and ethical beliefs, which does not require esoteric or inaccessible evidence; testimony from prominent spokespeople is often sufficient. 2) Public reason theories assume reasonable people live under conditions of pluralism, free association, and free speech, so they are likely to be exposed to evidence for a variety of traditions. And, 3) Given that public reason theories must see this divergence as inevitable (or that it is unreasonable to expect otherwise) in any liberal society, the ‘exposure to different evidence’ explanation seems too contingent to serve their purposes.

So, assuming that reasonable people are exposed to evidence for a variety of traditions, could it be correct to retain the one they grew up with or had salient experiences with? Is that why disagreement is inevitable, even among correct reasoners? This is a matter of some debate among epistemologists. Some have argued that sticking with the evidence that one acquired growing up is perfectly rational, even when confronted by alternative views, because one *should* evaluate the probability of alternatives in light of what one already believes. Moreover, it is often correct to change one’s views in response to people or experiences that present you with new evidence; this can be an epistemic virtue. On the other hand, one could argue that these explanations of disagreement result from incorrect reasoning or epistemic flaws. In the short term, it is an epistemic flaw to

²⁴ Susan Mendus and David Enoch have made similar arguments regarding impermissible bases of public reason theories generally. See Susan Mendus, “Pluralism and Scepticism in a Disenchanted World,” *Pluralism: The Philosophy and Politics of Diversity*, ed. by Maria Baghramian and Attracta Ingram (New York: Routledge, 2000) pp. 103-119. And: David Enoch, “Political Philosophy and Epistemology: The Case of Public Reason,” *Oxford Studies in Political Philosophy* (Volume 3) ed. by David Sobel, Peter Vallentyne, and Steven Wall (Oxford: Oxford University Press, 2017) pp.132-165.

²⁵ *Ibid.*, p.139.

weight evidence based on the order in which you receive it, so it may also be incorrect to favor the tradition you learn early on over alternatives that you encounter later.²⁶

Regardless of whether disagreement is a correct response to different backgrounds or experiences, however, we can at least conclude that this explanation of disagreement attributes certain realistic human qualities to reasonable people. Under this explanation, reasonable people acquire evidence at different times in their lives, which influences the weight they give that evidence and therefore influences their conclusions. Or, under this explanation, reasonable people have more and less salient experiences with people and experiences, which influence the weight given to those sources of evidence. These are ways that humans, and not computers, reason.

Any diversity argument – any explanation of why people inevitably disagree on religious and ethical questions – must conceive of reasonable people as having such realistic tendencies and (depending on your epistemological view) even limitations. Moreover, for the diversity argument to explain why reasonable people end up with a wide range of religious and ethical views, it must attribute *significant* realism to reasonable people. After all, the diversity argument must not only explain why reasonable people disagree slightly about ethics; it must explain why they endorse views as different as Buddhism and Utilitarianism.

So, for the diversity argument to do its job, it must conceive of reasonable people as realistic human reasoners. And, indeed, that is what we will see from diversity arguments such as John Rawls' burdens of judgment, which conceives of reasonable people as challenged by difficult moral questions and malleable when exposed to different arguments and experiences. These realistic qualities are why reasonable people hold different religious and ethical views, and why political norms should therefore take this pluralism into account.

With this conception of reasonable-but-realistic people in mind (hereafter just “reasonable people”), let us reformulate the first half of the public reason argument:

1. To be legitimate, political norms must be acceptable to reasonable people
2. Reasonable people have certain realistic human qualities and/or limitations that cause them to arrive at diverse religious and ethical views of the good
- 2a. So, political norms must be acceptable to people with certain realistic human qualities and/or limitations

²⁶ As for the influences of people or experiences on your conception of the good, this is only an epistemic virtue when we correctly assess those people or experience as sources of evidence. If we assumed that all reasonable people did that, however, then we would be back to explaining disagreement by suggesting that people are exposed to different evidence, which they correctly interpret. Instead, this explanation of disagreement seems to rely on the inevitable influence of salient experiences on our beliefs, and salience is unlikely to track epistemic value.

3. So, political norms must be acceptable to diverse religious and ethical viewpoints

Once we acknowledge the diversity argument, it is clear that public reason theories are committed to the claim that political norms must be acceptable to people with certain realistic qualities and possibly limitations. And we can now see that there is a sizable gap between steps 2a and 3. Finding norms that are acceptable to viewpoints is quite different from finding norms that are acceptable to realistic people. This is because turning people into viewpoints removes the very human tendencies and limitations that public reason theories are committed to accommodating.

To see this, consider the characteristics that a viewpoint must have for us to determine whether a political norm would be “acceptable” to that viewpoint. First, it must articulate a set of fundamental beliefs or values, so that we know whether those beliefs are compatible with the proposed norm or not. Second, it must rank those beliefs according to weight so that we know which beliefs take priority in cases of conflict. Third, it must assume that these ordered fundamental beliefs do not change in the course of deliberation or when their implications are revealed. And fourth, it must contain sufficient epistemic norms so that we know exactly how to interpret, extend, or otherwise apply those beliefs to the norms in question. Taken together, viewpoints must be precisely articulated, ordered, and fixed beliefs, which adhere to valid epistemic norms.

Given these characteristics of viewpoints, the requirements of 2a and 3 are now very different. The realistic qualities that lead individuals to rank their beliefs, in part, according to the order in which they were acquired or the salience of their experiences will lead to different results than the way in which the claims of a viewpoint will be ranked, whether that is according to perceived importance or through a process of reflective equilibrium. In addition, the diversity argument tells us that reasonable-but-realistic people change their fundamental views in response to environments and information, while viewpoints do not. Finally, if ethical diversity comes about due to the cognitive or epistemic limitations of reasonable people, then they will not necessarily correctly identify their fundamental beliefs or adhere to valid epistemic norms. In short, political norms that would be acceptable to people with realistic qualities and limitations, such that they diverge into views as different as Buddhism and Utilitarianism, are going to be very different than the norms that are acceptable to internally consistent, clearly articulated, unchanging, and rank-ordered viewpoints.

Public reason theorists do not seem to notice the significant gap between 2a and 3, let alone defend the move between them. Nevertheless, there are two justifications for the move that seem appealing and may underlie their arguments. First, there is reason to think that moving from people to viewpoints retains a relevant notion of respect. Here, respect still requires finding norms that realistic people could accept but, by turning people into viewpoints, we actually meet a more demanding standard in answering that question. It is more demanding to ask whether fixed viewpoints could accept a political norm, since we could otherwise claim that more realistic people would *change* their views to accept it. A model that turned malleable people into fixed viewpoints would be

more respectful of pluralism, make it harder to find norms that all could accept, and therefore, ironically, produce a more realistic result.

Second, one might argue that it is more faithful to what people believe to conceive of them as systematic and consistent reasoners – as viewpoints – than as realistic and limited people. To return to my previous example, it may be more faithful to my Hindu identity to represent me based on the collective wisdom on Hinduism, rather than my own piecemeal impressions. Perhaps this is the true meaning of “respect”: to find norms that would be acceptable to the best thinking of a doctrine that I aspire to act on.

While these two arguments certainly explain the appeal of a viewpoint model, they ultimately fail to justify the move from 2a to 3. It is plausible that a *certain* kind of viewpoint model would be more demanding than a more realistic representation. For example, if each religious or ethical viewpoint were sociologically accurate, including its most doctrinaire elements, it would be impressive (and surprising) to see that they could all converge on liberal norms. However, as we will see in the sections on Rawls and Gaus, that is not how the viewpoint model is deployed.

Three features of viewpoints make it far easier to connect any input with a desired output. First, viewpoints allow theorists to decide how to represent people’s views – at what level of abstraction, with what content, and based on which identity. Unlike more realistic formulations, these starting points do not have to be part of a person’s realistic self-understanding. Second, viewpoints fix priority rules, where realistic people struggle with consistency. This allows theorists to claim that people will reliably take the correct deliberative path from the starting point to the conclusion. And third, viewpoints can make sophisticated connections between claims, which more realistic people may not be able to, allowing theorists to argue that, if people truly understood the implications of their starting points, they would arrive at the right conclusion. These features make it easier, not harder, for theorists to find norms that reasonable people can accept.

These same features explain why viewpoints are often not more faithful to what people really believe. The “best thinking of the doctrine that I aspire to” may very well represent what I would accept if only I were reasoning better, as, for instance, when theorists say that an aspiring Christian would accept the freedom of conscience because her faith requires a voluntary belief in God. However, the “best thinking” approach also allows theorists to go further and claim that all reasonable people aspire to be liberals, and the best thinking on liberalism is that it takes priority over all other identities. Or, further still, that reasonable people aspire to be agents in the world and the best thinking on agency requires a Kantian liberalism, which is therefore “acceptable” to them. The problem with defining respect as the “best thinking of the doctrine I aspire to” is that it places no principled limits on what could be acceptable to me. Without any limits on starting points or reasoning capacities, there is no way to ensure a significant relationship between the person and the norms she is said to accept. In other words, this conception of respect lacks the very qualities that the viewpoint model removes.

Without an argument to justify the move from 2a to 3, public reason theorists cannot fulfill their commitment to finding norms for realistic people by finding norms for viewpoints. As I will argue, this gap poses a problem for even the best public reason theories, those of Rawls and Gaus. Both argue that norms must be acceptable to diverse groups because of the realistic tendencies and limitations of reasonable people, thereby committing to finding norms that would be acceptable to such people. Instead of seeking such norms, however, Rawls and Gaus argue that liberal norms are acceptable to viewpoints, which intentionally remove those human limitations. Without a justification for the move from people to viewpoints, the liberal norms they identify do not meet their own tests of public justification. We are therefore still in search of norms that could be acceptable to reasonable people, in all their humanity.

Conclusion

Public reason theories are faced with an impossible task: to show that people who think very differently could nevertheless think the same. They solve this riddle by equivocating on what it means to “think differently.” When they explain why political norms should accommodate diverse views of the good – their “diversity argument” – they acknowledge that people have different levels of reasoning ability and make unpredictable choices. However, when they argue that we can all accept the same liberal norms, they insist that we “think differently” in only one sense: we hold a variety of doctrines, which are fixed, precise, and ordered, enabling them to converge on the same norms. There is therefore an internal contradiction in public reason theories; their diversity argument commits them to the first sense of “think differently,” while their conclusion requires the second.

There are two ways of solving this problem, which are both worth pursuing, although neither is easy. The first option is to try to solve the problem I have identified by providing a justification for the move from people to viewpoints. I have considered and rejected two possibilities along these lines, but perhaps there are others. Any such justification must explain why acceptability to viewpoints preserves the sense of respect that motivates public reason in the first place.

The second way is to reject the viewpoint model entirely and to seek political norms that could be acceptable to more realistic people. The contours of such a theory will depend on how realistic reasonable people must be, which in turn depends on the diversity argument the theory employs. However, it is safe to say that, if liberal norms can be acceptable to realistic people, they are likely to be considerably thinner than the more robust set that Rawls and Gaus defend.

Chapter Two – Judgments versus Viewpoints: Minimal Idealization in John Rawls

In the previous chapter, I argued that the common source of the inconsistency between the justifications of public reason theories and their idealization is a conflict between “realistic people” and “viewpoints.” In this chapter, I describe how that conflict plays out in John Rawls’ theory. Rawls justifies his theory, in part, using a “diversity argument” that explains why reasonable people would inevitably disagree about the good, which necessitates the turn to public reason. That diversity argument – the burdens of judgment – commits the theory to seeing reasonable people as realistic in important ways. However, like other theorists, when Rawls explains how reasonable people could all accept liberal norms, he turns them into fixed and consistent viewpoints, which abstracts from the realism that the justification commits him to.

Rawls’ Diversity Argument: The Burdens of Judgment

Rawls’ diversity argument – which tells us why political norms must be acceptable to a range of religious and ethical views, rather than simply true, just, or sound – emerges from what he describes as the challenges that even reasonable people face in their ethical reasoning: the “burdens of judgment.”

- (a) “evidence – empirical and scientific – bearing on the case is conflicting and complex”
- (b) “we may disagree about [considerations’] weight”
- (c) “concepts...are vague and subject to hard cases”
- (d) “the way we assess evidence and weigh moral and political values is shaped by our total experience”
- (e) “there are different kinds of normative considerations of different force on both sides of an issue”²⁷

How do these “burdens” explain why political norms must accommodate diverse ethical views? One obvious possibility is that reasonable people disagree because they are flawed reasoners, so we should expect them to struggle with hard questions and therefore arrive at different answers. Since Rawls has already argued that political norms must be acceptable to reasonable people, and now argues that reasonable people inevitably diverge due to their flaws, we can conclude that political norms must be acceptable to diverse views.²⁸

²⁷ John Rawls, *Political liberalism* (New York: Columbia University Press, 1993) p.56-7.

²⁸ Leif Wenar suggests that Rawls’ argument can work without the burdens of judgment, which would evade my critique. However, for reasons described in the previous section, I take some diversity argument to be a crucial element of the public reason argument. Without it, there is no reason to think that “reasonable people,” an invented category, would have different views, even if current citizens do. Without something like the burdens of judgment, one might argue that, while political norms must be acceptable to reasonable people, reasonable people can all accept a Kantian ethics. See Leif Wenar, “Political Liberalism: An Internal Critique,” *Ethics*, 106, pp.32-62.

However, Rawls rejects this interpretation of the burdens of judgment and of reasonable people. In his view, the burdens of judgment are supposed to explain disagreement, even among “correct (and conscientious)” reasoners.

We might suppose, say, that most people hold views that advance their own more narrow interests; and since their interests are different, so are their views. Or perhaps people are often irrational and not very bright, and this mixed with logical errors leads to conflicting opinions. But while such explanations explain much, they are too easy and not the kind we want. We want to know how reasonable disagreement is possible, for we always work at first within ideal theory... An explanation of the right kind is that the sources of reasonable disagreement – the burdens of judgment – among reasonable persons are the many hazards involved in the correct (and conscientious) exercise of our powers of reason and judgment...²⁹

If the burdens are not claiming that reasonable people make mistakes on their way to different views, then they must be offering an explanation of how correct reasoners – those who are rational, bright, and do not make logical errors – will nevertheless diverge. There are three possible interpretations of the burdens that could explain how divergence is consistent with correct reasoning.

First, the burdens could claim that reasonable people disagree because they have access to different evidence. In other words, reasonable people are correct, but not omniscient, reasoners. And second, the burdens could claim that reasonable people disagree because they are influenced by different people and experiences. Both of these interpretations are consistent with burden (d), which Rawls describes more fully as:

...the way we assess evidence and weigh moral and political values is shaped by our total experience, our whole course of life up to now; and our total experiences must always differ. Thus, in a modern society with its numerous offices and positions, its various divisions of labor, its many social groups and their ethnic variety, citizens’ total experiences are disparate enough for their judgments to diverge...³⁰

Here, Rawls suggests many possible variations in conditions that could cause judgments to differ – experiences, occupations, social groups, and so on. And it is certainly true that a coal miner has access to different evidence than a banker on certain issues, so each could rightly reason to different conclusions based on the evidence they have. But, in interpreting the burdens, we must keep in mind the scope of disagreement among reasonable people; they differ not just by political opinions or party, nor by the weight they place on this value rather than that, but on fundamental ethical belief-systems. So, our different social positions or experiences must explain why we have access to different evidence on the good life, which seems far less plausible. Non-Christians do not reject

²⁹ John Rawls, *Political liberalism* (New York: Columbia University Press, 1993), p.55-56.

³⁰ *Ibid.*, p.56-57.

Christianity because they are not aware that the Bible exists or that it is supposed to be the word of God. They reject it because they deny that it is the word of God; awareness is not the issue.

A better explanation for divergence is that reasonable people learn their own conception of the good early on in life, and then they often stick with that conception, even when they encounter other faiths and philosophies. For this explanation to be consistent with correct reasoning, Rawls must then argue that it is perfectly rational to stick with the view you learn early in life.

As I have argued, whether it is rational is a matter of some debate, but there are good reasons to think it is not. After all, the claim that reasonable people acquire their fundamental ethical beliefs from their initial communities implies that reasonable people are responsive to new arguments, evidence, and environments. Unless reasonable people are then isolated from further experiences, however, they are likely to continue to encounter new arguments and evidence from other traditions. Of course they will evaluate new ideas in relation to the old, but given a lifetime of new encounters, why should we expect correct reasoners to stick with their original views?

Clearly, the answer cannot be that reasonable-and-correct people are more responsive to new ideas at the beginning of their lives because that would imply an epistemic flaw – either they were too suggestible early in life or too closed-minded late. Perhaps the answer is that it is impossible to convey some evidence for traditions other than one's own because they require some direct perception or insight.³¹ However, it is not clear whether I have more reason to trust insight revealed directly to me than others would have, when told of my experience. And, in any case, most fundamental beliefs are acquired through testimony, which everyone has access to. In short, the claim that reasonable people disagree about the good because they start with different beliefs may still attribute some epistemic flaw to reasonable people.

Even if burden (d) does not imply a flaw in reasonable people, however, it is hard to make sense of the other burdens without doing so. These other burdens (a, b, c, and e) offer a third possible argument for why reasonable people would diverge, despite correct reasoning. That argument is the following: Because of the inherent difficulties in ethical reasoning – vague concepts, hard cases, complex evidence, and so on – we cannot expect even correct reasoners to overcome them and arrive at the right conclusion.

In interpreting these other burdens, we must remember that, as I previously argued, the explanation of diversity cannot be based in a controversial view of the good, which Rawls acknowledges:

[Burden (e)] could, of course, be described from within a comprehensive moral doctrine, as when [Thomas] Nagel says that there are basic conflicts

³¹ David Enoch, "Political Philosophy and Epistemology: The Case of Public Reason," *Oxford Studies in Political Philosophy* (Volume 3) ed. by David Sobel, Peter Vallentyne, and Steven Wall (Oxford: Oxford University Press, 2017) pp.132-165.

of value in which...values are incomparable: they are each specified by one of the several irreducibly different perspectives within which values arise...Nagel's discussion is not implausible, but a political conception tries to avoid, so far as possible, disputed philosophical theses and to give an account of the burdens of reason that rests on plain facts open to all.³²

If the burdens cannot rest on a controversial philosophical view, then this third argument for disagreement (among correct reasoners) is not open to Rawls. Rawls cannot claim that correct reasoners *cannot* (or even "are unlikely to") overcome these burdens and get to the right answer, because if even correct reasoners cannot resolve them, then this implies one of three controversial theses: 1) There are multiple true answers; 2) The truth is indeterminate; or 3) The best human reasoners cannot be expected to arrive at the truth. The first two, we have already seen, are based on controversial ethical theories. And, most religious and ethical adherents would reject the third one, since they believe that they *have* overcome these burdens and that correct reasoning *is* likely to reach the right conclusion.³³ From their point of view, other reasonable people fail to reach the right conclusion because those others make (understandable) mistakes.

So, Rawls' claim that the burdens of judgment explain how correct reasoners would inevitably disagree does not hold. Burden (d) may attribute some epistemic flaw to reasonable people by suggesting that they give disproportionate weight to early experiences. And, if reasonable people were correct reasoners, then burdens (a), (b), (c), and (e) – which cite the difficulties of moral reasoning – would rely on a controversial view of the good by suggesting that even the best reasoners cannot get at the truth. So, if Rawls' diversity argument – the burdens of judgment – is inconsistent with the conception of reasonable people as correct reasoners, then the burdens must explain diversity, at least partly, by attributing epistemic flaws to reasonable people.³⁴ This may be what Rawls refers to in quotations like this:

Many conceptions of the world can plausibly be constructed from different standpoints. Diversity naturally arises from our *limited powers* and distinct perspectives...(Emphasis added.)³⁵

The burdens are therefore not (or not just) about the nature or conditions of ethical reasoning but about particular reasoners – what they can and cannot do. This means we

³² John Rawls, *Political liberalism* (New York: Columbia University Press, 1993) p.57, fn.10.

³³ David Enoch, "Political Philosophy and Epistemology: The Case of Public Reason," *Oxford Studies in Political Philosophy* (Volume 3) ed. by David Sobel, Peter Vallentyne, and Steven Wall (Oxford: Oxford University Press, 2017) pp.132-165.

³⁴ David Enoch also argues that the burdens of judgment attribute epistemic flaws to reasonable people: "Seeing that at least some of the burdens of judgment seem to highlight what are still epistemic *flaws*...it is not clear how relying on them is consistent with doing ideal theory." David Enoch, "Political Philosophy and Epistemology: The Case of Public Reason," *Oxford Studies in Political Philosophy* (Volume 3) ed. by Sobel, Vallentyne, and Wall (Oxford: Oxford University Press, 2017) pp.28, n.50.

³⁵ John Rawls, "Kantian Constructivism in Moral Theory," *Journal of Philosophy* 77, pp.515-542. Quoted from Susan Mendus, "Pluralism and Scepticism in a Disenchanted World," *Pluralism: The Philosophy and Politics of Diversity*, ed. by Maria Baghramian and Attracta Ingram (New York: Routledge, 2000) pp. 103-119, at p.108.

can infer the epistemic tendencies and flaws of reasonable people from the burdens they face. What are these tendencies and flaws? First, the fact that reasonable people diverge because of their different life experiences tells us that reasonable people are able to change their fundamental beliefs based on new environments, arguments, and information. To what degree are they responsive to new environments? If this burden is supposed to explain divergence into views as different as Buddhism and Utilitarianism, then new experiences must have *considerable* influence on the fundamental attitudes of reasonable people.³⁶

Second, the burdens tell us that reasonable people have difficulty weighing evidence and values in hard cases, which is why we cannot expect them to always arrive at the right answer. This difficulty implies that, in hard cases, reasonable people lack a definite guide to making trade-offs and ranking priorities. So, on fundamental questions, we cannot expect reasonable people to have clear priority rules, making them somewhat unpredictable.

And, third, this degree of responsiveness and unpredictability of reasonable people means that reasonable people will sometimes fail to draw the right conclusions from their premises. The influence of their total life experiences on reasoning and their difficulty with reading evidence and interpreting concepts means that, at times, they will fail to arrive at conclusions that are consistent with their priorities. So, even if reasonable people shared a background and fundamental values, we should not expect them all to end up with conclusions that are consistent with each other.

In short, the burdens imply that reasonable people are, sometimes and on hard cases, responsive, unpredictable and inaccurate reasoners. If these are the qualities that the burdens attribute to reasonable people, then we can now clarify Rawls' public reason argument:

1. To be legitimate, political norms – laws, principles, or social rules – must be acceptable to reasonable people
2. Reasonable people are sometimes responsive, unpredictable and inaccurate reasoners, which is at least partly³⁷ why they arrive at diverse religious and ethical views of the good
3. So, political norms must be acceptable to sometimes responsive, unpredictable and inaccurate reasoners who therefore end up with diverse views of the good

³⁶ As Rawls discusses, it may not just be major life experiences that influence reasonable people's fundamental views, but also political deliberation with others: "When citizens deliberate, they exchange views and debate their supporting reasons concerning public political questions. They suppose that their political opinions may be revised by discussion with other citizens; and therefore these opinions are not simply a fixed outcome of their existing private or nonpolitical interests." John Rawls, "The Idea of Public Reason Revisited," *The University of Chicago Law Review* 64(3): p. 772.

³⁷ As mentioned, part of the explanation may also be a rational response to different upbringings.

This third claim is the task that Rawls is committed to through his diversity argument, the burdens of judgment.

Rawls' Viewpoint Model

However, Rawls does not show us norms that would be acceptable to people with such tendencies. Instead, he argues that liberal principles of justice would be acceptable to viewpoints, which have the very opposite characteristics; they are fixed, predictable, and consistent. To see this, consider how Rawls' solution to the problem of reasonable disagreement is supposed to work.

Rawls argues that, despite coming from different religious and ethical traditions, reasonable citizens also *share* a background – that of being a citizen in a liberal-democracy. They grow up and work under the same political institutions and they are taught the same civic values, which influence their attitudes just as their private traditions do. So, Rawls sensibly argues, if reasonable citizens simply channeled their shared civic identity instead of their disparate private identities, they could arrive at a consensus on acceptable political reasons and, to some degree, on the norms they support. Whether or not this consensus were optimal from those citizens' all-things-considered points of view, at least the consensus would be acceptable to all citizens *as* citizens.

...in public reason, ideas of truth or right based on comprehensive doctrines are replaced by an idea of the politically reasonable addressed to citizens as citizens. This step is necessary to establish a basis of political reasoning that all can share as free and equal citizens... Since we seek a shareable public basis of justification for all citizens in society, giving justifications to particular persons and groups here and there until all are covered fails to do this.³⁸

What is a “basis of political reasoning that all can share as free and equal citizens?” Among other things, it is a set of reasons that do not rely on any religious or ethical tradition for support.³⁹ Because citizenship does not depend on having a particular view of the good, the reasons that one can accept as a citizen – and not as a Christian or Utilitarian – cannot rely on a particular view of the good.⁴⁰ For example, if my Hindu views told me that eating beef should be prohibited, I would have to ask myself whether other citizens could accept that norm as citizens. Since the only reason to support the prohibition is based on Hindu doctrine, other citizens cannot have any reason to support it as citizens. In proposing a political norm, I would therefore have to prioritize alternative norms that others could accept over the norm that my religion requires.

³⁸ Ibid., p.481-482.

³⁹ These are reasons that “it is at least reasonable for others to accept... as free and equal citizens, and not as dominated or manipulated, or under the pressure of an inferior political or social position.” Ibid., p.446. In other formulations: reasons that “can be presented independently from comprehensive doctrines... can be worked out from fundamental ideas seen as implicit in the public political culture of a constitutional regime, such as the conceptions of citizens as free and equal persons, and of society as a fair system of cooperation.” Ibid., p.453.

⁴⁰ John Rawls, *Political liberalism* (New York: Columbia University Press, 1993) p. 223-225.

However, that only tells us the norms that other citizens cannot accept; what *can* be the basis for public reasoning? Public reasons, Rawls argues, are drawn from fundamental ideas that are said to be implicit in the public political culture and the workings of liberal-democratic institutions.⁴¹ Roughly, these ideas revolve around conceptions of freedom, fairness, and equality. I must therefore ground my arguments in these ideas before reasoning my way to conclusions that are consistent with them. Rawls even argues that my reasoning must follow particular “guidelines of inquiry,” which are publicly acceptable ways of reasoning.⁴² If I fail to follow these guidelines, or even fail to reason well enough to arrive at conclusions consistent with my premises, then I have not accomplished my (or Rawls’) goal of grounding political norms on reasons that all citizens could accept as citizens.

So, for Rawls, the possibility of reasonable citizens arriving at liberal political norms depends on their willingness to, at least initially,⁴³ restrict their deliberation to shared facts,⁴⁴ theories,⁴⁵ and values – to a single, shared identity – rather than to rely on their diverse private traditions.⁴⁶ He then imagines those citizens as reasoning “correctly” from this shared identity to the conclusions that rationally follow.⁴⁷ Because this shared identity is based in the public political culture, and because that political culture is liberal, we can then say that the only norms that are acceptable to all citizens as citizens are liberal norms – norms consistent with the values of freedom, fairness, and equality.⁴⁸

For this solution to work, however, reasonable citizens must have the following characteristics. First, these citizens must generally prioritize their identity as citizens over all their private identities – as mothers and Muslims and Kantians and Quakers – when

⁴¹ Ibid., p.223. See also n. 17.

⁴² “Now it is essential that a liberal political conception include, besides its principles of justice, guidelines of inquiry that specify...principles of reasoning and rules of evidence in the light of which citizens are to decide whether substantive principles properly apply and to identify laws and policies that best satisfy them.” Ibid., p.223-224.

⁴³ There is a further requirement of an overlapping consensus in which citizens inhabit their private identities, but I will return to this later.

⁴⁴ “...ascertainable evidence and facts open to public view...” John Rawls, “The Idea of Public Reason Revisited,” *The University of Chicago Law Review* 64(3): p. 786.

⁴⁵ “...nor to elaborate economic theories of general equilibrium, say, if these are in dispute...plain truths now widely accepted, or available, to citizens generally.” John Rawls, *Political liberalism* (New York: Columbia University Press, 1993) p.225.

⁴⁶ “To put the case another way, we distinguish between the point of view of people as citizens and their point of view as members of families and of other associations.” John Rawls, “The Idea of Public Reason Revisited,” *The University of Chicago Law Review* 64(3): p.790.

⁴⁷ “Public justification is not simply valid reasoning, but argument addressed to others: it proceeds correctly from premises we accept and think others could reasonably accept to conclusions we think they could also reasonably accept.” Ibid., p.786.

⁴⁸ More particularly, any liberal conception has “three things: first, it specifies certain basic rights, liberties, and opportunities (of the kind familiar from constitutional democratic regimes); second, it assigns a special priority to these rights, liberties, and opportunities, especially with respect to claims of the general good and of perfectionist values; and third, it affirms measures assuring all citizens adequate all-purpose means to make effective use of their basic liberties and opportunities.” John Rawls, *Political liberalism* (New York: Columbia University Press, 1993) p.223.

the two conflict. Put another way, reasonable citizens must prioritize the importance of public reasoning over all other values, when there is a conflict. Second, their fundamental shared values, beliefs, and priorities must remain largely fixed. If the perspective of citizen calls for them to hold values like freedom, fairness, and equality, they cannot be persuaded to abandon any of those core grounds, or their conclusions will no longer be acceptable to all. And, third, reasonable citizens must have sufficient reasoning ability and imperviousness to non-rational influences to follow the guidelines of inquiry and arrive at conclusions that are consistent with their shared grounds.

However, fixed sets of fundamental beliefs and values with clear priority rules and consistent conclusions – these are the characteristics of *viewpoints* and not of reasonable people (as implied by the burdens of judgment). Rawls is considering reasonable people as if they were fixed citizen-viewpoints, rather than somewhat responsive, unpredictable, and limited reasoners. If Rawls' solution can only show liberal principles to be acceptable to such viewpoints, then he has failed to meet the task he set for himself.

Of course, I have so far left out a seemingly crucial piece of the puzzle, the “overlapping consensus” (OC) requirement, which ensures that the norms acceptable to citizens as citizens are also acceptable to them based on their private traditions – as Mormons and Utilitarians and so on. The idea is that, after citizens have arrived at political norms based on their citizen-viewpoints, they then evaluate those norms from their private perspectives. If the norms remain acceptable then, and only then, are the norms fully justified.⁴⁹

One could argue that the requirement of acceptability to private viewpoints reintroduces the realistic human qualities of reasonable people in two ways. First, the OC requirement might imply that reasonable people *are* able to change their fundamental beliefs in response to new ideas, both public and private. For example, one might say that the political norms that are fully justified are the product of a negotiation between each reasonable person's private and public perspectives. Reasonable people initially propose the norms by channeling the viewpoint of a citizen, but when they make sure that these norms are acceptable to their private viewpoints, they alter and adapt both private and public perspectives until both are in harmony (or “reflective equilibrium,” as Rawls calls it). In this way, the overlapping consensus requirement implies that reasonable people can change their fundamental views.⁵⁰

⁴⁹ John Rawls, “Reply to Habermas,” *The Journal of Philosophy*. 92(3), p.143.

⁵⁰ There is another sense in which the OC might imply the malleability of fundamental beliefs. Private viewpoints may be fixed, but this merely represents the way in which reasonable people's views are also fixed for a period of time before changing into something else. But, since that something else will be just another viewpoint, the model can account for malleability by requiring acceptability to all viewpoints. However, even if this division of reasonable people into neat doctrines were faithful to the way people change (which seems unlikely), this does not change the dilemma that Rawls faces. Either the private viewpoints have an influence on the public consensus or they do not. The former implies that reasonable people have realistic qualities, but produces illiberal norms. The latter produces liberal norms, but removes the realistic qualities.

Second, one might argue that the OC requirement implies that reasonable people are not required to prioritize the citizen-viewpoint over their other values and identities. Perhaps private and public identities change each other, as described above, or both must be able to jointly accept the political norms. In either case, as long as reasonable people are entitled to reject or alter the public consensus based on their private views, neither identity seems to have priority over the other, which means reasonable people do not have to be as predictable and consistent as viewpoints to accept Rawls' conclusions.

However, as Jonathan Quong argues, this conception of the overlapping consensus (and therefore the implied picture of reasonable people) puts Rawls' theory in a bind.⁵¹ If the OC requirement implies a mutual negotiation between reasonable people's private and public beliefs, and allows for the rejection of the public consensus based on private views, then the public consensus will likely be illiberal, because it is influenced by illiberal private views. On the other hand, if the public consensus is not influenced by private views, and reasonable people are expected to change or de-prioritize their private beliefs when they conflict with the public consensus, then the overlapping consensus requirement has no effect on political norms. And if it has no effect, then it does not reintroduce the realistic qualities of reasonable people, as this objection claims. In other words, this bind forces Rawls to choose between failing to justify liberal norms to reasonable people, or justifying liberal norms to viewpoints.

It is possible that Rawls would be content with acceptability to viewpoints. He seems to believe that fulfilling such a requirement would be more important to political stability than acceptability to reasonable people:

Moreover, these doctrines have their own life and history apart from their current members and endure from one generation to the next. The consensus of these doctrines is importantly rooted in the character of various associations and this is...crucial in providing a deep and enduring basis for its social unity.⁵²

Meanwhile, the public reason theorists inspired by Rawls are clear in choosing acceptability to viewpoints. To avoid the problem of the failure to prioritize public reasons, Quong concedes that we must address viewpoints:

We must be aware, however, that it is possible (indeed very likely) that citizens who are generally reasonable may sometimes make unreasonable demands, and vice versa. Strictly speaking, therefore, it is not the unreasonable *citizen* who is excluded by public reason, but rather unreasonable *views or claims*.⁵³

⁵¹ Jonathan Quong, *Liberalism Without Perfection* (Oxford: Oxford University Press, 2011) p.166-167.

⁵² John Rawls, "Reply to Habermas," *The Journal of Philosophy*. 92(3), p.145-6.

⁵³ Jonathan Quong, *Liberalism Without Perfection* (Oxford: Oxford University Press, 2011) p. 291.

And, to avoid the issue that reasonable people may be inconsistent reasoners, Gerald Gaus argues that we must consider only the “rational natures” of reasonable people, rather than their “reasoning found in common sense.”⁵⁴

When another demands that you comply with a rule, she is demanding that you do what you have sufficient reasons to do; she is appealing to your rational nature, not demanding that you put it aside. She must be saying: “You have reasons to comply that you are ignoring. My demand is not simply a demand that you live as I see fit, but as you would see fit if you adequately employed your reason.”⁵⁵

Whether Rawls and his successors choose reasonable people or viewpoints, however, they run into problems. Justification to reasonable people fails to validate liberal norms, while justification to viewpoints is inconsistent with the aim of the theory, set by the diversity argument. As we will see, the other major public reason theory, Gaus’ view in *The Order of Public Reason*, faces the same dilemma.

⁵⁴ John Rawls, *Political liberalism* (New York: Columbia University Press, 1993) p. 224.

⁵⁵ Gerald Gaus, *The Order of Public Reason* (Cambridge: Cambridge University Press, 2011), p.263.

Chapter Three – Extraordinary People: Moderate Idealization in Gerald Gaus

In this chapter, I argue that there is a version of Gerald Gaus' theory that escapes my previous critique and remains internally consistent across the four elements of public reason theories described in the Introduction. However, such a theory would only justify very "thin" political norms, which would not ensure the liberal freedoms that Gaus tries to defend. So, when Gaus argues that his model would justify "thicker" liberal and libertarian norms, he must re-describe two elements of the theory in ways that push it into inconsistency with his underlying justification.

Introduction

In the *Order of Public Reason*, Gerald Gaus offers a compelling justification for why people with different views of the good ought to determine social and political norms. In fact, he offers several compelling justifications, grounded in anti-authoritarianism, equality, positive liberty, and our moral practices. In this chapter, I will focus on the justification from our moral practices because that is the most well-developed of his accounts. I will argue that this justification tells us why we ought to accommodate people of diverse views in the first place, which also tells us who and how we ought to accommodate them. Specifically, it tells us that we must accommodate a very broad group of people – ordinary, "moral" people – and their very diverse views. This is a strength of his theory since it promises to offer a high level of responsiveness to citizens' attitudes and it is, along with many other elements of his theory, consistent with his underlying justification.

However, the breadth of his group of "reasonable people" forces his theory to confront the familiar public reason dilemma. A broad and diverse group of reasonable people will have a hard time accepting any norms at all, let alone liberal ones. To solve this problem, Gaus employs two other strategies over the course of his book, both of which, I will argue, make his final theory inconsistent with its underlying justification. First, Gaus narrows the group of reasonable people itself in several ways, so it is ultimately less broad and diverse than initially claimed. And second, Gaus transforms the remaining reasonable people into viewpoints, which removes the ordinary tendencies and limitations that justified his theory in the first place.

The Justification and Elements of the Theory

For Gaus, a functioning society needs social rules to structure cooperation, and those social rules in turn ground the kinds of political norms that I have been discussing. Because social rules have moral authority over us, he argues, they must be ones that we could accept.⁵⁶ The reason for this serves as the underlying justification of his public reason theory.

⁵⁶ "Moral rules, we have seen, are social rules, and social rules exist;...our social life occurs in the context of existing social rules. Members of the Public are thus defined as the participants in the relevant rule

Gaus argues that the claim that social rules must be acceptable to citizens is inherent in our moral practices. When others violate legitimate social rules, we blame or shame them, and when we violate legitimate rules, we feel guilt. These practices only make sense when the rules we hold people to are ones that violators could accept. Gaus quotes Gabriele Taylor approvingly:

To feel guilty he must accept not only that he has done something which is forbidden, he must accept also that it is forbidden, and thereby accept the authority of whoever or whatever forbids it. The person who accepts the authority does not merely recognize its power and so thinks it simply prudent to obey its commands; he also accepts its verdicts as correct and binding.⁵⁷

More specifically, Gaus argues that we only engage in the moral practices of blame, shame and guilt for breaking a rule, when they have *sufficient reason* to accept that rule. Blaming someone for breaking a rule that he did not have sufficient reason to accept – that he could not have accepted – would be like blaming a two year old who cannot separate right from wrong. Annoyance might be appropriate, but blame would not be. So, Gaus concludes, a social rule that we would hold others responsible for obeying – a legitimate social rule – must be one that all could accept.

...the Basic Principle of Public Justification is not an ideal imposed upon morality by a philosophical commitment to the idea that moral agents are free and equal but is itself grounded in our reflective understanding of a bona fide social morality – one in which the moral emotions are well grounded and for which it is appropriate to feel guilt for violations. Unless each has sufficient reasons to internalize the relevant rule, and unless each has sufficient reasons to see that the rule requires ϕ in this context, and to typically act on the rule in the appropriate circumstances, the moral emotions that are necessary for social morality cannot be rationally sustained.⁵⁸

When do we have sufficient reason to accept a social rule? Let us think again about when we blame a person for breaking the rules. It is not merely when there are reasons for that person to accept the rules, because those reasons might be inaccessible to her. Given her own beliefs and values, she might not be able to see why that rule is important, no matter how hard she tried. Instead, we should only blame her when she *could have accepted* that

governed practice. Our analysis of social morality in Chapter IV showed us that we must justify our demands to those on whom we press them. The problem of justification arises when moral authority is claimed: it is our fellow participants in the rule-governed practice on whom we make demands in its name. Our question, confronting a social rule, is whether those whose lives are governed by the rule – those who make demands on each other grounded on the rule – all have sufficient reasons to internalize it and endorse it as binding.” Gerald Gaus, *The Order of Public Reason* (Cambridge: Cambridge University Press, 2011), p.268.

⁵⁷ Ibid., p.204.

⁵⁸ Ibid., p.263.

rule. At the same time, the rule does not have to be one that she actually does accept. When we say that someone “should have known better,” we do not imply that she *does* know better. Instead, we mean that, if she had thought about it more, she would see why she should have obeyed that rule.

This middle ground is how Gaus conceives of “having a sufficient reason.” According to Gaus, a reasonable person, Alf, has “a sufficient reason *R* if and only if a “respectable amount” of good reasoning by Alf would conclude that *R* is an undefeated reason (to act or believe).”⁵⁹ So, people have sufficient reason when they have “accessible”⁶⁰ reasons for accepting the rule – when they would accept it after a respectable amount of good reasoning, from wherever they start.

What is that starting point? Gaus again considers our moral practices. We feel guilty when we break a rule, even if we did not desire to follow that rule. So, if we want to know which rules engage our moral emotions, we should not look at what we desire. Instead, we should start from the set of beliefs, values and motivations that we use to evaluate social rules, which Gaus calls our “evaluative standards.”⁶¹ When people evaluate a rule as good (or would after a respectable amount of reflection), then we blame them for failing to obey it. Taken together, then, we have sufficient reason to accept a rule when we have reasons that are accessible to our evaluative standards.

This underlying justification produces a theory that is remarkably consistent across all the elements laid out in the introductory chapter: the “reasonable” group, the level of idealization, the object of acceptance, and the intensity of acceptance. Let us go through each of these in turn. First, since the purpose of Gaus’ theory is to find rules that allow for our common moral practices of blame and guilt, we would expect that the group that must be able to accept social rules – the “reasonable group” – would include nearly everyone. We expect most people to follow social norms and blame them when they do not. And, indeed, that is what Gaus argues. The group of reasonable people consists of what Gaus calls “moral” people – people who are capable of reflecting on their desires and subordinating those desires to their evaluative standards. Gaus calls such people “normal socialized individuals”⁶² and they include most people but not, for instance, children or psychopaths, who are incapable of looking beyond their selfish interests. Since we do not blame such people for their violations (though we may still jail the latter), they do not have to be able to accept social rules and are, therefore, unreasonable.⁶³

⁵⁹ Ibid., p.250.

⁶⁰ Ibid., p.255.

⁶¹ Ibid., p.278.

⁶² Ibid., p.145.

⁶³ “We do not confront young children and psychopaths as moral persons to whom authoritative moral prescriptions can be sensibly addressed. To be sure, we may still seek to regulate their behavior through social rules by threatening punishment for failure to comply, but it makes no sense to address them as moral agents who can see themselves as committed to acting on strong rules (§ 9.2a). Because of this our normal moral emotions are undermined in our interactions with them. We can treat them as patients to be helped, potential agents to be trained, or threats to be neutralized, but not as fellow moral persons who are able to grasp the authority of moral prescriptions. To feel resentment toward a four-year-old’s violation of a

Second, given that social rules must be ones that reasonable people have sufficient reason to accept, but need not actually accept, we would expect the level of idealization to be moderate – not identical with each person’s reasoning but not too far away either. And, that is largely the case.⁶⁴ Because desires and interests are irrelevant to whether someone has sufficient reason to accept a rule, moral people are represented by only their relevant reasons, which are their evaluative standards. To decide whether moral people have sufficient reasons, furthermore, we must represent those views in a way that makes sense of their conclusions. Gaus calls this making the views “intelligible.” To make sense of people’s reasoning, we would have to resolve glaring contradictions within it and loosely rank important values, but not put it into a precise order.⁶⁵ Finally, because people have sufficient reason when they have accessible reasons, we represent reasonable people’s views as if they had undergone a respectable amount of reasoning and reflection from wherever they began.⁶⁶ So, our moral practices appear to justify idealizing moral people’s views to their intelligible evaluative standards after a respectable amount of reasoning.

Third, the object of acceptance of such a theory must be the norms themselves and not the reasons behind the norms.⁶⁷ We blame and feel guilt for violating rules when we have sufficient reasons to obey those rules, not when we merely have sufficient reason to accept the reasons behind the rules. It would not make sense to blame someone for violating a rule because it is based on reasons they accept, if they ultimately believe those reasons require a very different rule than the one you are holding them to. And, Gaus argues just that; the object of acceptance must be social *rules*, as opposed to reasons or principles. The former is too attenuated from what you are being asked to do, and the latter is too abstract, which leaves room for multiple interpretations.

Fourth, the intensity of acceptance need only be moderate. We hold people to rules that those people do not think are optimal, or even particularly good. As long as they do not reject the rule outright, it seems as if they could recognize the authority of the rule and we would blame them for violating it. Whether Gaus adheres to this logic is a matter of some debate. In some sense, even a socialist who rejects private property can go to the market and be willing to exchange money for goods because he recognizes that this rule is operative there. He would abide by the rule, Gaus argues, because he recognizes that it is better than having no rule.⁶⁸ Gaus interprets this level of intensity as considering the rule to be better than a state of “blameless liberty.”⁶⁹ Critics have argued that preferring a rule to a state of anarchy does not show that one finds the rule acceptable, just that the

moral rule is akin to being afraid of a harmless rabbit. Given our practices and concepts, these reactions just do not make sense: they are, as McDougall said, absurd.” Ibid., p.210.

⁶⁴ Ibid., p.276.

⁶⁵ “Imagine Alan sitting on a public beach, a pebble in each hand, splitting one pebble by striking it with another. Betty, a casual observer, asks him what he is doing. She can see, of course, that he is splitting pebbles; what she is asking him to do is to explain it, to redescribe it as an activity with an intelligible point, something he could have a reason for doing.” Ibid., p.344.

⁶⁶ Ibid., p.249.

⁶⁷ Ibid., p.271.

⁶⁸ Ibid., p.322.

⁶⁹ Ibid., p.319.

alternative is so much worse. Instead, they argue, we should ask whether the person prefers the rule to a feasible alternative.⁷⁰ The critics may be right that the intensity required is inconsistent with Gaus' underlying justification. However, I will leave this debate aside to focus on two aspects that seem more problematic.

Overall, however, Gaus' theory is fairly consistent with its justification on all four aspects that that I have identified. A public reason theory justified by the need to legitimate our moral practices would indeed have something like these elements: 1) All "moral" citizens would be considered reasonable; 2) We would idealize the evaluative standards of these moral citizens so that they are intelligible and have undergone a respectable amount of reflection; 3) The object of acceptance must be rules, not reasons or principles; and 4) The intensity of acceptance must be at least "tolerable." A theory like this – one that seeks social norms that could be acceptable to a very broad group of reasonable people – is likely to identify fairly thin social rules, which may just meet Gaus' stated goal of "structuring social life." If that is all that is required, Gaus' theory can be both internally consistent and deliver modest political norms.

Accepting Liberal Norms

However, when Gaus argues for the conclusions of his model, the social rules are far more robust than this description would allow for. To justify such conclusions, Gaus changes many of the elements described above in ways that are inconsistent with his original justification. This results in a very different kind of public reason theory – one that starts with a much narrower group of reasonable people who can accept more robust, and often liberal, norms.

Gaus changes his theory in two ways. First, he changes the boundaries of "reasonable people" from "moral people" with very diverse evaluative standards to a group that largely agrees on important moral values and methods of reasoning. And second, like many of the other public reason theorists, he transforms his version of reasonable people into sophisticated viewpoints, which are capable of accepting more liberal norms. I will trace each of these developments, in turn.

To see how Gaus narrows the group of reasonable people, let us recall the kinds of views that are included and excluded among this group. Because his underlying justification is to make sense of our moral practices of blame and shame, among others, the group is defined by all those that we currently hold morally responsible for their actions and who are capable of subordinating their desires to care for moral rules ("moral people"). This excludes children and psychopaths, but not many others. Gaus then idealizes this moral group in three ways: First, they are represented by their evaluative standards and not their desires, since the latter are not *relevant* to blame and guilt.⁷¹ Second, they are represented by only an *intelligible* version of their views because only when we make their views intelligible can we say whether they have sufficient reason to accept a rule. Third, we

⁷⁰ For instance: David Enoch, "The Disorder of Public Reason," *Ethics* 124 (1): p.153.

⁷¹ "I am a burglar and I have a reason to keep on burglarizing" could, perhaps, be a sufficient reason for some to act, but that too is irrelevant to the justification of a moral rule against theft." *Ibid.*, p.281.

represent their views as if they had undergone a *respectable* amount of reasoning and reflection from wherever they began because, if they can access reasons to accept a rule, then we would blame them for violating it.

Despite all the restrictions and idealizations mentioned, however, there are still no substantive limits on citizens' views and, unless we slip in a substantive interpretation of "respectable amount of reasoning," there will be idealized citizens with undesirable, but intelligible, evaluative views on the right moral rules. Presumably, this group will include racists, anarchists, collectivists, and elitists – all kinds of inegalitarian, illiberal and undemocratic views. Other than very basic rules of security and basic welfare, perhaps, there seem to be very few rules that these viewpoints could agree on, even if they only had to agree that having the rules would be better than a state of blameless liberty.⁷² Anarchists, for instance, would choose the latter.

However, Gaus ultimately argues that such reasonable people, suitably idealized into what he calls "Members of the Public,"⁷³ can accept basic liberal norms such as free speech, freedom of conscience⁷⁴, the right not to be deceived⁷⁵, basic welfare rights⁷⁶, and strong property rights.⁷⁷ How can such a seemingly broad group of merely "moral people" accept such specific conclusions? Part of the answer, it seems, is that the group of reasonable people is much narrower than we thought.

Under what might be called "radical pluralism" we would so characterize the deliberations of Members of the Public as to allow for just about any evaluative standard that rational agents have endorsed – including, say, those that value the suffering of others and subjugating them. But this cannot be correct, for some limits are implicit in the very idea of public justification of a morality. Our aim in seeking public justification is to treat others as free and equal moral persons while also claiming moral authority over them. We are moral persons seeking justified mutual authority.⁷⁸

Under the banner of "limits [that] are implicit in the very idea of public justification of a morality," Gaus narrows the group of reasonable people beyond the idealizations mentioned above – beyond, therefore, what follows from his underlying justification. He does so in three ways.

First, Gaus defines "intelligible" and "relevant" reasons in a way that ensures a much greater degree of agreement among reasonable people than is required by the justification. Gaus is certainly right that some level of intelligibility and relevance are

⁷² It is a further question whether this level of acceptance is justified by the goal of assigning people moral responsibility for violations.

⁷³ Ibid., p.26.

⁷⁴ Ibid., p.354.

⁷⁵ Ibid., p.352-353.

⁷⁶ Ibid., p.359.

⁷⁷ Ibid., p.381.

⁷⁸ Citation needed.

required. If we are trying to figure out which norms reasonable people could accept, and we want to take the public reason route of representing their reasons and then seeing which norms are consistent with those reasons, then we must be able to understand people's reasons and relate them to some conclusions. As Gaus says, "...to attribute a reason to a person is to suppose her deliberations are intelligible and relevant to the enterprise of justifying a moral rule. To attribute reasons to others is partly an interpretive enterprise, and to interpret is to make sense of what the other is doing. I cannot (successfully) interpret your action in a way that renders it unintelligible."⁷⁹

However, the amount of intelligibility and relevance required to interpret reasoning is fairly minimal. Let us say that I wanted to know whether a member of the Westboro Baptist Church could accept norms against discrimination based on sexual orientation. If we took them at their word that they believe that the deaths of soldiers is punishment for allowing homosexuality, then we could represent their evidence as "American military casualties" and their conclusion as "homosexuality should be prohibited." We might find that reasoning absurd and abhorrent, but it is intelligible in one sense; we can understand the reasons and see what conclusions they would draw. For public reason's purposes, this meets the minimum requirement. Moreover, from the Westboro Baptist's perspective, the claim that God disapproves of homosexuality is clearly "relevant to the enterprise of justifying a moral rule" about discrimination based on sexual orientation. (It is not relevant to "us" but that can hardly count in accommodating pluralism.) So, the minimum amount of intelligibility and relevance is set by considering what "we" can understand and what "they" consider relevant.

However, there is a stronger interpretation of 'intelligibility' and 'relevance' that considers only how "we" see others' beliefs. This is how Gaus appears to interpret these terms.

If each can be confident that all other moral persons, from their own distinctive evaluative perspectives, also have reason to accept common authoritative requirements, the moral claims of all are consistent with the Principle of Moral Autonomy...and so we respect the moral autonomy of others. Suppose, though, Alf seeks to justify rule **X** to Betty by appealing to her standard ΣB , which leads her to endorse **X**. But suppose also that Alf believes that ΣB is not an intelligible basis for endorsing **X**: he simply cannot understand how there is a sound deliberative route from ΣB to **X**. He cannot see how ΣB is a standard relevant to endorsing **X**. In this case, if Alf demands that Betty comply with **X** because she endorses it on the basis of ΣB , he has no assurance from his own first-person perspective that she actually has a reason to comply...A plausible conception of evaluative pluralism must accept what we might call "mutually intelligible evaluative pluralism" at the level of Members of the Public.⁸⁰

⁷⁹ Ibid., p.280.

⁸⁰ Ibid., p.279.

Here, Gaus interprets intelligibility as understanding how “there is a sound deliberative route” from premise to conclusion. But, of course, while we can understand how a Westboro Baptist could get from military casualties to opposition to homosexuality, it is certainly not a sound deliberative route, which means that such views are effectively excluded from Gaus’ public reason model in the name of ‘intelligibility.’ Similarly, while the minimum amount of relevance required would be that the Baptist himself thinks his reasoning is relevant, Gaus argues that relevance should be determined by the group. Since casualties are clearly irrelevant to whether discrimination is permissible, this too would seem to exclude the Baptist’s views.

However, this objection would probably not worry Gaus because of a second move he makes to narrow the group of reasonable people. At least initially, he simply excludes those with objectionable views, including presumably the Westboro Baptists, by adding a seemingly substantive component to reasonableness. Reasonable people must now be “good-willed”:

Our real – or, at a minimum our first – problem is the nature of a publicly justified social morality among boundedly rational moral persons committed to the moral enterprise, and who confront each other as deeply disagreeing on the ends of life and the ultimate justification of each other’s foundational commitments, and yet who conceive of each other as reasonable and intelligible (though often deeply wrong-headed). Such people do not include pure egoists, though many Members of the Public put great weight indeed on their own interests and plans; neither do they include monomaniacs who are solely committed to one and only one value, such as one who cares about nothing but counting blades of grass, or his stamp collection. Such persons are the stuff of philosophical stories and objections, but in many ways they are beyond (or at the edge of) the common human horizon. In the end we should have something to say about such people, but we only confuse ourselves if we think moral inquiry commences with them. The general problem of social morality under pluralism among moral persons who confront each other as intelligible agents is severe enough; even those who think a crucial job is what we say to the Nazi, should at least allow that what we say to good-willed intelligible others is an important first step.

Of course, Gaus frames this exclusion as merely a first step. If we can solve this difficult, but manageable, problem of cooperation among good-willed people, then we can worry about extending our solution to Nazis and Westboro Baptists. However, any solution we find to cooperation among a narrower group will not tell us anything about the solution for a broader group. They are simply different questions. Solving the problem of cooperation among Kantians, for instance, does not make the problem of accommodating Nazis any easier; it merely postpones the true challenge, which is finding norms that all “moral people” could accept. Until Gaus can meet that challenge, it is fair to say that the conclusions that he draws are inconsistent with the task set by his underlying justification.

The third move Gaus makes to narrow the reasonable group is a consequence of the second. By excluding people with objectionable views and motivations from those who must be able to accept social rules, Gaus leaves us with a group who hold a much narrower range of views. So narrow, in fact, that he characterizes the remaining views as largely shared among all reasonable people, differing only by weight. Again, Gaus leaves some wiggle room here by suggesting that not all views are shared, but the implication is clear: in finding justified social norms, we need only think about the general consensus – what most decent people think.

As Isaiah Berlin might say, the range of plausible pluralism in the evaluative standards of Members of the Public is limited by the “common human horizon.” One way of thinking of this is that most of the disagreements of Members of the Public will be about how some Σ relates to the evaluation of x (e.g., whether it is outweighed by other evaluative standards) rather than whether it is a bona fide standard relevant to the evaluation of moral rule x . This restriction is bolstered by empirical research which indicates that the main source of our disagreements is not about *what* is valuable, but about what is *more* valuable. According to Milton Rokeach, Americans largely agree in affirming a set of thirty-six values; what they differ on is “the way they organize them to form value hierarchies or priorities.

At this point, we have come a long way from Gaus’ original theory, which promised to start with a broader group of reasonable people and still find norms they could all accept. When Gaus draws the conclusions from his model, it looks increasingly like the Rawlsian “shared reasons” model he argues against. Now, it seems, we ought to start with a group of good-willed people who differ mostly in the weight they give their views⁸¹ and who find each other’s reasoning to be sound and relevant.

Compare this to the group of reasonable people that Gaus’ underlying justification requires. Roughly, Gaus argues that we need to find norms that people can accept in order to hold those people responsible for violating the norms – in order to blame them. So, the group of people who must be able to accept the norms – the “reasonable” people – are those we tend to blame, who he calls “moral people” (as opposed to those, like children and psychopaths, who we tend to excuse or restrain). Through the three moves just described, however, Gaus has excluded many “moral people” – people who are capable of subordinating their desires and coming to care for moral rules. Nazis and Westboro Baptists are capable of adhering to social rules; they just adhere to bad ones.

⁸¹ Gaus does insist that, in his model, not all views are shared. “Now while they generally agree on what constitutes a good standard they do not always do so. We do not wish to go so far as to suppose that each endorses all the evaluative standards of others, or even that each thinks all the standards of others would be warranted at higher levels of deliberation. But each must accept that the standards of others are warranted at least at the minimum requisite level and that all can see the intelligibility of employing these standards when thinking about the justification of a moral rule.” However, as we will see in the next section, any views that are not shared end up dropping out at the level of abstraction in which he considers “Members of the Public.”

And even if we can justifiably exclude them, there are still many people who differ from others in more than just the weight of their beliefs. Such people are not simply outcasts, whom we must fight, but people we tend to hold responsible for violating social norms. If Gaus' underlying justification tells us that such people must be able to accept social norms, but his model goes on to exclude them, then his theory is internally inconsistent.

Gaus has a ready response to this argument, which he draws from Rawls. Gaus might argue that the moral practices behind the initial group of reasonable people – people we tend to blame and those we do not – are just one data point for our analysis, just as, for Rawls, our intuitions about the “original position” are only one data point. The other important data point is our intuitions about the kinds of norms we ought to have – in this case, basic liberal or libertarian norms. If there is any inconsistency between these two data points, that does not scuttle the theory. It simply means we must bring the two into reflective equilibrium by contracting the group of people we blame, or expanding the norms we consider sufficient, until the two are consistent. In this case, Gaus may argue, that means contracting the group of reasonable people from those we currently blame to a much smaller group. But, this is necessary in order to bring the group into line with the kinds of norms that Gaus believes we must have in order to effectively structure social life, which turns out to include certain liberal and libertarian norms.

In constructing the Deliberative Model we start out with a very broad range of moral persons who disagree, and construct Members of the Public as idealized representations. We then see what sorts of moral rules can be justified to them. We then may go back and consider those value systems that stretch intelligibility, which we have not yet modeled. We shall want to inquire whether these persons are committed to the moral enterprise and treating others as free and equal; if they are, we must see whether we can model any remotely intelligible Member of the Public to represent them. If we can see how some such persons might be described, we may iterate our deliberative model to determine whether the results radically change. Do we find that our justified social morality is fundamentally transformed in a way that is manifestly unacceptable (say, it shrinks to a very small core or even disappears)? If we find this, we may conclude that we have reached the limits of the reconciliation of reasonable pluralism with a morality and free and equal persons, and must confront the conclusion that, in the end, moral relations with such persons is not possible.⁸²

This is an effective response, to some extent. If you see our current moral practices as something we ought to continue in general, but not as infallible, you might think that certain kinds of people we currently blame are not, in fact, blameworthy. When we bring all of our moral practices into line, including the practices that structure our social life, we will see that some of our current practices are inconsistent with a publicly justified system and must be changed.

⁸² Ibid., p.282.

However, while consistency with our intuitions about norms that are not “manifestly unacceptable” is important, the internal consistency of the theory cannot require an inconsistency with its underlying justification. The underlying justification provides a limit to how much we can change the reasonable group, no matter how absurd the conclusions of the theory may be. As I argued in the Introduction, this limit is necessary in justifying a public reason theory. Without such a limit, a public reason theory has no reason to claim that this specific pairing of reasonable group and political norms is the right one. The underlying justification explains why the reasonable group must be just this broad and not broader or narrower.

In Gaus’ theory, that limit is set by our moral emotions. So, each time we exclude a person from the group of the reasonable, we are implicitly claiming that we cannot have moral relations with that person, which means we must treat them as we would psychopaths:

...to exclude moral persons is also a terrific cost: our moral relations with such people would be transformed into the relations that obtain between us and those who are not capable of moral autonomy...such as the psychopath. We may treat them as strategic partners, patients to be helped, or dangers to be contained – but not as fellow participants in a moralized social life. The moral emotions and the practice of blame would no longer be appropriately directed at them.

If this is the (necessary) limit, then we cannot shrink the reasonable group to: ‘good-willed people who differ mostly in the weight they give their views and who find each other’s reasoning to be sound and relevant.’ Doing so requires too drastic a change in our moral practices, such that we must treat many ordinary citizens as we would psychopaths. If this is required by reflective equilibrium, then Gaus must concede that we are permitted to ignore our moral emotions a significant amount in order to justify “better” social norms. Once he concedes this, however, more liberal theorists such as Rawls can argue that we should ignore even more of our moral emotions in order to justify more robust liberal norms. There is therefore no longer a reason to prefer Gaus’ pairing of a broad reasonable group and “thin” social norms to any other pairing a public reason theorist might offer. This is the consequence of an inconsistency with one’s underlying justification.

Gaus’ Diversity Argument: Moderate Expectations

Gaus ultimately argues that certain liberal and libertarian norms are justifiable to his idealized reasonable people, “Members of the Public.” As I have argued, it is difficult to see how the group of reasonable people that Gaus begins with – “moral people” – could all accept such particular norms. Gaus makes this claim far more plausible in two ways. First, as I have described, he shrinks the group of reasonable people to make it far more homogenous than it initially appeared. But, this only gets us part of the way to liberal and libertarian norms. To get from this smaller group of reasonable people to such norms, Gaus must further idealize this group, which he does by turning those realistic people into

“viewpoints.” And, just like the first move, this second step is also inconsistent with the underlying justification of his theory.

I have already discussed part of that justification. To blame people for violating social rules, Gaus argue, those rules must be ones that people could understand. And, since we apply moral praise and blame to all but the worst deviants, social rules must be within the grasp of all decent and ordinary people. But, there is a second aspect to the justification, which I call the “diversity argument.” As described in earlier chapters, the diversity argument is essential to any public reason theory; it tells us why norms must be acceptable to people of diverse faiths and philosophies, rather than to just the correct conception of the good. The diversity argument explains why pluralism matters to political morality.

Even if social norms must be acceptable to decent and ordinary people, as Gaus argues, we still lack a diversity argument. After all, even if norms must be within the grasp of ordinary people, perhaps ordinary people are capable of arriving at the singular, sophisticated, whole truth. Gaus’ denial of this possibility serves as his diversity argument. Gaus argues that, even if ordinary people reasoned as well as humanly possible, our ordinary human tendencies and limitations will prevent all of us from arriving at the same truth.

There are three limitations that cause this divergence: First, even extraordinary human reasoners are limited by the path-dependent nature of reasoning. Because humans are not capable of scanning all of our beliefs at once, we must start with some subset of our total beliefs. Which subset we start with will necessarily be a non-rational choice (if it is a choice at all), since this is the start of the reasoning process, not its conclusion. So, all reasoners will generally diverge because their starting points do.⁸³

The second and third limitations are particular to ordinary-but-reasonable people, who will inevitably diverge because: 1) They have limited time and cognitive ability⁸⁴ and, 2) They only have access to “easily obtained relevant information.”⁸⁵ Because of these limitations, we cannot expect too much of even reasonable people. Given limited time, ability, and information, the conclusions of abstruse economic, scientific or philosophical theories are not norms that reasonable people can accept. And because of our path-dependent reasoning, we cannot claim that reasonable people would all accept the same coherent philosophy, since each person would take a different path of reasoning that would alter her final conclusion. Instead, we must expect that ordinary-but-reasonable people will arrive at a diverse range of familiar moral views.

⁸³ Ibid., pp.242-243.

⁸⁴ “Given the realities of life, the demands on their time and resources, and differences in the sophistication of their reasoning, most simply would never understand the reasons this practice ascribes to them...” Ibid., p.255. Also see: “Even Members of the Public, who recognize their sufficient reasons, are of limited rationality, and as creatures of limited reasoning powers, when asked whether they have sufficient reason to endorse a rule, they must ask, “what are the alternatives we are deciding between, and what are the costs of refusing to endorse any of them?” Ibid., p.269.

⁸⁵ Ibid., p.237.

The maximum – what sets an upper limit on how difficult it can be to access the reasons which we suppose all participants to have – is that the practice of morality is not an elite practice such as physics or moral philosophy, but a basic human practice... We cannot ascribe to moral agents reason to accept infinite utility calculations, the noumenal self, or the original position. These may be elements of philosophical theories that explain or further justify people's moral reasons, and the philosophers who advocate them may argue that they are in some way the upshot of what normal moral agents do believe, but they are the result of specialist constructions based on long deliberations, and even their teaching is difficult (and alas, not convincing to many of our students). Our account of moral reasons cannot suppose that a person must grasp them in order to identify her reasons.⁸⁶

As with the Rawlsian version, however, Gaus' diversity argument does more than tell us that norms must be acceptable to diverse views. It also implies that the reasonable people holding the diverse views will have certain realistic qualities; namely, limited cognitive ability, time, and information. And these qualities have further consequences for what reasonable people are, and are not, capable of.

First, as Gaus argues, reasonable people's limited cognitive ability means that political norms cannot be overly sophisticated. We cannot expect them to reason from the original position, for instance. However, if reasonable people are cognitively limited, then it is not just their understanding of the norms that is constrained; their understanding of their *own beliefs* must be equally limited. They will have trouble identifying, ranking, and applying their own fundamental beliefs. Second, and similarly, if reasonable people have limited time and ability to evaluate norms, they presumably have limited time and ability to examine their own beliefs, clarify their own reasons, and prioritize them. And, third, if reasonable people reason path-dependently, they presumably cannot form a coherent whole from their own various beliefs and motivations. Gaus' diversity argument therefore constrains how much he can idealize people in seeing which norms they can accept. Just as they are limited in their ability to appreciate sophisticated norms, they are equally limited in their ability to understand, prioritize, and apply their own views.

Gaus' Viewpoint Model

That is not, however, how Gaus represents people in his idealized model; instead of representing limited and ordinary people, he turns them into viewpoints. To find out which norms they can accept, he first idealizes ordinary moral reasoners into what he calls "Members of the Public," who have a much more sophisticated understanding of their own beliefs and ability to apply them. Based on these improvements, he then claims that all Members – even, presumably, those that start off as non-liberals – can accept the same substantive liberal rights. As I will argue, the reasoning that he uses to show that they can all converge on these liberal rights may well be correct, but it is far too sophisticated to attribute to the limited reasoners that Gaus is committed to representing.

⁸⁶ Ibid., pp.254-255.

In an important step of his argument for the acceptability of basic liberal norms, Gaus argues that all reasonable people have sufficient reason to accept what he calls “the presumption in favor of liberty” (hereafter ‘the presumption’):

The Presumption in Favor of Liberty: (1) agents are under no standing moral obligation (in social morality) to justify their choices to others; (2) it is wrong to exercise one’s liberty so as to interfere with, block, or thwart the agency of another without justification.⁸⁷

Essentially, the idea is to place the burden of justification on those that would interfere with another’s liberty, rather than those who seek to exercise their liberty. When two claims collide, therefore, the default solution is to side with liberty over interference. And since, for Gaus, any justification for interference must be acceptable to all reasonable people, the interferer must provide a justification that is acceptable to the person he interferes with, which means the presumption is likely to support strong individual rights. As mentioned, Gaus argues that it strongly supports free speech, freedom of conscience⁸⁸, the right not to be deceived⁸⁹, basic welfare rights⁹⁰, and even capitalism⁹¹.

So, in order to defend the claim that all reasonable people have “sufficient reason” to accept these basic liberal rights, Gaus must first argue that all reasonable people have sufficient reason to accept the presumption and reject its alternatives, which are a presumption in favor of restraint or no presumption on either side. Again, a reasonable person, Alf, has “a sufficient reason *R* if and only if a “respectable amount” of good reasoning by Alf would conclude that *R* is an undefeated reason (to act or believe).”⁹² So, Gaus must argue that all reasonable people would conclude that there is a presumption after a respectable amount of good reasoning.

What is a respectable amount? This condition comes from Gaus’ diversity argument; since morality must be acceptable to all reasonable people, and reasonable people include ordinary reasoners, moral standards cannot require more than a respectable amount of reasoning by people of ordinary abilities, time, and access to information. So, for Gaus to claim that all reasonable people have sufficient reason to accept the presumption, he must argue that they would all arrive at the presumption from whatever views they start with, given only a respectable amount of reasoning that ordinary people are capable of.

Since, for Gaus, reasonable people are not necessarily liberal by definition, we can assume that at least some of them initially deny the presumption in favor of liberty. Who are these presumption-deniers? Act Utilitarians, it turns out, would reject the presumption because they believe that *all* actions must be evaluated for their utility, rather than only

⁸⁷ Ibid., p.341.

⁸⁸ Ibid., p.354.

⁸⁹ Ibid., p.352-353.

⁹⁰ Ibid., p.359.

⁹¹ Ibid., p.381.

⁹² Ibid., p.250.

evaluating actions that interfere with liberty.⁹³ More generally, this group includes all those who believe that there is no presumption in either direction; that is, those who believe that there are moral rules and we must follow them, whether we are interfering with, or taking, liberties.⁹⁴ At least some religious people would presumably belong in this group as well.⁹⁵

So, to make this more concrete, let us imagine someone who initially disagrees with the presumption: a Christian who believes in censorship to prevent blasphemy. Gaus' claim is that this "censor" has sufficient reason to accept the presumption anyway because her view that there is no presumption will be defeated once she engages in a respectable amount of good, but ordinary, reasoning. With that in mind, let us look at Gaus' actual argument for why all have sufficient reason to accept the presumption.

The argument begins by turning reasonable people into viewpoints, which Gaus calls the "method of abstraction." Gaus begins by abstracting a reasonable person's beliefs to a single, highly abstract perspective – the "perspective of agency." Like other viewpoint arguments, this represents reasonable people as if they were capable of identifying their core reasons at a very abstract level and granting them deliberative priority in all cases.

Suppose, then, that Members of the Public abstract to evaluate proposals simply from the perspective of agency. They deliberate only on the basis of the importance of being an agent – someone who deliberates, makes choices, and whose actions normally are based on those choices. Now this last point is important. An agent is one who understands her action as based on her choices: to choose is not simply an intellectual, but a practical, exercise. As Benn put it...an agent is one who sees "one's own thought as the prolegomenon to intended action that underlies a person's conviction that he makes decisions – that, unlike skids or lightning strikes, they do not just happen to him."⁹⁶

Gaus then cites Gewirth's argument that, if we start with this abstract element within the Christian's viewpoint – the perspective of agency – the presumption follows logically. Again, like other viewpoint arguments, this represents a reasonable person as if she can not only accurately identify and prioritize her core reasons but also argue faithfully from those reasons to their conclusions.

In Gewirth's view, insofar as we consider ourselves simply from the abstract perspective of agency, we must consider the aims and purposes of our action as good; we must value those aims and purposes that form the

⁹³ Ibid., p.342.

⁹⁴ "Many of those who dispute the Presumption of Liberty do so on the grounds of a devotion to a first-person perspective on morality (§12.4), according to which if Alf thinks ϕ is morally required he should go ahead and do it whether or not it can be justified to Betty." Ibid., p.343.

⁹⁵ In fact, since Gaus recognizes an absolute morality external to social morality, it may be the case that anyone who agrees with that does not believe in the presumption, as Gaus has described it.

⁹⁶ Ibid., p.341.

basis of our decision to act. Consequently, we cannot help but claim noninterference to act on these aims and purposes. How could one have no objection to thwarting what one sees as good? Freedom is a proximate necessary condition for acting on one's aims, and one's devotion to one's aims implies a valuing of the necessary conditions to act on one's choices...[which include] (ii) recognizing that this claim derives from the generic features of human agency as such and so (iii) all agents have the same ground for claiming noninterference as one does..."⁹⁷

My point is not whether this is a valid argument or not. The point is just that it is hard to imagine that our Christian censor, after a respectable amount of good reasoning and given the limitations of ordinary people, could replicate any argument like this and thereby conclude that there must be a presumption in favor of liberty. The whole point of public reason is that we cannot impose the conclusions of such sophisticated arguments on those who cannot accept, or even understand, them. Gaus recognizes exactly that when he argues that we cannot expect ordinary reasoners to accept the noumenal self or the original position. It is hard to see how an argument like Gewirth's – from the perspective of agency to claims of non-interference – is any less a "specialist construction based on long deliberation" than Rawls' original position. So, if ordinary people do not have sufficient reason to accept the original position because the latter is too complex, then the same would be true for Gewirth's viewpoint argument. Gewirth's perspective of agency is far too abstract and his subsequent argument too sophisticated to provide ordinary people with sufficient reason to accept the presumption. And this problem is not limited to an argument like Gewirth's; it is inherent in all viewpoint arguments. They abstract deeply and then extrapolate accurately – thereby turning people into viewpoints – in direct contrast to the limitations assigned by the diversity argument.

Gaus' Response

In part, Gaus recognizes this problem. Gaus acknowledges that the acceptance of a single abstract viewpoint does not tell us what diverse reasonable people could accept, so we must remove the abstraction, return to the diversity of viewpoints, and see whether the presumption is "overturned."

...when the abstraction is lifted, and the deliberators are aware of the full range of their evaluative standards, the conclusion reached via abstraction must not be overturned... we must show that, given the full range of reasonable pluralism of evaluative standards, all moral persons still have sufficient reason to endorse the claims of agency. If they do not, then appeal to these claims does not treat some as free and equal; some are subjected to the authority of moral claims that they cannot, as rational moral persons, endorse.⁹⁸

⁹⁷ Ibid., p.345-346.

⁹⁸ Ibid., pp.359-360.

This concession to realism, however, is limited to a concern with diversity; Gaus wants to make sure that the presumption is still acceptable to “the full range...of evaluative standards.” But, like other diversity arguments, acceptability to diverse views is only the *conclusion* of Gaus’ diversity argument. The premise of that argument is that reasonable people are limited in their reasoning ability – that is *why* norms must be acceptable to diverse views. So, to fully ensure that the presumption is acceptable to reasonable people, Gaus must not only remove the abstraction, he must remove the extra idealization as well. Gaus’ question should not be what “rational moral persons” would accept; it is what reasonable and limited people would accept.

So, while Gaus then argues that the presumption would not be overturned by diverse viewpoints, this is only part of the question. We must also ask whether it would be overturned by diverse, and limited, reasoners – ordinary moral reasoners who engage in a respectable amount of reasoning, under normal constraints. In that light, let us consider Gaus’ “non-abstract” argument for the presumption and see whether it would be acceptable to limited reasoners. Here is that argument:

When others seek to claim authority over them – to instruct them in what they must do – a free person insists that she is an equal interpreter of the demands of morality, and *her* understanding of morality is the one that should guide *her* action...But note that this is itself a quintessential claim of agency – it is a demand for free *moral* agency. The free and equal person insists that her reasons, as determined by her deliberation, determine what she ought to do; her reason is for her the definitive judge of what she, as a moral agent, must do. She understands herself as a self-directed moral agent, and only a moral practice that gives her sufficient reasons will suffice. This is not simply an abstracted feature of her set of reasons that might be overturned under full justification, for to be a free moral person is to, all things considered, demand reasons for one’s compliance with morality...In a society of rational evaluative pluralism, however, in determining whether he can freely endorse a moral rule, a rational moral person consults his evaluative standards, seeking to determine whether the moral rule conforms to them. Rational moral persons, then, are committed to more than simply free moral action; they are committed to their evaluative standards, and ensuring that moral demands are endorsed by them. They understand these standards as providing them with reasons to judge and to act, and so the claims of free agency include the freedom to act on these evaluative standards.⁹⁹

If I understand this argument correctly, the argument begins from the fact that a reasonable person seeks to act on her own beliefs about what she should do. Because she seeks to act on her own beliefs, she is implicitly claiming that her beliefs determine what she should do. Even if her beliefs told her to follow the directives of a priest, for instance, it is still her beliefs that ultimately direct her. So, to seek to act on one’s beliefs about what to do implies that one’s beliefs are the ultimate authority on what to do, and any

⁹⁹ Ibid., pp.360-361.

alternative proposal must be acceptable to one's beliefs. In other words, one's beliefs about what to do have a presumptive weight against any claim to interfere.

If this is indeed the argument, then there are two problems with claiming that ordinary moral reasoners could accept it. First, it seems possible for a reasonable person to arrive at the opposite conclusion, as one of Gaus' critics, David Enoch, does. Enoch argues that a reasonable person could see the authority of their beliefs coming, not from the authority of their person, but from God or truth or nature or the content of the beliefs themselves. If so, then it would not follow from having beliefs that one should give those beliefs extra weight against the potential interference of others:

Compare this to cases of, say, taste. Suppose that in the actual world you prefer vanilla ice cream over chocolate. Now imagine a possible world in which you prefer chocolate. And suppose that you – in the actual world – have to order ice cream for your counterpart in that possible world. What flavor do you order? Clearly, chocolate. What this shows is that in the actual world too, when you order vanilla ice cream, your reason for action *is* indexical, it *is* about you and your preferences, it's simply *that you prefer vanilla*. When you act on a principle or something like the value of autonomy, this simply isn't so. Your reason for action in such cases is *that-p*, not *that-you-believe-that-p*. It's just not about you at all. So there is no sense in which you're giving extra weight to your beliefs over others' – you're giving *no* weight to your beliefs here. You're just giving weight to the value of autonomy, or the truth of Catholicism, or some such.¹⁰⁰

But, the larger point is not whether reasonable people would side with Gewirth or with Enoch, but that not all would be capable of siding with either. Even in Gaus' second, "non-abstract" argument, he requires all reasonable people to abstract to the fact of having evaluative standards, recognize that this fact has priority over the content of those standards, and then draw the right conclusion from these premises, which is supposed to be the presumption in favor of liberty. This seems hardly more realistic than the "method of abstraction" that Gaus concedes is not fully justificatory.¹⁰¹ To find the norms that would be acceptable to reasonable people, it is not enough to find the norms that would be acceptable to diverse but highly rational viewpoints. You must find the norms that limited human beings can reason their way to.¹⁰² The latter, and not the former, is the project that public reason theories are committed to.

¹⁰⁰ David Enoch, "Against Public Reason," *Oxford Studies in Political Philosophy* (Volume 1) ed. by David Sobel, Peter Vallentyne, and Steven Wall (Oxford: Oxford University Press, 2015) p.131.

¹⁰¹ Gaus' argument that reasonable people cannot all accept autonomy could also apply to the presumption: "As such, they are not committed to personal autonomy as self-authorship or perfectionism. To be sure, the perfectionist can argue that they should be: he can insist that there is a reason for them to care, and they should see it. At the appropriate level of idealization, however, all Members of the Public do not have reasons to pursue this ideal. Members of the Public see themselves as committed to religious and other doctrines that can defeat the claims of such perfectionist autonomy." Gerald Gaus, *The Order of Public Reason* (Cambridge: Cambridge University Press, 2011), p.340.

¹⁰² Gaus could respond: 1) That this reasoning is accessible to ordinary people via experts, or 2) That there is another path from "no presumption" to "presumption" that ordinary people can easily grasp, or 3) That

Conclusion

This chapter begins by describing the underlying justification of Gaus' theory, which is based in our moral practices of blame and guilt, among others. I then argued that he could construct a theory in which the four elements of public reason theories – the scope of the reasonable group, the level of idealization, the object of acceptance, and the intensity of acceptance – are all consistent with this underlying justification. The result would be a compelling public reason theory that would show how fairly “thin” social norms could be publicly justified.

However, when Gaus describes the liberal and libertarian conclusions that he draws from his model, he changes two of the elements of his theory. This makes the conclusions more plausible, but also makes these two elements inconsistent with the underlying justification in our moral practices. The first element that changes is the scope of the reasonable group, which goes from all “moral people” – those who are capable of following moral rules – to a much narrower group in which members mostly share evaluative standards, differ mostly in the weight they assign those standards, and reason in ways that all others consider sound and relevant. The second element that changes is the idealization of those reasonable people from realistically described people with limitations of time, ability, and information to “viewpoints,” which precisely identify, rank, and apply people's views in ways that realistic and limited people would not be able to. As with the other public reason theories I have considered, the twin goals of justifying political norms to a diverse public and of validating liberal norms pull the theory into inconsistency.

this argument is intended to convince philosophers, not ordinary reasoners. Very briefly, I don't find these responses convincing for the following reasons: 1) if Gaus's theory is accessible via experts, then almost any philosophical theory should be – including the original position and noumenal selves, which Gaus denies – so there is effectively no limit on what ordinary people could accept. 2) If there is another, simpler path, Gaus has not offered one. It's hard to imagine a simple and uncontroversial argument that could convince, for instance, all act Utilitarians that they must be liberals. And, 3) If this argument is aimed at philosophers, then we still need one that could apply to ordinary reasoners. Otherwise, we have no reason to think that all ordinary reasonable people have sufficient reason to accept the presumption in favor of liberty.

Chapter Four – A Reasonable Hope: Maximal Idealization in Jonathan Quong

In this chapter, I consider whether we can justify public reason liberalism even if we begin by assuming that reasonable people already accept liberal values; that is, if we idealize them as much as possible. Dropping a realism requirement succeeds in avoiding the inconsistency between justification and idealization, but then gives us no reason to prefer a political to a perfectionist liberalism. Once we reinstate even a minimal realism requirement, however, the inconsistency returns. If reasonable people must be public reason liberals, then they are far too idealized to meet even the most minimal realism requirement.

Introduction

As I have argued so far, there is a fundamental conflict between the justification of a public reason theory, which requires a significant amount of realism, and the goal of justifying liberal political norms, which requires substantial idealization. This conflict plays out in a different way in each public reason theory, but one of the common ways it shows up is in the inconsistency between an aspect of public reason's justification – the “diversity argument” – and the claim that reasonable people could accept basic liberal norms. The diversity argument tells us why political norms should be acceptable to “reasonable people” who hold diverse views of the good, which requires attributing certain realistic qualities to those “reasonable people.” However, in order to show that liberal norms are acceptable to such people, public reason theories must turn them into diverse viewpoints, which lack those realistic qualities.

One way to solve this problem is to offer a diversity argument that is consistent with the maximum idealization of reasonable people. This is the appeal of Jonathan Quong's theory. Quong can fend off the objection that his reasonable people are insufficiently realistic by arguing that he is only seeking political norms for the ideal liberal society. Because it is ideal theory, we need not worry about whether his reasonable people sufficiently resemble real people, so we can conceive of them as viewpoints or ideal liberals or whatever matches our intuitions about the kinds of views that political norms must be compatible with.

Quong cannot move to fully ideal theory, however, because he would no longer have any argument against perfectionist liberals, who simply prefer their ideal to his. Instead, he must, and does, support a minimal realism requirement that allows him to reject both more realistic public reason theories, such as Gaus', and perfectionist liberal theories, such as that of Joseph Raz. According to this minimal realism requirement, reasonable people must be sufficiently realistic that they could be part of the “best foreseeable liberal”¹⁰³ society – an idealized, but humanly possible, future.

¹⁰³ Jonathan Quong, *Liberalism Without Perfection* (Oxford: Oxford University Press, 2011) p.143.

Even this minimal realism requirement causes problems for public reason theories, however, including Quong's. The first problem is the "asymmetry objection," which asks why realistic people (of any sort) would disagree about the good life but not about liberalism, as public reason theories suggest they do. I will argue that Quong could supply an answer to this objection, which is that shared political institutions in the best foreseeable liberal society could lead citizens to converge on liberal values, but not on a single conception of the good life.

However, this response would not help Quong's theory to answer a second objection: "self-application." Self-application requires that the demands of public reason theories apply to the theories themselves. So, for instance, the justification of public reason must itself be based on public reasons. I will argue that the self-application requirement is appropriate; reasonable people must indeed accept the public reason requirement in order to accept politically liberal norms. Accepting the public reason requirement, in turn, requires them to accept very specific beliefs. The conception of reasonable people that Quong's theory requires therefore ultimately fails the minimal realism test because we cannot reasonably expect that specific an agreement among minimally realistic people. This is not necessarily fatal to his theory, but it means that Quong no longer gives us a non-intuitive reason to prefer political to perfectionist liberalism. Both can argue that their political principles could be acceptable to some idealized group of liberals, but neither can claim that this group is realistic or that such a society could actually come about.

Ideal Theory

Public reason theories try to answer the question of what makes social or political norms – such as laws, constitutional principles, or social rules – legitimate; that is, why do states have the right to enforce these norms or, alternatively, why do citizens have to obey them? What separates contemporary public reason theories from other theories of legitimacy is the method of justifying these norms. The fundamental thought is that, to be legitimate, political norms do not have to be acceptable to all citizens, nor do they have to be fully just or true. Instead, the norms need only be acceptable to "reasonable people" and be sufficiently just.

The main work of a public reason theory is therefore to explain why we should prefer one particular pairing of a "reasonable group" with the norms it can accept to another such pairing. As I have argued, Rawls and Gaus both begin with a relatively broad, diverse, and realistic group of reasonable people and therefore struggle to show how such a group could accept liberal political norms. Quong has a potential solution to this: fully ideal theory.

By engaging in fully ideal theory, Quong can argue that reasonable people do not have to be realistic; they are merely a device, like the original position, for representing the kinds of characteristics that we think are relevant to deciding political norms. Quong quotes Rawls in support of this view:

The aim of the original position “is to uncover a public basis for a political conception of justice...In describing the parties we are not describing persons as we find them. Rather the parties are described according to how we want to model rational representatives of free and equal citizens.”¹⁰⁴

So, if we, as liberals, believe that the kinds of characteristics that are relevant include accepting basic liberal values and being fully rational, then reasonable people are liberal and rational.

...to say that certain principles of justice could be endorsed by all reasonable people is to say that those principles can be validly constructed from a normative ideal of society as a fair system of social cooperation between free and equal citizens. Reasonable citizens are a hypothetical constituency defined in terms of their acceptance of this ideal, and so addressing our justifications to this constituency is one way of specifying the reasons that can ground political justification in liberal societies.¹⁰⁵

This high level of idealization allows Quong to avoid two related problems attributed to public reason theories. The first is the problem that I have identified throughout this dissertation: the conflict between the “diversity argument” and the “viewpoint model.” The diversity argument tells us why political norms should be acceptable to people with a range of views of the good life. I have argued that explaining why reasonable people would disagree in this way requires attributing to them some realistic qualities. On the other hand, explaining how reasonable people could consistently accept liberal political norms requires turning them into fixed and sophisticated “viewpoints,” producing a conflict between the theory’s justification and its methods.

The second problem is similar in citing a conflict between realism and idealization and is often leveled at public reason theories; it is called the “asymmetry objection.” As I just described, public reason theories begin with the fact of disagreement about the good life and end by suggesting that reasonable people can agree on a liberal conception of justice (or the reasons that support it). However, if disagreement about the good is inevitable, even among reasonable people, then we should also expect disagreement about justice. After all, we debate matters of justice like the correct distribution of wealth just as much as any fundamental religious or ethical question. If our political norms can legitimately rely on a disputed idea of justice, then the theory is treating the two types of disagreement asymmetrically without good reason. Because Quong’s entire theory hangs on the differential treatment of liberal values of justice from ideas about the good life, this is a serious problem for his view and other Rawlsian theories of public reason.

Quong can potentially respond to both these problems by insisting that his theory is crafting political norms for the ideal society and not for realistic people. In his words, “The idea of the reasonable citizen is thus constructed from a conception of an ideal

¹⁰⁴ Ibid., p.293.

¹⁰⁵ Ibid., p.144.

liberal society, and not from empirical facts about actual citizens.”¹⁰⁶ In the ideal society, reasonable people would be as consistent and rational as viewpoints. And, in the ideal society, people would agree on certain liberal values of justice but disagree on questions of the good. We know that because they are defined that way.

Reasonable citizens must... be defined in terms of their shared commitment to a certain fundamental view of society as fair, and citizens as free and equal. Thus, any disagreements about justice between reasonable citizens necessarily presuppose certain fundamental political commitments. Although... those shared commitments will not yield a unique conception of justice on which all reasonable citizens can agree, it does provide an objective standard to assess competing claims and arguments... The asymmetry objection is defused because none of these commitments or beliefs can be subject to reasonable disagreement for the simple reason that... political liberalism has defined reasonable persons to include only those who accept those ideas.¹⁰⁷

Quong is therefore able to respond to the two problems that plague more realistic public reason theories by claiming to do ideal theory from the very beginning, which therefore allows him to define reasonable people as already accepting certain abstract liberal values: “freedom, equality, and fairness.”¹⁰⁸

While reasonable people are defined to accept these abstract liberal values, however, they are not defined to accept any particular liberal conception of justice, whether that is Rawls’ “justice as fairness” or a Kantian or Millian liberalism.

Even if all or most citizens in liberal democracies do endorse the basic ideas of freedom, equality, and fairness that are central to the political liberal project, they may draw very different conclusions about what a conception of justice based on those value requires. The values of freedom, equality, and fairness are notoriously open to differing interpretations. Libertarians, luck egalitarians, socialists, Republican flat-taxers, and Democrat tax-and-spenders can all be plausibly seen as developing conceptions of justice where freedom, equality and fairness play central roles. What are we to say about the vast majority of citizens in current liberal democracies who might reject... any particular liberal conception of justice.¹⁰⁹

Quong’s idealized liberals, therefore, are too ideal to reject liberal values in the abstract, but not so idealized that they share a liberal conception of justice. The result is a defense of *political* liberalism, which rejects both more realistic public reason theories that cannot

¹⁰⁶ Ibid., p.144.

¹⁰⁷ Ibid., p.214.

¹⁰⁸ Ibid., p.149.

¹⁰⁹ Ibid., p.148.

guarantee liberal norms and perfectionist liberal theories that insist on a more robust liberalism (that is not necessarily neutral between conceptions of the good).

However, while invoking ideal theory may be an effective response to more realistic public reason theories, it creates problems for the argument against perfectionist liberal theories. Once Quong argues that reasonable people do not have to be realistic, but rather reflect our pre-theoretic liberal commitments, then it is open to perfectionist liberals to assert their more robust liberal commitments. Perfectionist liberals could agree that legitimate political norms should be acceptable to people with the right liberal values, but just define the “right liberal values” differently. Reasonable people, they might say, should be defined to agree on the supreme importance of Kantian autonomy, which is why they are all able to accept robust liberal norms that seek to maximize individual autonomy. To defend political liberalism against this perfectionist challenge, Quong must give us some reason to think that reasonable people should be defined to accept the abstract liberal values of fairness, freedom, and equality, rather than more robust interpretations of those values.

Reflective Equilibrium

Quong suggests that we ought to attribute abstract liberal values to reasonable people because these values are the conclusion of a process that Rawls calls “reflective equilibrium.” He does not offer any details on how or why reflective equilibrium would favor these values over perfectionist ones, but I will suggest three arguments he could make. In the end, however, I will argue that none of these amounts to anything more than an intuition that it is more important to guarantee abstract liberal values than perfectionist liberal values.

According to Quong, the abstract liberal values are chosen for their important role by reflective equilibrium. This process begins with “ideas implicit in our public political culture...because we are most confident about their validity.”¹¹⁰ Presumably, these judgments are about rights and values that liberals generally believe in – free speech, limited government, equal voting rights, and so on. Then, the reflector comes up with principles that best explain these judgments. In this case, Quong argues that reflecting on our liberal judgments yields the conclusion that we can be most confident in the claim that liberalism is a normative ideal of “society as a fair system of social cooperation between free and equal citizens.”¹¹¹ In other words, we can be most confident that liberalism is some combination of the abstract values of fairness, freedom and equality, though we do not yet know more than that.

However, as when Rawls’ used the method of reflective equilibrium, it is not clear why we should start with the judgments that “the group” the most confidence in, rather than starting with those that reflect the liberal truth. Here are a few possibilities: First, it could be because reflective equilibrium tells us that abstract liberal values are most likely to be widely accepted in an ideal liberal society. Starting with judgments that are widely

¹¹⁰ Ibid., p.155.

¹¹¹ Ibid., p.155.

accepted by current liberals, and then improving them through philosophical analysis, gets us liberal values that could be accepted by more reflective liberals – ideal ones. This tells us directly what ideal liberals could accept.

Second, it could be because reflective equilibrium reveals the core, or essence, of liberalism. By starting with just our most confident judgments about liberalism, we arrive at what is truly non-negotiable about it. Defining reasonable people as holding only this liberal core therefore ensures the broadest possible acceptability without giving away the essential elements of liberalism, without which it would no longer be a liberal theory.

However, it is hard to see why someone who did not already share Quong's assumptions would find these further arguments convincing. These first two possibilities seem straightforwardly question-begging. The claim that abstract liberal values are most likely to be acceptable in an ideal liberal society depends on already knowing the character of an ideal liberal society. If the ideal liberal society were one in which all citizens were Kantian perfectionists, then autonomy would be likely to be widely accepted, rather than merely abstract liberal values. It is only Quong's definition of the ideal liberal society that makes the abstract liberal values more widely shared.

As for the second option – abstract liberal values are the essence or core or necessary minimum of liberalism – it seems initially plausible. But, what constitutes the core of liberalism is precisely what this debate is about. Is the essence of liberalism a balance of the value of broad acceptability with a more minimal liberalism, or is the essence about maximizing some perfectionist liberal value like autonomy, or something else entirely? As Quong admits in the previous quotation, socialist and libertarian views have claims to be called 'liberal,' so the concept is contested. But, once we drop the idea that the word 'liberal' provides any guidance, the claim reduces to the idea that acceptance of the abstract liberal values is the minimal level of liberalism that is necessary for legitimacy. That also begs the question against perfectionist liberals, who say otherwise.

A third possible justification is that reflective equilibrium tells us that abstract liberal values are more likely to be true than other liberal values, such as the value of autonomy. Reflective equilibrium starts with the judgments that we have the most confidence in and works them up into principles that we should have the most confidence in. If the result of this reflection is merely abstract liberal values, as Quong claims, then we ought to think that they are more likely to be true than more robust liberal values.

However, the claim of truth is unavailable to public reason theories. Political liberalism rejects perfectionist liberalism, not by showing that the latter is false, but by arguing that there is something more important than the whole truth – namely, some degree of acceptability. If Quong argues for that degree of acceptability by saying that the abstract liberal values are more likely to be true than perfectionist liberal values, then he is right back in the philosophical arena, competing over the truth of rival values.

None of these three justifications gives us a reason to define reasonable people as holding abstract, but not perfectionist, liberal values. More generally, there is something odd

about relying on reflective equilibrium to defend political liberalism. The idea behind political liberalism is that we have irreconcilable conflict among fundamental values about the good life. These disagreements are foundational, meaning there is no common standard to resolve them. The idea behind reflective equilibrium, on the other hand, is that theories are not justified by starting from foundational values and following them to conclusions; they are justified by how all our judgments hang together. So, we may be very certain about particular conclusions and unsure about foundations and end up jettisoning the foundations. Or, vice versa. The point is that reflective equilibrium is based on the assumption that foundational values are not merely givens that we must take or leave; we can evaluate them by their fit with our judgments at other levels of generality. If we have more confidence in our conclusions about justice than our ultimate foundations, we can use the latter as fixed points to see which foundational values fit best. It is therefore strange that the theory would claim that the clash of foundational values is irresolvable, despite agreement in other areas, when it relies on a justificatory method that uses agreement in other areas to resolve foundational disputes.

In any case, it is clear that Quong cannot argue for political over perfectionist liberalism from within fully ideal theory, as I have described it. If the characteristics of reasonable people are determined solely based on a conception of ideal liberals, then there is no more reason to define this conception to be politically liberal than perfectionist liberal. Instead, Quong must, and does, argue that we should prefer political over perfectionist liberalism because reasonable people must be minimally realistic. This minimal realism requirement allows Quong to argue that, while it is minimally realistic to expect agreement at a very abstract level of liberal values (at least in an ideal liberal society), it is too unrealistic to expect agreement on a robust value like autonomy or a controversial theory like Kantian perfectionism. Just as idealization allows Quong to reject more realistic public reason theories, the minimal realism requirement allows him to reject perfectionist liberal theories. Despite its necessity, however, the minimal realism requirement leads to two other problems for Quong's theory. The first, I will argue, can be solved, while the second may be fatal.

Minimal Realism

To find out how realistic reasonable people must be, we must know why Quong thinks they have to be realistic at all. As I have argued, Quong must attribute a minimal realism to reasonable people to fend off arguments from perfectionist liberals. So, Quong must explain why ideal liberals would still disagree about the good life and therefore cannot agree on a perfectionist liberalism. In other words, Quong must provide a "diversity argument."

Quong's diversity argument is that even the ideal liberal society will still contain disagreement because liberalism requires free conditions – free speech, conscience, and association. Under such conditions, even reasonable people will inevitably disagree on fundamental questions. The reason they will inevitably disagree will be important in a moment, but for now let us just grant that they will. Given this fact about any liberal society, Quong argues, even "ideal" liberal theory must take into account this minimal

amount of realism: ideal liberals will disagree about the good life. We cannot therefore define ideal liberals in any way we want to, even in ideal theory; we must define them according to what we could expect in the “best foreseeable liberal” society (hereafter BFLS).

...[political liberalism] does not see the fact of reasonable pluralism as something external to liberal theory. It is, instead, a fact about liberalism. The fact of reasonable pluralism is a fact about the exercise of rationality under liberal conditions. As Rawls says, this kind of pluralism is ‘the inevitable long-run result of the powers of human reason at work within the background of enduring free institutions’. Reasonable disagreement about the good life is part of liberal theory itself, since reasonable disagreement is what we must expect in a society well ordered according to any liberal conception of justice... This kind of pluralism is thus an essential feature of liberal theory in the sense that it will be produced under the best foreseeable liberal conditions.¹¹²

Defining reasonable people according to the best foreseeable liberal conditions is therefore a matter of internal consistency for liberalism. For liberalism to be consistent with “realizing its own principles under the best foreseeable conditions,”¹¹³ those principles cannot assume anything that is not “foreseeable.” So, if liberalism must be acceptable to all (or most) of its citizens, as public reason theorists argue, then it must be acceptable to all of its citizens under foreseeable conditions. Since it is not foreseeable that all citizens under free conditions will come to agree on Kantian autonomy, even in the best possible liberal society, liberal theory cannot assume such agreement. Therefore, Quong concludes, we cannot imagine reasonable people (or ideal liberals) to hold such robust liberal values. The best we can foresee is that reasonable people will agree on liberal values at an abstract level; they will agree on the importance of freedom, fairness, and equality, in some form and combination, but not on anything more specific.

Given that Quong must assume this minimal realism, however, idealization can no longer solve the problems with realistic public reason theories in the way that Quong suggests. In particular, we can no longer solve the asymmetry problem by simply defining reasonable people to agree on justice but not on the good. Instead, Quong must argue that, under the best foreseeable liberal conditions, reasonable people will agree on abstract liberal values but not on any values about the good life. And making that argument depends on knowing why people will inevitably disagree, even in the ideal liberal society.

Justifying Asymmetry without Definitions

Why is agreement on liberal values, but disagreement about the good, inevitable under the best foreseeable conditions? Most public reason accounts follow Rawls in starting with the same cause of disagreement – the burdens of judgment – but then drawing

¹¹² Ibid., pp.142-143.

¹¹³ Ibid., p.143.

different conclusions from that cause regarding justice and the good. The burdens of judgment are a series of challenges involved in moral reasoning, such as the difficulties of evaluating evidence and weighing values, as well as the fact that people's various life experiences and backgrounds will invariably color their judgment.

(a) empirical and scientific evidence is often complex and conflicting; (b) we may reasonably disagree about the relative weight of different considerations; (c) concepts are vague and subject to hard cases; (d) the way we assess evidence and weigh values can be shaped by our total life experience; (e) different normative considerations on different sides can make overall assessment difficult; and (f) the number of values any social institution can incorporate is limited.¹¹⁴

These burdens are supposed to explain why reasonable people inevitably disagree on questions of the good, even under the best foreseeable conditions.

At first glance, however, the burdens would seem to make agreement on abstract liberal values just as unlikely as agreement on the good. Let us say that we consider ideal liberals as those who would only accept some reasonable combination of freedom, fairness, and equality, rather than emphasizing one of those values at the expense of the others. Let us then say that some libertarians emphasize freedom at the expense of fairness, while some socialists do the opposite. If these libertarians and socialists do not accept all the abstract liberal values, then they might fail to do so because of the burdens of judgment, such as the burden that there are "different normative considerations on different sides that can make overall assessment difficult."

Or, if we consider agreement on the abstract liberal values to mean that one always prioritizes liberal over non-liberal values, then many ordinary citizens would not qualify as ideal liberals. For instance, there are citizens who would accept liberal conclusions most of the time, but insist that their view of the good has priority on certain issues, such as abortion. Kevin Vallier calls these viewpoints 'all things considered reasoners' (or ATCRs) because they sometimes consider their overall views of the good in political decisions.¹¹⁵ If we considered opposition to abortion to be an illiberal position, in that it violates freedom or equality, for instance, then this could be expected due to "complex and conflicting" empirical evidence or "vague" concepts. In these ways and others, the burdens of judgment could lead us to disagree on justice as well as the good.

To defend Quong's asymmetry, therefore, we need a quality that disagreements about the good possess and that disagreements about justice do not. Quong has given us part of such a story. He argues that the reason we must take into account disagreement on the good is because such disagreement is an inevitable product of liberal institutions, such as freedom of speech and association, even in the BFLS. Since it is an inevitable product of

¹¹⁴ Jonathan Quong, "Public Reason," *The Stanford Encyclopedia of Philosophy* ed. by Edward Zalta (Fall 2017 edition), URL = <<https://plato.stanford.edu/entries/public-reason/>>.

¹¹⁵ Kevin Vallier, "On Jonathan Quong's Sectarian Political Liberalism," *Criminal Law and Philosophy* 11(1).

liberal institutions, the liberal ideal must accommodate it. This argument only gets us part of the way, however, because it still cannot meet the second condition to justify asymmetry. Disagreement with abstract liberal values is also an inevitable product of free institutions, but Quong would not therefore seek to accommodate illiberal views. If we incorporate an idea from Joshua Cohen, however, this institutional story can give us a reason to think that, while disagreement about the good is inevitable, disagreement about justice may not be.

One of the burdens of judgment is that, even if we share the same values, each of us will be coming from different backgrounds and life experiences, and therefore interpret and weight those values accordingly. This is especially true of our views of the good, since we learn them from very different sources. We are taught different beliefs, attend different religious institutions with different moral leaders, and live in communities with different moral expectations.

As Cohen argues, however, this is not as true of political beliefs, at least in centralized liberal democracies.¹¹⁶ We grow up under the same general laws, constitutional principles, and political institutions, which exert an influence on us to accept the ideas implicit in them. For the most part, we go to schools that teach the same civic values. To some degree, we share sacred days and objects, representatives, and political heroes, past and present. And, unlike churches and temples, political institutions can use state power to shape our reality in accordance with the values they promote, which makes us even more likely to absorb them. For instance, in addition to being grounded in the value of equality, the laws *show* us that we are equal by then distributing votes and redistributing wealth in certain ways. These features of any liberal society make liberal ideas on justice “familiar and attractive”¹¹⁷ in a way that ideas of the good that belong to one tradition will not be to members of another. So, while the burdens of judgment and citizens’ different backgrounds will produce disagreement on the good, their shared membership in political institutions will produce some agreement about justice, at least in the best foreseeable liberal society.

Moreover, these convergent influences counteract the one burden of judgment that produces the greatest divergence in views: the claim that different backgrounds, upbringings and life experiences lead to very different conclusions. The other burdens of judgment presuppose such a level of agreement and explain disagreement from those shared premises, which may lead to some divergence, but not nearly as much as membership in different institutions. Sharing political institutions and experiences will therefore do the most to get us to converge on a shared set of values, at least at an abstract level. So, we might say that the burdens explain why, on questions of justice, we agree on abstract liberal values but disagree on their particular interpretations. And that is exactly the type of agreement that Quong attributes to reasonable people.

¹¹⁶ Joshua Cohen, “Democracy and Liberty” *Deliberative Democracy* ed. by Jon Elster (Cambridge: Cambridge University Press, 1998) pp.189-191.

¹¹⁷ Joshua Cohen, “A More Democratic Liberalism,” *Michigan Law Review*, 92(6) p. 1532.

Moreover, even if this convergence on justice is not likely to be as widespread or strong as this account implies, even in the best foreseeable liberal society, Quong claims he only needs to make a weaker claim: that such a convergence on liberalism is *possible*, while a convergence on the good is not.

The legitimacy of political principles does not depend on whether current liberal citizens do accept them, or whether the principles are congruent with their current beliefs. Instead principles are defined as legitimate if it is *possible* to present them in a way such that they *could* be endorsed by rational and reasonable citizens. This is why Rawls says, with regard to the overlapping consensus, that his theory ‘makes no attempt to prove, or to show, that such a consensus would eventually form around a reasonable political conception of justice’ ...¹¹⁸

Of course, this is a convenient story in two ways. First, it ignores the fact that modern liberal democracies have a constant influx of immigrants who may not grow up under the same institutions or be subject to the other convergent influences. Second, if all that matters is that such a convergence *could* come about, then we can say the same about a view of the good. We can imagine a homogenous liberal democracy in which all citizens attended the same churches or temples and, as a result, shared an idea of the good life, at least at an abstract level. This brings up the more general question of what value there is in showing that a convergence on justice *could* come about. If the claim is not that such convergence is likely to come about, is there anything to this claim beyond the sense that anything is possible?

Quong’s theory may have the resources to respond to both of these concerns. On the first point, he might say, it is true that immigrants would not experience all of the convergent influences mentioned above, but that does not mean that they would not experience any, particularly the longer they lived in their new society. They would still feel the influence that positive law and state enforcement of those laws can have on public opinion. For instance, newly naturalized citizens to the United States are required to know the contents of and swear to uphold the Constitution. It is possible that widely accepted, codified, and manifested political norms can influence the views of immigrants fairly quickly.

Even if not, however, Quong’s theory can survive a small minority that cannot converge on abstract political values. The purpose of his theory is to ensure that some version of liberalism will be practical, in the sense that it could be broadly accepted for moral reasons in an idealized, but possible, society. The idealization involves assuming that a controlling majority of the citizens are liberal but not that all citizens are. One might even argue that, just as the constraints on idealization require us to moderate our conception of liberalism, they should also require us to moderate the acceptability requirement. An idealized, but possible, society should not require unanimous acceptability because that is not possible. Of course, whether immigration would lead to enough dissidents to violate even this modified acceptability requirement would require more theoretical and empirical work on the standard of practicality and the likelihood of immigrant uptake.

¹¹⁸ Jonathan Quong, *Liberalism Without Perfection* (Oxford: Oxford University Press, 2011) p.144.

I think Quong could have a stronger response to the second concern – that there *could* be convergence on views of the good, such as a particular religion. It is true that, if everyone attended the same religious institutions, it could generate a convergence on religious values. Quong and Rawls would not deny that this is possible at some point in the history of a liberal democracy. What they would deny, I think, is that it could last. The contention of the burdens of judgment is that, under free conditions, people will inevitably disagree about the good life and so a unified church will inevitably splinter. The only way to retain the convergence would be to coerce its members to stay together or mandate a set of beliefs, which would then violate liberal freedoms of association and conscience. So, while a religious convergence in a liberal society is possible, it cannot be stable. A sustained political convergence, on the other hand, would not require any violation of free conditions. It is not a violation – and may even be a requirement – of a functioning liberal democracy to mandate shared participation in political institutions and the teaching of civic values.

More generally, the claim that a limited political convergence could come about is not just the claim that such a thing is possible, in the way that anything is possible. It is the claim that such an agreement is possible, given permanent features of liberal societies. More specifically, Quong must claim that nothing, including the burdens of judgment, precludes such agreement, which he can do by adding this story about shared political institutions. Full agreement on a single view of the good is precluded even under the best foreseeable conditions, but agreement on abstract liberal values is not. This allows Quong to justify the asymmetry without artificially insulating liberal values from the “foreseeability” requirement. If he can do that, then he can justify asymmetry without begging the question against perfectionist liberals.

This may not be the only story Quong could tell to get complete divergence on the good and limited convergence on justice. But, if any story satisfies, it must go something like this. It must acknowledge that even the BFLS would contain broad disagreement about the good, but there is nevertheless a common influence to converge on abstract liberal values. This influence cannot be too strong however. If it could produce more foundational or concrete agreement on liberal values, then perfectionist liberals could claim that the constituents of the BFLS would agree on autonomy or other views of the good. So, any useful justification of asymmetry can only produce a limited and abstract agreement on justice.

Self-Application

Even if this new defense answers the asymmetry critique, however, it also makes Quong’s theory vulnerable to another major critique of public reason theories: self-defeat or, as I will call it, self-application.¹¹⁹ The answer to the asymmetry critique, I claimed, could only justify a thin agreement; any thicker an agreement and perfectionist liberals could claim that reasonable people could also agree on perfectionist values. As I will

¹¹⁹ I change the name because I am not arguing that self-application considerations actually defeat the theory.

argue in this section, however, the self-application critique requires Quong to assign a much more substantive agreement to reasonable people than we have discussed to this point – far more than shared political institutions could ever produce, even in the BFLS. If reasonable people must agree on such substantive matters, and no foreseeable liberal society could ever produce such a substantive agreement, then Quong’s reasonable people cannot meet the foreseeability requirement and his liberal ideal is no more “practical” than the perfectionist ideal. In that case, they are both doing fully ideal theory with no considerations other than which theory endorses more normatively correct conclusions.

The self-application requirement is the idea that the requirements of a public reason theory must apply to elements of the theory as well. For instance, if public reason requires that each reasonable person limit herself to only public reasons in political decision-making, then that requirement must also be grounded only in public reasons. To many critics, the self-application requirement makes most public reason theories obviously untenable, especially theories like Quong’s that insist on such a narrow range of acceptable public reasons. Given that even well-informed citizens do not think they should restrict themselves to public reasons – indeed, most are not even aware of such a theory – a public reason theory would have to define “reasonable people” with a much greater degree of sophistication and agreement than the normal citizens they are supposed to resemble. Quong’s response is to say that the public reason requirement cannot be publicly acceptable in *our* society, but it can be acceptable to the ideal liberals of the best foreseeable liberal society. To see if that is true, let us consider exactly what those ideal liberals must accept in order to adhere to the public reason requirement.

There are three ways in which reasonable people might have to accept public reason itself: First, they might have to accept the public reason requirement itself. Quong’s theory requires that reasonable people only rely on reasons that other reasonable people can accept in political argument.¹²⁰ Does that requirement also have to be acceptable to reasonable people? Second, reasonable people might have to accept the boundaries of reasonableness. Quong argues that only those who prioritize abstract liberal over non-liberal values are reasonable and therefore qualified to accept or reject political norms. Do such “political liberals” also have to be able to accept that all, and only, political liberals are so qualified? Third, reasonable people might have to accept the underlying justification of public reason. That is, even if reasonable people have to accept the public reason requirement, do they also have to accept that requirement based on the same justification? I will argue that reasonable people must accept public reason in all three ways. Meeting each self-application requirement forces Quong to ascribe more and more particular views to his ideal liberals, which ultimately makes their level of agreement “unforeseeable.”

¹²⁰ With the exception of Rawls’ “proviso,” which allows for the use of private reasons, as long as public reasons could be given at a later time.

Quong rightly acknowledges the first self-application requirement: reasonable people must accept the public reason requirement itself.¹²¹ To simplify all these abstractions, I'll discuss them through an example. Imagine that Anu is one of Quong's ideal liberals, in that she accepts abstract (but not foundational) liberal values, which lead her to endorse the right to free speech. She is also a committed Christian who believes that saying "God damn it" is blasphemy and a grave sin. Quong requires Anu to set aside these views on blasphemy in political advocacy because they are not based on reasons that other constituents can accept. Since Anu believes that her views on blasphemy are correct, the only reason she would bracket this view is because she accepted the public reason requirement to offer others acceptable reasons. So, the public reason requirement is itself a reason that leads to her conclusion that there is a right to free speech; it is one of the things she must believe to accept liberal laws. Anu, and other constituents, must therefore accept the public reason requirement as one of their qualifications for being a reasonable person. To this degree, then, self-application requires Quong to attribute one further belief to reasonable people: they must also accept the public reason requirement itself.

It is less clear whether Quong would endorse the second form of self-application: that these public reason liberals must also be what David Estlund calls an "insular group,"¹²² which means they must accept that all and only other public reason liberals deserve reasons that they can accept. Whatever Quong's view, however, this too is required of reasonable people. If Anu thought that those who rejected abstract liberal values belonged in the group of reasonable people, then she would not have to limit herself to just liberal public reason. If she believed, for instance, that Christians who prioritized blasphemy also deserved reasons they can accept, then those Christians could not accept a blanket right of free speech. At this point, neither a blanket free speech nor anti-blasphemy law would be based on reasons that Christians and non-Christians could accept, so neither law would be legitimate, leading to an indeterminate theory. To take another quick example, let us say that Babu believed that only a very narrow group should get to determine political norms – the group of Kantian liberals. In this case, autonomy would be a reason that they could all accept, so they could base their case for free speech on autonomy, leading to a more robust liberal theory.

In other words, if Anu and Babu agree to limit themselves to just reasons drawn from (or compatible with) abstract liberal values, this must be because reasonable people see all and only those who limit themselves to abstract liberal values as reasonable. This insularity requirement adds another very particular view to Quong's ideal liberals.

Accepting the Justification of Public Reason

Already, we have seen that answering self-application requires Quong to attribute two substantive views to reasonable people: acceptance of the public reason requirement itself and acceptance of the boundaries of the reasonable group. It already seems implausible

¹²¹ Jonathan Quong, "Liberalism Without Perfection: Replies to Gaus, Colburn, Chan, and Bocchiola," *Philosophy and Public Issues - Filosofia E Questioni Pubbliche*. 2(1), p.53.

¹²² David Estlund, "The insularity of the reasonable: Why political liberalism must admit the truth," *Ethics* 108(2): pp.252-275.

that sharing political institutions could ever produce such substantive agreement under any “foreseeable” conditions. The third form of self-application, however, pushes the required agreement even further and makes it seem truly impossible. This is not an objection to political liberalism, per se; Quong can still claim that insisting on only politically liberal (i.e. “neutral”) reasons provides the best balance of acceptability to a broad group and quality of political norms. But, he no longer has any other advantage over perfectionist liberals, since they could claim that their balance of a narrower group and more robust norms is superior. Since neither theory is “foreseeable,” there is no advantage gained in being minimally realistic.

The third form of self-application asks whether reasonable people must accept the same justification for political liberalism. If they did, it would certainly increase the level of agreement necessary and threaten the pluralism that public reason wants to preserve. For that reason, Quong denies that his theory must apply to itself in this way. Quong claims that his reasonable people do not have to accept any particular underlying justification for public reason. In fact, he argues that, while each reasonable person holds some justification for public reason, this foundation is rooted in each person’s particular conception of the good and is, therefore, not part of his theory. Quong calls this the “buck-passing” account because he passes the buck to each reasonable person to incorporate public reason into her own view of the good.¹²³

As Elvio Baccharini argues, however, this buck passing account does not seem consistent with two other aspects of Quong’s public reason theory: the accessibility and sincerity requirements.¹²⁴ The reason that citizens must rely only on public reasons in political discussion, Quong argues, is that political proposals made to others must be justified to them. What does it mean for Anu to justify a proposal to Babu? Two things: 1) All the steps in the argument for the proposal must be “accessible” to Babu, from first premise to conclusion. That means that the steps cannot be based on evidence that Babu has no reason to trust or would reject, like other people’s religious experiences or sacred texts. And, 2) All the steps in the argument must also be “sincere,” which means that Anu must consider them to be valid reasons, not merely persuasive ones.¹²⁵ So, Anu could not offer Babu reasons that were accessible to his point of view, but that she rejected, because she would then be offering those reasons insincerely. These two requirements, when combined with Quong’s claims that fundamental views of the good life will never be accessible to all (because of the burdens of judgment), lead to the conclusion that only

¹²³ Jonathan Quong, *Liberalism Without Perfection* (Oxford: Oxford University Press, 2011), p.232.

¹²⁴ Elvio Baccharini, “Public Reason: The Consensus and the Convergence View,” *Filozofija i drustvo*. 25(1): pp. 75-95.

¹²⁵ “What does it mean to justify something to a particular person or group of people: Why not say that a given principle is, or is not, justified *simpliciter*? The answer, very briefly, is that whether or not some principle is justified to a person, Peter, depends on what else Peter is justified in believing. Suppose the question is whether proposition Q can be justified to Peter. Suppose the only valid justification for Q depends on premises A, B, and C. In order for Peter to be justified in believing Q, Peter must therefore also be justified in believing A, B, and C. But what if there is no way to justify A to Peter? The grounds for premise A might, for instance, be an eyewitness account of some event that Peter has no good reason to accept. Thus, A might well be true, but the grounds of A’s truth are not accessible to Peter, and there is no way to make them accessible to Peter.” Jonathan Quong, *Liberalism Without Perfection* (Oxford: Oxford University Press, 2011), p. 141.

abstract liberal reasons can be both sincere and accessible in the best foreseeable liberal society.

The buck-passing account, however, seems to suggest that abstract liberal reasons will be either inaccessible or insincere. Buck-passing, recall, means that each reasonable person grounds abstract liberal values in her own view of the good. So, if Anu supports abstract liberal reasons based on her Christian values and Babu does so based on his Kantianism, then justifying these views to each other will force them to either rely on reasons the other rejects or insincerely offer reasons from the other's tradition.¹²⁶ Therefore, if abstract liberal reasons are supported by each person's view of the good, and those views conflict, then abstract liberal reasons cannot be justified to other reasonable people. That is, if justification requires accessible and sincere arguments, then Anu cannot justify her liberalism to Babu because it is grounded in her Christianity. So, Quong cannot avoid the question of how reasonable people justify abstract liberalism merely by passing the buck.

Quong's response to this question of self-application is similar to his response to the asymmetry critique – we simply define reasonable people to see abstract liberal reasons as justified to all:

...(the failure to be consistent with PJS [Principle of justificatory sincerity]) does not afflict the Rawlsian consensus model of justification because...this model takes the fundamental ideas of the overlapping consensus as given, that is, as the justified starting points from which all reasonable public justifications can proceed. All suitably public justifications thus necessarily begin from fundamental and shared political ideas that we sincerely assume all other reasonable citizens are justified in accepting.¹²⁷

These definitional solutions are questionable in themselves, but they are particularly problematic in this case, because it seems that Quong is defining reasonable people to believe contradictory things. To see that this is the case, let me trace reasonable people's reasoning from their varied beliefs about the good to the abstract liberal values.

Reasonable people hold various views of the good. Nevertheless, their shared belief in abstract liberal values like fairness leads them to endorse the importance of acceptability, which says that all reasonable people deserve laws that are justified to them. Justifying laws to them means offering them sincere and accessible reasons for those laws and, since fundamental views of the good cannot be both sincere and accessible, they cannot be justifying reasons. This is the reasoning that reasonable people must use to get from

¹²⁶ "Reasonable Catholic citizens, for example, may view reasonable Muslim citizens as unjustified in endorsing Islam, and thus believe reasonable Muslims are unjustified when they take themselves to have sufficient reasons grounded in their Islamic doctrine to endorse the fundamental political ideas of freedom, equality, and fairness (and vice versa)." Jonathan Quong, "Liberalism Without Perfection: Replies to Lister, Kulenovic, Zoffoli, Zelic, and Baccarini," *Filozofija i drustvo*. 25(1): p.120.

¹²⁷ *Ibid.*, p.120.

their diverse views of the good to the belief that they must bracket these views of the good in determining political norms.

So, what happens when these reasonable people encounter Baccarini's worry that even their abstract liberal views cannot be sincere and accessible due to their foundations in conflicting views of the good? Quong argues that they simply assume that abstract liberal views are justificatory. However, this is in conflict with reasonable people's belief that justification requires sincere and accessible reasons. They must, therefore, have some way of resolving this conflict to avoid holding inconsistent beliefs.

Quong's arguments suggest two ways to square the sincerity and accessibility requirements with the foundations of abstract liberal values. The first is to modify the concept of justification so that the foundations of abstract liberal values do not have to be sincere. This strategy, I will argue, is not available to Quong because he rejects the rival convergence theories of public justification for doing just that. The second way is to make abstract liberal values their own foundation (self-justifying), so reasonable people that agree on abstract liberal values can access their foundations and sincerely recommend them to others. This second way meets all of Quong's requirements, but only by attributing another very substantive view to reasonable people, in violation of Quong's "foreseeability" condition.

The first way to make justification consistent with abstract liberal values is to modify the conception of justification, which, to this point, has required sincere and accessible reasons all the way down. In his reply to Baccarini, Quong suggests what might be seen as a two-stage concept of justification. The second stage is the one we are now familiar with, in which reasonable people argue from abstract liberal values to particular laws without relying on any insincere or inaccessible reasons (i.e. views of the good). The first stage is about the justification of the abstract liberal values themselves and allows for that justification without requiring sincerity.

...on the internal conception of political liberalism, the pre-OC¹²⁸ stage—the reasoning that leads individuals to endorse the fundamental political values from within their own comprehensive perspective—is not part of the process of public justification at all, and so the standards of public justification (like [Principle of Justificatory Sincerity]) do not apply to this stage. The comprehensive deliberations of citizens may be a necessary precondition for a well-ordered liberal society to emerge and for public justification to begin, but those deliberations are not subject to the same standards of sincerity and civility that apply when we engage one another in public discussion in our roles as citizens.¹²⁹

Essentially, Quong is suggesting that we can drop the sincerity requirement for justification prior to the agreement on abstract liberal values and only apply it after that point of agreement. The problem is, because the concept of justification is a crucial

¹²⁸ OC is short for overlapping consensus, which, for Quong, is the agreement on abstract liberal values.

¹²⁹ Ibid., p.120.

argument in reasonable people's support for prioritizing abstract liberal values, reasonable people must also accept this two-stage concept of justification. Anu offers abstract liberal values to Babu, rather than her deep Christian values, because she believes that justifying political norms to Babu requires only giving him reasons that he can access and that she sincerely believes. But, when she recognizes that her abstract liberal values are grounded in her Christianity, which Babu cannot access, she must adopt a two-stage conception of justification to continue to prioritize abstract liberal values over her Christian values. Changing the concept of justification, however, is precisely the solution that Quong rejects in his critique of convergence theories of public justification, so it is not available to him now.

Quong's argument against convergence theories is complicated¹³⁰, but the important thing for my purposes is his insistence that any form of justification that merits the name must meet the sincerity condition. Quong claims that any concept of justification without sincerity amounts to a "relativism about reasons"¹³¹ because it requires each reasonable person to see others' reasons for a norm as justified, even if she does not believe that they are actually reasons in favor of that norm. In other words, she must act as if other people's conception of the good were good reasons, even if she rejects them. Relativism about reasons, Quong continues, is far too controversial a belief to require of reasonable people, given the fact of pluralism. Put another way, it is far too substantive a belief to imagine citizens converging on, even in the BFLS. Quong's critique of convergence theories as requiring an insincere and controversial conception of justification therefore makes it impossible for him to rely on such a conception in his own theory.

If getting rid of the sincerity requirement is unavailable to Quong, then reasonable people must sincerely believe the reasons that they offer others, all the way down to the

¹³⁰ Consensus theorists like Quong argue that constituents should agree on abstract liberal reasons and then use only those reasons to justify political norms. "Convergence" theorists like Gerald Gaus and Kevin Vallier also insist on a point of agreement, but it occurs further down the chain of argument. In convergence theories, reasonable people "converge" on the political norms directly from their different views of the good, rather than first agreeing on abstract liberal reasons and then on political norms. So, for instance, under a convergence theory, Anu might find Christian reasons to accept freedom of speech directly – maybe it allows her to spread the word of God – rather than first accepting the general priority of liberal over Christian values, as Quong's theory requires. Babu, meanwhile, finds Kantian reasons to directly support free speech.

Quong's critique of convergence theories is that Vallier's "reasonable people" cannot justify norms like freedom of speech to each other because justifying *to* others requires offering sincere reasons for proposals. Because Anu does not accept Babu's Kantian view of the good, she cannot sincerely think that Babu is justified in accepting their convergent norm. From her point of view, Babu is joining the convergence for bad reasons, which makes her proposal to him mere persuasion, or even manipulation, rather than justification. So, Quong's critique of convergence theories is based on the importance of sincerity to justification. In fact, Quong argues that if reasonable people accept any concept of justification that does not require sincerity, then they are being required to accept an overly controversial concept of justification. For Anu to consider herself as justifying free speech to Babu, she must believe that being justified in accepting X means having sufficient reasons to accept X, *from his perspective*. Quong calls this concept of justification, the "moderate form relativism about reasons," and he argues that it is far too controversial a belief to require of constituents, given the fact of pluralism.

¹³¹ *Ibid.*, p.269.

foundations of their arguments. Given that views of the good differ, the only other option is to claim that abstract liberal values are themselves foundational values; reasonable people need not go any deeper. Reasonable people in the BFLS could then support their proposals with abstract liberal values and consider them both sincere and accessible to all other reasonable people. This second option may be what Quong is suggesting in the following paragraphs but, because it is unclear, I will quote him at length:

The objection I press against the convergence model succeeds in part because proponents of the convergence model—like Gaus—adopt a different picture of the public reason project. On Gaus’ account, there can be no normative commitments that are taken as given—as beyond the test of public justification. Rather, any normative claim that is alleged to ground a demand on others must be subject to the test of public justification, and so it is always appropriate to ask, about any individual’s normative commitments—even the most fundamental liberal values—whether we sincerely believe the individual in question is justified in endorsing those commitments.

My view of public reason is different. I begin with a puzzle that arises for those who share certain moral commitments and aim to live together with others on just terms that can be reasonably and willingly accepted by all, but who recognize that among those who share this aim there is no agreed religious, moral, or philosophical framework that can determine what justice requires. Individuals who share this aim, and recognize this kind of disagreement, should realize that the terms of justice will have to be grounded in public reasons: in reasons we all share by virtue of our common normative commitments to the values of freedom, equality, and fairness. Public reason is thus not, on my view, as foundational to moral reasoning as it is on Gaus’ account. As I see it, public reason is further downstream in the justificatory structure of moral and political philosophy. It is essential in helping us to understand what justice requires given a commitment to values such as freedom, equality, and fairness, but it cannot vindicate or undermine those prior normative convictions. As a result, we do not worry about whether individuals are justified in endorsing the fundamental liberal values—those values are not up for debate—they are rather the grounds from which we begin thinking about what political justification must look like in a well-ordered liberal society.¹³²

It is difficult to pin down the exact view Quong is suggesting, but there is one possibility that I think we can reject. Quong might say that we are starting with a conception of reasonable people as ideal liberals, so it is irrelevant whether they consider the foundations of liberal values to be accessible or not. Since they all agree on abstract liberal values, he could argue, we do not have to worry about whether they also see those values as accessible all the way down.

¹³² Ibid., p.120-121.

This response does not work because Quong's reasonable people include all forms of liberals, including perfectionist liberals. The reason that perfectionist liberals put aside their perfectionism is because they recognize that their perfectionist views are inaccessible to others. The accessibility requirement is the entire reason perfectionist liberal constituents rely only on abstract liberal reasons in political decision-making, rather than a thicker liberalism. So, reasonable people must think that even liberal beliefs need accessible foundations.

A more consistent interpretation would start from the thought that abstract liberal values serve as their own foundations or occupy a different sphere from our views of the good. Either way, views of the good would not "vindicate or undermine" abstract liberal values. In this picture, our political views exist off to the side of our fundamental views of the good life, so there can be no conflict. Abstract liberal values do not require any support from our other views and, as long as we know whether we are in the private or the political sphere, we will have no trouble knowing what to believe and how to act. Abstract liberal reasons would then be accessible all the way down because there is nothing below them. This is what it seems to mean to assume that freedom, equality and fairness are justified to all; we assume that "they are the grounds from which we begin thinking" in the political sphere.

This response meets the self-application requirement. Reasonable people can now accept the foundations of public reason because those foundations meet the public reason requirements – sincerity and accessibility – applied consistently. Answering the self-application critique comes at a cost, however. Requiring reasonable people to accept abstract liberal values as their own foundation means rejecting the buck-passing account in which each constituent supports those values from her own view of the good. In itself, this is no problem. But, it clouds the picture we previously had of people of many different ethical traditions accepting these abstract liberal values. Now, all those Christians and Kantians who saw abstract liberal values as an outgrowth of their foundational views are no longer considered reasonable. Only those who see political values as being their own foundation, or see political values as needing no foundation in the good, belong there.

We can now see all the beliefs that Quong must attribute to reasonable people as a result of the public reason requirement applying to itself. First, I argued that reasonable people must accept the public reason requirement itself; they must believe that other reasonable people deserve reasons they can accept. Second, they must agree on the precise boundaries of reasonableness: that all and only public reason liberals are qualified to decide political norms. Finally, they must agree on the foundations of political liberalism: that abstract liberal values are their own foundation or need no further grounding.

As Quong has emphasized, of course, he is doing ideal theory, so he is free to attribute very particular beliefs to reasonable people. The only requirement is minimal realism; that permanent features of liberal societies do not preclude this level of agreement, so that it represents a liberal ideal for a possible society. Some of those permanent features are

the burdens of judgment, which we know preclude any agreement on fundamental views of the good. We also know, from Quong's critique of convergence theories, that they preclude particular epistemic views, such as the relativism of justified belief. These are too controversial to ever be agreed upon in any liberal society.

It is hard for me to see, then, how these moral and epistemic views would be too controversial, but the views of Quong's idealized liberals would not. Even if we consulted only liberals, there would be disagreement on all three of the views Quong must attribute to reasonable people. Perfectionist liberals, of course, reject the public reason requirement itself. The claim that abstract liberal values serve as their own foundation is rejected by many of the reasonable views that Quong claims to accommodate – public reason liberals of all religions and ethical philosophies. Finally, even among public reason liberals, there would be disagreement over which “insular group” the laws should respond to – over who should be considered reasonable. The claim that ‘all and only public reason liberals who accept that abstract liberal values are self-justifying’ are qualified to determine political norms is not accepted by anyone, as far as I know – not even Quong, since he favors the buck-passing view. All of these views and their denials seem well within the burdens of judgment; reasonable people, and philosophers, can disagree about them.

Maybe, however, the other permanent feature of liberal societies, which I argued could forge some political agreement, can produce a society in which all or most citizens agree on such substantive views. The fact that citizens share political institutions, schools, and leaders could plausibly narrow the range of disagreement about political values, while broad disagreement about the good remained. So, Quong may use this as a way to answer the asymmetry critique and to attribute agreement on abstract liberal values to the ideal liberals of the BFLS.

However, as I also argued, the agreement forged by shared institutions could only be a limited one. For one thing, shared political institutions could not eliminate disagreement on political values because the rest of the burdens of judgment remain in force. We would still disagree about how to interpret and weigh those values, leading to different liberal interpretations of justice. More importantly, if shared institutions could produce a very substantive degree of agreement, then we can also imagine them to produce agreement on perfectionist liberalism, which would negate any advantage of Quong's theory.

If the agreement on justice in the BFLS is necessarily limited and abstract because anything more substantive is not foreseeable, and self-application requires Quong to attribute very substantive beliefs to reasonable people, then the conception of reasonable people that Quong needs is precluded by foreseeable conditions. Even if shared institutions can counteract some of the effects of the burdens of judgment on our political values, they could not lead us to converge on a single view of public reason's requirements, its foundations, or the boundaries of the reasonable. If Quong's conception of idealized liberals is not foreseeable, then it no longer holds any advantage over a group of perfectionist liberals; both are fully ideal, unconstrained by human limitations. If we still have a reason to prefer political over perfectionist liberalism, it is purely intuitive.

Conclusion

Quong calls trying to justify liberalism to non-liberals “an attempt to do the impossible.” It seems, however, that he has set himself an equally impossible task: to insist on political liberalism while rejecting perfectionist liberalism. Quong rejects perfectionist liberalism because it lacks a measure of realism; it must be possible, he argues, for a liberal society to agree on its foundational values. When that same realism requirement threatens political liberalism, first in the asymmetry and then in the self-application critiques, Quong is forced to further define and narrow his conception of reasonable people to the point where it can no longer meet any plausible realism requirement. Attributing very substantive beliefs to reasonable people, while declaring any further attribution to be unrealistic, seems like an impossible task. If neither is realistic, then political and perfectionist liberalism are back in the same boat: fully ideal theory.

This would not be a problem if there were non-intuitive reasons to prefer Quong’s conception of reasonableness to that of perfectionist liberals – if it were more likely to be accepted by idealized liberals or more likely to be true. But, any such reason ends up appealing to the intuition that a political liberalism is a better balance of acceptability and correctness, which is a strategy that is equally available to perfectionists. Quong’s only true argument against perfectionist liberalism remains its lack of “foreseeability,” which, if it is a problem, is a problem for both types of liberalism.

So, perhaps we can leave all that argumentation aside and offer a very simple case for Quong’s political liberalism: it represents the best balance of acceptability and liberalism available.¹³³ That is, it not only guarantees values that are very important to liberals (even if they are not the only important values), it also justifies them to a broad range of views of the good (even if others might prefer broader). Once we drop the sincerity argument against convergence theories, and the foreseeability argument against perfectionist theories, we are left with many possible balances of the values of acceptability and liberalism. Quong could simply claim that his theory offers a better intuitive balance than more realistic public reason theories because the latter cannot guarantee liberalism. And his theory offers a better balance than perfectionist liberal theories because the latter would only be acceptable according to an even narrower range of viewpoints.

At this point, we are left with a judgment call. Quong’s political liberalism offers a liberalism that is compatible with a range of views of the good so, in that sense, it does accommodate pluralism. And that is clearly the sense of pluralism that Quong intends when he phrases his puzzle as one:

¹³³ Quong seems to suggest something like this when he says: “Of course there’s much more that needs to be said in response to Gaus’s important challenge. All I can do here is conclude by emphasizing that I don’t think there’s a coherent and morally attractive alternative to my, admittedly, sectarian form of political liberalism. We can have a theory of public reason that won’t be sectarian, but then we can’t be sure it will be a liberal theory. Insofar as the public reason project is a distinctively liberal project, a certain amount of sectarianism is both unavoidable and, indeed, desirable.” Jonathan Quong, “Liberalism Without Perfection: Replies to Gaus, Colburn, Chan, and Bocchiola,” *Philosophy and Public Issues - Filosofia E Questioni Pubbliche*. 2(1), p.58.

...that arises for those who share certain moral commitments and aim to live together with others on just terms that can be reasonably and willingly accepted by all, but who recognize that among those who share this aim there is no agreed religious, moral, or philosophical framework that can determine what justice requires.¹³⁴

If the only pluralism that matters is the diversity of views of the good, then Quong's theory is indeed an admirable balance. In my view, however, Quong's political liberalism accommodates pluralism in the same way that Isaiah Berlin's value pluralism would: by mandating another very particular belief. Quong accommodates disagreement on the good by mandating agreement on public reason. This may be the best balance available; diversity of the good and conformity of the right is one way to accommodate pluralism. But, it does not seem obviously better than the reverse, which is what perfectionist liberals could offer.

¹³⁴ Jonathan Quong, "Liberalism Without Perfection: Replies to Lister, Kulenovic, Zoffoli, Zelic, and Baccarini," *Filozofija i drustvo*. 25(1): p.121.

Chapter Five – Liberal Democrats: Procedural Idealization in Joshua Cohen

Rather than justify liberalism by attributing substantive liberal beliefs to reasonable people, as Jonathan Quong does, perhaps we can justify it by attributing a value for a deliberative procedure. If successful, this strategy may avoid the inconsistency between justification and idealization by building the method of idealization into the justification itself. If we start with the value of deliberation, then it makes sense to consider the views of reasonable people after an ideal deliberation. While this may get us closer to explaining how reasonable people could agree on liberal norms, however, it does not get us all the way there. Joshua Cohen's argument that public reasoning would yield certain liberal norms illustrates the difficulty of justifying liberalism from mere reasonableness.

Introduction

There is a conflict between the justification of public reason theories, which conceives of the group of "reasonable people" as diverse and limited reasoners, and their conclusion, which suggests that this broad group can all accept liberal political norms. The theories that I have considered so far ultimately idealize reasonable people in ways that are inconsistent with the theories' justifications, such as by turning reasonable-but-realistic people into sophisticated "viewpoints." To avoid this problem, we must therefore find a method of idealizing reasonable people that would be consistent with its own justification.

One promising possibility is to ground a public reason theory in the value of a deliberative procedure. If the justification of a public reason theory argued that all those who valued democratic deliberation were reasonable, then this justification would not only give us a reason to accommodate this diverse group of people, but also a way of idealizing them. We could say that legitimate political norms are those that could be accepted by deliberative democrats after undergoing an ideal deliberative procedure. If we could then claim that such people could ultimately accept liberal political norms, we would have an internally consistent public reason theory that can still justify appealing norms.

To evaluate this possibility, I will draw on the work of Joshua Cohen, who argues that reasonable people that accept a democratic ideal would also therefore endorse certain liberal norms. I am primarily interested in his effort to draw liberal norms from the values inherent in a procedure, so I will focus mostly on the article, "Privacy, Pluralism, and Democracy,"¹³⁵ where he argues for particular liberal norms. To see whether his view could solve the problem that I have posed, however, I will go beyond his stated arguments in ways that – ironically for a public reason theory – he may not accept.

¹³⁵ Joshua Cohen, *Philosophy Politics Democracy*, (Cambridge: Harvard University Press, 2009), pp.269-303.

Hopefully, the view I elaborate will be independently interesting, even if it is not fully faithful.

I will begin with a conception of reasonable people as those who accept the democratic ideal, which means we are starting with a broad group of citizens. However, this group is quickly idealized in two ways: First, by attributing to them the right understanding of the democratic ideal: they accept that they ought to offer each other reasons that they can reasonably expect to persuade others so motivated, under ideal deliberative conditions. Second, by idealizing these “democrats” as if they had undergone an ideal deliberative process, which, plausibly, improves their views both epistemically and morally.

However, there is a limit to this transformation, which is what makes it a public reason theory. The ideal deliberation would not produce a full consensus on the good life. For either predictive or normative reasons, ideal deliberators will never agree on certain fundamental ideas of the good. Instead, they will hold a restricted, but still diverse, set of views of the good, even as they share a democratic ideal.

What kinds of reasons, and therefore norms, can such people accept? Given the starting point, they can certainly accept reasons based in the value of the democratic process – equal political rights, for instance, or free political speech. They can also accept any values that all views of the good could agree on: public health and safety, common defense, efficiency, the welfare of future generations, and so on. Because the representatives still hold conflicting views of the good, however, they cannot accept any reasons that depend on fundamental ideas of the good. In Cohen’s model, these fundamental commitments serve as vetoes on public reasons.

If the theory ended there, we would have an interesting but somewhat limited conclusion. Political norms ought to be based on public reasons, which are reasons that enable the democratic process or that all systematic views of the good could accept. Furthermore, public reasons are valid unless they conflict with the fundamental commitments of these views of the good, in which case they are invalid reasons.¹³⁶ In other words, we have legitimate guidelines to address all questions that do not interfere with any of these fundamental commitments and a reason to veto any state action that does.

However, the limitation of this theory is that it cannot tell us what to do in cases of conflicting fundamental commitments. If I have a religious commitment to prevent blasphemous speech and you have a liberal commitment to free speech, then it seems that the law could neither allow nor forbid blasphemous speech. A law that allowed such speech would be vetoed by my commitment, and a law that prohibited it would be vetoed by your commitment.

One way of solving this problem, consistently with the procedural justification, would be to settle such conflicts democratically. However, Cohen argues that reasonable people would always side with those whose freedom is restricted over those who would want to

¹³⁶ Unless they are “especially compelling [public] reasons,” Cohen says. This is similar to the compelling reasons test the Supreme Court uses to infringe individual rights. *Ibid.*, p.312.

restrict freedom, even if both are acting out of fundamental commitments. In this way, the democratic ideal supports liberal rights seemingly unconnected to the democratic process, such as the freedom of religion, expression, sexual intimacy, and even the right to choose an abortion. And, it does so without relying on any liberal norms.

Why would the representatives support the liberal side of these conflicts in all cases? I consider two possible answers: 1) That the state must justify coercion, but not the lack of coercion, and 2) Cohen's answer, which says that the representatives will all agree that the person with the strongest reasons is the one who is burdened by a law, so only *her* commitments serve as vetoes. I will argue that Cohen's theory cannot support either of these answers. Neither the values implicit in democracy (as Cohen sees them) nor the various idealizations result in a presumption of freedom or in the priority of self-regarding commitments over other-regarding ones.

I will then argue that Cohen could re-frame the burdened vs. non-burdened distinction as a distinction between sins of commission and omission that would not be distinctively liberal and therefore more congenial to all views of the good. While this may resolve some cases of fundamental conflict, however, it still leads to two problems: first, in certain hard cases, even this basis of agreement would break down, and second, there will be other cases in which this distinction does not produce a result that liberals would favor.

If reasonable people would not always side with individual freedom in these cases, then this theory cannot draw liberal conclusions from a procedural ideal. Nevertheless, a procedural theory avoids the inconsistency between justification and method that seemed fatal to the previous theories. In the next chapter, I will therefore suggest a way to retain this virtue and still justify appealing political norms, even if they are not as robust as public reason liberals might want.

Reasonable Democrats

We can begin a procedural theory of public reason by arguing that, just as we should decide ordinary policy through a democratic process, we should determine fundamental political norms in a similar way. Whatever outcomes citizens may desire, most can agree that we should decide them democratically. To those who are willing to abide by democratically determined laws and not impose their own wishes unlawfully, they get some positive say in deciding the laws. So, if we begin with the question of who deserves a say in deciding fundamental political norms – who is a “reasonable person” – we should give the same answer: democrats.

This is a promising place to start a public reason theory for two reasons. First, it is a broad conception of reasonableness, which ensures the values that I have called “responsiveness,” or more simply, broad “acceptability.” The value of democracy is not only widely held but, as Cohen employs it, it does not require accepting any prior norms that are not implicit in the democratic ideal. So, while Cohen will argue that the public reason process will result in norms consistent with liberal values such as fairness and

equality, he does not begin with those values. “Reasonable people” therefore begin with a wide range of other views, as long as they accept the priority of a democratic process in deciding fundamental political norms.

Second, as we will see, the democratic ideal, properly understood, does double duty as a way of improving people’s beliefs. Accepting democracy means accepting both a way of deciding and a way deliberating correctly, which gets us from this broad group of reasonable people to more particular conclusions, without the additional assumptions that Rawls and Quong require. In this way, a procedural idealization could be consistent with its grounding values.

From Democrats to Deliberative Democrats

A procedural view can define its reasonable group – those who deserve justification – as sharing the value of a democratic procedure and no values outside of that procedure. However, while the acceptance of democracy defines the reasonable group, it does not fully tell us how to conceive of reasonable people. To find out what reasonable people can accept, we must know how we should represent them.

Cohen begins by arguing that the best interpretation of the democratic ideal is deliberative democracy, which he defines as “reasoning together as equals.”

“The deliberative democrat emphasizes that democracy is not simply about treating people as equals in a process of collective decision-making, or about fair bargaining among groups, but also about reasoning together as equals on matters of common concern...”¹³⁷

Deliberative democracy is a unified ideal, but I will separate it into the three components to help explain its requirements: reasoning, reasoning *together*, and equality.

That democracy involves some form of reasoning is no longer a controversial claim, thanks largely to deliberative democratic theorists. Alternative theories have argued that democracy involves bargaining among competing interests, in which the procedure aims to get the answer that suits their interests or desires, in contrast to reasoning in which participants try to get the answer “right.” Or, that democracy involves a struggle among groups competing for a larger share of political power. But, proponents of both would acknowledge that the democratic process does not just translate given, pre-political interests into representatives and policy. People’s interests, values, and beliefs are transformed through political discussion, and there is no such thing as our views prior to this process. So, whatever its end goal, a democratic ideal must consider the norms of public deliberation that would serve its purposes.

However, these alternative theories *would* deny what deliberative and epistemic democrats claim about the aim of deliberation; that its goal is partly to get the decision right, in some sense that is independent of the process itself. For example, David Estlund

¹³⁷ Ibid., p.7.

argues that there is an epistemic component to common democratic procedures.¹³⁸ If it were merely a matter of fairness or of bargaining, then we could make decisions by lottery or take turns ruling each other. Instead, we have a process that includes free debate, secret ballots, a bureaucracy, and a Supreme Court that can only be explained by the objective of reasoning about the right answer and not by fairness alone. Perhaps based on something like Estlund's arguments, we can argue that democracy implies reasoning toward the right result.

Putting reasoning at the center of the democratic process means representing reasonable people in a second way: abstraction. We omit all other aspects of reasonable people, such as their interests and desires, to consider only their reasoning about what is good or right. This is why Cohen contrasts deliberation to bargaining or persuasion; it's not about providing others with reasons that *they* should choose a policy, or providing them with reasons that a policy would be in *my* interest. Instead, reasoning and deliberation are about providing reasons that *we* should choose a policy – about getting the decision right. To that end, we abstract to reasonable people's views about what is good for all.

Representing reasonable people by just their views of what would be good for all, and attributing to them the value of collective reasoning, means that reasonable people should hold a number of attitudes about democratic deliberation: 1) They aim to get the decision "right." As a form of reasoning, deliberation presupposes that all are not simply supporting the proposal that best serves their particular interests but are offering reasons that their proposal would be correct by the standard of appraisal they happen to hold, such as truth. 2) Since the object is to arrive at the right decision, reasonable people offer what they believe are genuine reasons and information, rather than engaging in strategic deception or manipulation, as they might in a negotiation.¹³⁹ And, 3) Since they are aiming at the right answer rather than just expressing their pre-deliberative preferences, reasonable people are open to changing their views when presented with what they recognize as better reasons.¹⁴⁰

The reasoning process is constrained, however, by the second and third aspects of deliberative democracy: democracy is a form of reasoning *together* among equals. Reasoning together tells us to seek a *collective* answer about the good life – an answer that we could reason to together. I am not reasoning with you if I insist you take my conclusion on faith; I have to try to persuade you. To try to persuade is to offer you reasons that I reasonably expect could persuade you. For example, I should not expect to persuade you of a decision by saying that it would make me famous; that would be an unreasonable expectation. So, deliberative democracy requires us to offer reasons that we reasonably expect could persuade others. As Cohen puts it, "Deliberative democracy is about reasoning together among equals, and that means not simply advancing

¹³⁸ David Estlund, *Democratic Authority: A Philosophical Framework* (Princeton: Princeton University Press, 2009)

¹³⁹ Joshua Cohen, *Philosophy Politics Democracy*, (Cambridge: Harvard University Press, 2009), p.158, n.9.

¹⁴⁰ *Ibid.*, p.28.

considerations that one judges to be reasons, but also finding considerations that others can reasonably be expected to acknowledge as reasons.”¹⁴¹

Deliberative theorists often draw an analogy to the process of decision-making in the U.S. Supreme Court. When the Court deliberates on legal or constitutional questions, the lawyers and justices offer each other reasons for their preferred verdict. But, they do not appeal to just any reasons. They draw from a set of reasons that they reasonably expect to be persuasive: reasons based in constitutional language, common values, precedent, prudence and accepted rules of inference. This serves a dual purpose. For one, basing the decision on these kinds of reasons means that even the losing party could agree that the decision is not based on power or interest but on reasons, which makes the decision more acceptable. But, just as importantly, since the ostensible goal of the parties is to reason together to the right answer, they each aim to offer reasons that are typically accepted by the other parties. Cohen’s deliberative democracy requires that citizens and public officials show each other the same respect that the justices do by basing collective decisions on reasons that they reasonably expect to persuade, or “public reasons.”

What are reasons that we could reasonably expect to persuade? How do we know what others will accept as reasons before we offer our reasons and see what they say? This is where the third aspect of deliberative democracy comes in: equality. For Cohen, equality is a procedural value, which means it is about situating each deliberator equally within the reasoning process. Truly collective reasoning could not come out of deliberation in which differences in power, status, or wealth had any influence in the process. So, acceptable reasons are those that equally situated deliberators could be persuaded by.¹⁴²

¹⁴¹ There are many different ways to draw this conclusion from Cohen’s work. He cites many other values in defending this restriction on acceptable reasons. He says that inherent in democracy is the idea of collective authorship of the laws. Laws are made in the name of all citizens, not just the majority that voted for them (or the candidate who voted for them). When the laws are based on reasons that all can accept, all citizens can see the laws as a collective product, as their own product. As Cohen puts it, “All who are subject to decisions must find their basis acceptable...Political decisions are not simply a product of power and interest; even citizens whose views do not win out can see that the decisions are supported by good reasons. As a result, members can – despite disagreement – all regard their conduct as guided, in general terms, by their own reason. Establishing such political deliberation would realize an ideal of self-government or political autonomy under conditions of reasonable pluralism.” Ibid., p.336.

Cohen also argues that deliberative equality demands more than an equal input into the process; the inputs themselves must be consistent with equality. As Cohen says, “if one accepts the democratic process, agreeing that adults are, more or less without exception, to have access to it, then one cannot accept as a reason within that process that some are worth less than others or that the interests of one group are to count for less than those of others.” Ibid., p.162. In other words, we require equal participation because we consider everyone’s interests to count for the same. Since we have abstracted to only reasonable people’s views of what is good for all, equality then requires not offering proposals that favor the conception of the good of one group over another. Even if a majority voted for it, such a proposal would still be undemocratic.

Whether we describe the value as political autonomy, equality, or collective authorship, however, the requirement on deliberators is the same: they agree to limit themselves to reasons that all can (reasonably be expected to) accept in political deliberation.

¹⁴² Ibid., p.171.

Of course, in real deliberation, it is not possible to eliminate all of these pernicious influences, so we cannot simply run a real procedure and see what others can be persuaded by. Instead, we must imagine what it would be like without these influences by putting reasonable people in an ideal deliberative situation. But even then, how can we know what could be persuasive in the course of ideal deliberation? Maybe you are not initially persuaded by my reasons but, after hearing my reasons again, you come around. In that case, what was initially unpersuasive later became persuasive. So, to know what could persuade other reasonable people, we must consider what they could accept *after* ideal deliberation and what they could not.

This is the third, and most complicated, type of representation – procedural idealization. We represent reasonable people as already having undergone an ideal deliberative process, in which they offer each other reasons, criticize the views of others, and are open to improving their own views.

Before I turn to how deliberation might improve reasonable people's views, I want to emphasize how far these reasonable democrats have already come. A procedural theory can start by saying that those who accept democracy deserve public justification. We can then draw on Cohen's work to argue that the best interpretation of the democratic ideal is deliberative democracy. If we represent reasonable democrats as accepting deliberative democracy, then, we ought to represent them as accepting the elements of that ideal: 1) Reasoning requires that we abstract to their views of the good; 2) Collective reasoning means that they accept the requirement to offer only ideas of the good that they could reasonably expect to persuade other representatives in an equal setting; 3) They accept that an equal setting is an ideal one; and 4) That we can only know what would be persuasive in ideal deliberation after it has occurred.

If we accept that the deliberative ideal is the best understanding of democracy, then we can accept that we ought to idealize reasonable people in this way. While this already attributes fairly specific views to reasonable people, it is also important to emphasize how much diversity remains upon entering the idealized deliberation. The representatives are not otherwise limited in what they think or value. As long as their views do not deny the value or implications of the deliberative process, they will believe all of the many varied things that actual people believe. They will differ in their philosophical and religious views of the good and, crucially, they do not enter the ideal deliberative process accepting any existing system of rights or principles of liberalism. If we accept the argument, they are, so far, just a product of holding a procedural value, which means there are no substantive values prior to the value of the procedure itself. The next step is to see how their views change after undergoing an idealized version of that procedure.

Procedural Idealization

“Reasonable people” are defined by their acceptance of the democratic procedure and they are then idealized by going through a hypothetical version of such a procedure – an ideal deliberation. The question is how significantly this idealization by ideal deliberation can plausibly change reasonable people's views. In the end, Cohen will argue that

reasonable people will favor certain liberal rights, so we must see how one can get from here to there.

In what ways does this ideal deliberation change the views of reasonable people? As you would expect, the deliberative process furthers both aspects of the goal of deliberation: reasoning that is collective. The reasoning process, which involves offering sincere reasons for a proposal that serves the good, being open to changing those reasons in light of new arguments, and sincerely presenting each other with reasons and information will, it is plausible to suggest, improve the content of reasonable people's views. At the end of deliberation, if reasonable people are truly open to changing their views in light of better reasons, their views will have been criticized and adjusted accordingly, resulting in more accurate and internally consistent beliefs. Crucially, this does not mean that the views that emerge from deliberation will either be the same or all true. But, Cohen argues, at the end of the process, there will be a range of beliefs and values that the representatives are "stably disposed to affirm...as they acquire new information and subject it to critical reflection."¹⁴³ In other words, further deliberation will not change these views. Surviving rational scrutiny under these ideal deliberative conditions makes these beliefs epistemically and normatively significant for Cohen. They are not necessarily true or fully rational, but they are what he deems "fully reasonable," and which I am calling 'procedurally idealized.'

In addition to improving the epistemic content of reasonable people's views, and therefore of their decisions, the deliberative process also has the potential to make the reasoning more collective. The idea is that, because the deliberative process requires reasonable people to only offer reasons that could be persuasive to other reasonable people, this act of offering impartial reasons can transform their own views and possibly make their views of the good more impartial.

Assuming a commitment to deliberative justification, the discovery that I can offer no persuasive reasons on behalf of a proposal of mine may transform the preferences that motivate the proposal. Aims that I recognize to be inconsistent with the requirements of deliberative agreement may tend to lose their force...¹⁴⁴

And, beyond improving my own views, it is likely that the norm of offering reasons that others can accept has a positive moral effect on others' reasons as well. For example, the recent erosion of norms against racist statements on some parts of the internet does not simply allow racists to speak. It also encourages provocateurs and young people to test the moral limits, which further degrades the quality of deliberation. Having a norm of public reason giving prevents some of this slide into immorality for its own sake.

The conditions on reasoning that deliberation imposes – sincere reason giving and information exchange, being open-minded, and possibly even the persuasive reasons requirement – work from different ends toward the same goal: a consensus on reasons

¹⁴³ Ibid., p.52.

¹⁴⁴ Ibid., p.26

(though they do not necessarily achieve it). Collective reasoning improves the epistemic quality of participants' views, so they are more likely to see the right reasons as the right reasons. And it is possible that the restriction to persuasive reasons moves reasonable people closer to seeing less partial views of the good, not as a concession to fairness, but as the strongest reasons there are. Together, the two push reasonable people closer toward a consensus on a fund of public reasons.

After Ideal Deliberation

None of this has yet answered our original question, however: which laws or principles can reasonable people now accept? We know that the representing mechanisms have moved them toward a consensus, but how far? The answer seems to depend on how optimistic we are. If we believe the effects of rational debate among the open-minded are very strong, then the representatives may end up with fully rational and perfectly impartial views after the transformation. If we are more pessimistic about the effects of ideal deliberation, then there may be little change in their views, which will contain as much diversity as they began with.

This is the crucial question for public reason theories, so I will take some time to explain Cohen's answer. It is crucial because the amount of diversity that would survive the idealizing process defines it as a public reason theory. If the full range of diversity remains after idealization, then it is hard to see how reasonable people could accept *any* particular set of laws or policies. Any particular content would be supported by some and vetoed by others. The effect of the idealizing process is supposed to be to narrow that diversity so that some political norms would be acceptable to all reasonable people.

On the other hand, if no diversity survives idealization, then we have to wonder whether the idealizing process is implausible. So, the success of a public reason theory depends on its ability to hold this middle ground – to claim that political norms should be decided by reasonable people that are neither fully diverse nor single-minded. As the defining feature of a public reason theory, then, we must offer a strong argument for the diversity of views that would remain after transformation.

I will begin by describing the kind of views that public reason theorists believe should survive transformation and ultimately determine which political norms are acceptable. Theorists call these “stubborn” views by many names – philosophies of life, conceptions of the good, and reasonable comprehensive doctrines – but they all refer to the same¹⁴⁵, somewhat vague, concept. A comprehensive doctrine provides us with a complete guide to our ethical lives, both individual and collective; it is a system of beliefs and values that tells us what we, and others, ought to pursue and avoid. And the “reasonable” comprehensive doctrines – those that we should not expect to change in ideal deliberation

¹⁴⁵ Joseph Chan argues that they all have slightly different connotations, but I do not think the terms are used precisely enough to take anything from which phrase a theorist uses. Joseph Chan, “Legitimacy, Unanimity, and Perfectionism,” *Philosophy & Public Affairs*, 29(1), 5-42.

– will be our familiar and long-standing ethical and religious traditions, such as various forms of Christianity and Buddhism, Kantianism and Utilitarianism.

Here are three arguments that these familiar doctrines would survive, while other views would not. Each turns on the meaning of the requirement that ideal deliberators must offer each other reasons that they can reasonably expect to persuade others. As R.J. Leland and Hans van Wietmarschen have pointed out, there is an ambiguity in the phrase “reasonable to expect.”¹⁴⁶ This phrase can use ‘expect’ in a predictive or a normative sense. If it is predictive, then ‘reasonable to expect’ means making an accurate-enough prediction about what others would accept. If it is normative, then reasonable to expect means not demanding too much of others – not violating any norms, or expecting others to violate any norms, in changing their minds.

We can argue for the predictive sense of the claim by looking at history. We could argue that the current and prominent comprehensive doctrines would be stably affirmed after new information and critical reflection because they always have been in the past. We have seen that certain ethical and religious doctrines have not tended to converge or disappear, despite millennia of debate among “deliberators” of all levels of competence. No historical deliberative procedure or institution (that is not objectionably authoritarian) has been able to persuade the different camps to converge on a single doctrine. On the other hand, while all comprehensive doctrines have not converged into a single comprehensive doctrine, some have converged and others have disappeared. Some religious and ethical beliefs have faded over time, while others have gained more adherents. Taken together, these facts suggest that certain familiar doctrines have staying power, but they are unlikely to converge or disappear through any free process of deliberation, formal or informal.

These historical developments provide some evidence of the kinds of views that a real deliberative procedure would encourage and discourage and therefore give us a reason to validate the views that have survived. To claim this as evidence, we must claim that at least some of these historical developments were due to deliberative procedures such as reasoned debate and not to morally pernicious factors such as power or ill will. Cohen argues that, even if most of these “debates” took place under non-deliberative conditions, the fact that we do not need to rely on anti-deliberative factors to explain the persistent disagreement suggests that true deliberation would have the same results. Even if we could eliminate the effects of power and ill will, people would still disagree on their conceptions of the good because of the different systems of thought and backgrounds of the deliberators, which would be part of even ideal deliberative conditions. If we can assume that at least some of the historical development of comprehensive doctrines is due to an exchange of reasons and information, then we can claim that these debates have “idealized” the doctrines, to some degree, and that they have survived for good reason.¹⁴⁷

¹⁴⁶ R.J. Leland and Hans Van Wietmarschen, “Reasonableness, Intellectual Modesty, and Reciprocity in Political Justification,” *Ethics*, 122(4), pp.725-726.

¹⁴⁷ Joshua Cohen, *Philosophy Politics Democracy*, (Cambridge: Harvard University Press, 2009), p.227.

Even if we can accept history as evidence of the results of a real deliberative procedure, however, we must go further. After all, we need to know what they would accept after going through an *ideal* deliberative process with other idealized deliberators. To that end, Cohen offers two further arguments to support the idea that long-standing comprehensive doctrines would survive the deliberative process – one predictive and one normative.

A second predictive argument considers the limits of an ideal deliberation among reasonable people. Reasonable people, you will recall, are motivated to reason together, which means providing others with reasons that those others could be persuaded by. The deliberative process can therefore improve comprehensive doctrines epistemically by criticizing their internal consistency and factual accuracy, as suggested earlier, because reasonable people can expect each other to be persuaded by invoking shared epistemic standards and the uncontroversial conclusions of science (in most cases).

However, it *is* unreasonable to expect others to be persuaded by, for instance, my belief that the Bible is the word of God. This is not a source of evidence that they share, nor can I show it to be a consequence of any reasons they currently hold. From their perspective, I am not offering a reason but merely a statement of my belief.

...if others accept the idea of reasonable pluralism, then they notice what we also notice, namely, that what lies between our taking our views to be reasonable (about which there may be no disagreement) and our taking them to be true (about which there is disagreement) is not a further reason, but simply our (rationally permissible belief) in those views.¹⁴⁸

So, claiming that one's doctrine is not just reasonable but the whole truth has only "limited force."¹⁴⁹

If it is unreasonable to expect others to be persuaded by one's fundamental beliefs about the good, then a reasonable person will not offer those as reasons. Instead, a reasonable person will stay on a shared terrain of reasons. But, from that shared terrain alone, there is no definitive argument to say that another's comprehensive doctrine is wrong. Once the appropriate inconsistencies and factual errors have been pointed out, there remains a core of beliefs and values in a range of comprehensive doctrines that can only be contradicted by asserting that one's own beliefs are true and others are therefore false. Without that possibility, rational scrutiny in deliberation will be exhausted when it runs out of shared standards, which means the remaining views will be stably affirmed under the maximum rational scrutiny that is possible in ideal deliberation.

Put another way, this predictive version of the argument goes as follows: there are versions of each long-standing doctrine that would meet basic and common epistemic standards such as consistency and factual accuracy and therefore survive the rational scrutiny of a deliberative process. Any further scrutiny, however, would not offer

¹⁴⁸ Ibid., p.55.

¹⁴⁹ Ibid., p.54.

conclusive evidence that all could accept to reject these doctrines, so they would survive deliberation in some form.

Whether or not this defense succeeds, Cohen offers a third reason that ideal deliberators would not expect comprehensive doctrines to change – this time a purely normative argument. Here again, Cohen argues that it is unreasonable to expect our comprehensive doctrines to persuade others. This time, however, it is not because that would be a poor prediction but because others should not have to change their comprehensive doctrines to accept our reasons. Cohen puts it this way: “All we should say in response to the fact of reflective divergence is that in matters of comprehensive morality, the truth, if there be such, transcends the exercise of practical reason appropriate to expect of others, as free and equal.”¹⁵⁰

Why would it be inappropriate “to expect of others, as free and equal?” Drawing from a few different places, this interpretation seems the most compelling. First, we must understand what “free” means in the previous statement. To say a citizen is free, Cohen says, is to acknowledge that “no comprehensive moral or religious view provides a defining condition of membership...”¹⁵¹ So, treating others as free means not expecting others to adopt or abandon their comprehensive doctrines as part of political life.

Of course, comprehensive doctrines provide a complete guide to ethical life, so changing any one of our ethical ideas cannot mean *abandoning* our comprehensive doctrine. Some parts of a doctrine are more essential to it than others. Cohen refers to these essential beliefs and values as “fundamental interests,”¹⁵² “obligations,”¹⁵³ or “nonnegotiable commitments.”¹⁵⁴ Commitments are fundamental because their content contains their own importance. ‘Fundamentality’ refers to the strength of the demand placed upon the believer by the content of the belief, rather than the intensity with which the believer holds the commitment. So, for instance, the content of a religious obligation includes the belief that it is among the most important priorities of a believer, which makes religious obligations particularly fundamental commitments. Because of this stringent content, fundamental commitments are some of the most important reasons we have; they are “overriding obligations” or “nonnegotiable commitments.”

Moreover, unlike intense preferences, all can agree that one’s fundamental commitments are essential to us or nonnegotiable.¹⁵⁵ Why? The answer seems to be that the content of intense preferences does not contain their own importance, so they can be seen by others

¹⁵⁰ Ibid., p.226.

¹⁵¹ Ibid., p.156.

¹⁵² Ibid., p.264.

¹⁵³ Ibid., p.246.

¹⁵⁴ Ibid., p.310.

¹⁵⁵ “There is not general agreement about the relative weight of different reasons. But there are degrees we can acknowledge. If it is a matter of obligation, it has support for an especially weighty reason. If I prefer it, even if I prefer it intensely, that is a less weighty reason. If I say it is essential to human life, it has considerable weight. If I say it is a good thing to do, it is weightier than a matter of preference (even intense), but it is not as substantial as an obligation or a necessity for a decent life.” Ibid., p.313.

as self-imposed. My preference for collecting art, for instance, does not tell me that I ought to prioritize it above most things. If I do so, it is because I choose to, and I would acknowledge that. Prioritizing a religious obligation, on the other hand, is not something I would say is my choice. From my point of view, it is not a self-imposed obligation. Others can recognize that, from the point of view of the commitment-holder, fundamental commitments are not seen as self-imposed and so cannot be abandoned without abandoning the comprehensive doctrine itself. And, of course, the requirement that we treat people as free means that we cannot ask them to abandon their doctrines. So, while there are certain ethical claims that we can reasonably expect to persuade people out of, fundamental commitments are not among them.

Let us accept that one of these defenses succeeds and see what it tells us about the reasons that are and are not acceptable in deliberation. We started with the norms inherent in the deliberative ideal that told us that we are seeking the good of all, which, in turn, meant reasoning only according to considerations that all idealized deliberators could accept. To know what idealized deliberators could accept, we must know what they accept after the idealization process, which is an idealized deliberation. The views that would be stably affirmed after deliberation turned out to be the familiar comprehensive ethical and religious doctrines, as long as they meet certain basic and shared epistemic standards. These views are part of the idealized constituency because they meet a sufficient level of epistemic quality – they are able to survive rational scrutiny – and their fundamentality – they contain what are agreed to be very weighty reasons for the person who holds them.

So, these are the views that reasonable people hold after ideal deliberation: long-standing but conflicting conceptions of the good but also the desire to reason according to shared ideas about the good. Together, these two factors produce the basic requirement of Cohen’s theory: “A consideration is an acceptable political reason for free and equal people just in case it has the support of the different comprehensive views that might be endorsed by reasonable citizens.”¹⁵⁶ Acceptable public reasons, then, are considerations that Cohen’s idealized deliberators could accept, based on their acceptance of the value of deliberation and their ‘procedurally rational’ comprehensive doctrines, such as: the value of public order, efficiency, political equality, free speech and assembly, the welfare of future generations, and so on.

The Case For and Against Liberty

Reasons are public reasons when the various idealized religious and philosophical traditions can accept them. What reasons do not count as public reasons? Because public reasons are reasons that all can agree to, all “private” reasons are excluded: interests, convictions and preferences that not all people share. Those may be interests for you, Cohen argues, but they are not interests for the rest of us and so cannot be the basis of legitimate public decisions.

¹⁵⁶ Ibid., p.234.

If this were the whole story, however, Cohen's theory would have unappealing consequences, at least for Rawlsian public reason theorists. Representatives would only defend proposals based on common reasons, and without any substantive limits on the outcomes, there would be no way to defend many non-procedural rights that liberal societies currently enjoy, such as freedom of religion. Consider the *Employment Division v. Smith* case. When the state of Oregon passed a law banning the use of peyote, it could have relied on a number of acceptable public reasons, including public safety or health. Smith's counter-claim that he used it for a (let us assume, obligatory) religious ritual, however, would not be supported by any public reasons. His religious obligation, while it would survive the idealizing process, would not be a shared reason for reasonable people. And, freedom of religion in general would not be a value that reasonable people shared because it is not inherent in democracy, deliberative democracy, or a necessary result of ideal deliberation.

Cohen has two options to defend Smith and freedom of religion. He could go back and attribute to reasonable people a value for these non-procedural rights, like religious freedom. As we have seen, this is what Jonathan Quong does. But, for Cohen, that would defeat his purpose of beginning with only a procedural value and justifying from it these non-procedural rights. He explicitly denies that his defense of religious freedom relies on any ideas drawn from liberalism, so he cannot begin with such rights.¹⁵⁷

Instead, Cohen argues that reasonable people would not choose any law or principle that prevented the fulfillment of one of their fundamental commitments: "Religious people cannot accept, as sufficient reasons in support of a law, considerations that would preclude their compliance with those demands."¹⁵⁸ What is a consideration that would preclude Smith from fulfilling his commitment to take peyote? Cohen's argument is that a law banning peyote will be based on reasons of public safety, which is, ordinarily, a genuine public reason. But, if the law interferes with a fundamental commitment, like a religious obligation, then one of the implicit reasons favoring that law *must be* that public safety is a weightier reason than Smith's religious obligation to take peyote as part of his ritual. Since Smith can't accept *that* as a reason, the value of public safety no longer carries any weight in support of this policy. The religious obligation here serves as a veto on the public reason *and* on the public decision.¹⁵⁹ In this way, the requirement of a

¹⁵⁷ "More generally, we can see, at least in general terms, how a case for personal liberties emerges without resting on ideas of personal autonomy, individuality, or self-ownership drawn from liberalism, understood as a general philosophy of life." Ibid., p.311.

¹⁵⁸ Ibid., p.164. "First, because of the pluralism of philosophies of life among politically reasonable citizens, some bases for regulating conduct are politically weightless. To take the clearest case, people hold some commitments on faith, and take those commitments to impose overriding obligations. Such nonnegotiable commitments are not as such unreasonable, nor is there anything unreasonable about embracing them as true. But because they are expressly held as truths known through faith, they are matters on which reasonable people disagree, and adherents cannot reasonably expect others to accept those considerations as having any weight and, therefore, cannot use them in justifying regulations." Ibid., p.309-310.

¹⁵⁹ "But it will not be acceptable to suppose that, as a general matter, the value of public order transcends all other political values. Except perhaps in the most extreme circumstances, for example, a state may not impose a blanket prohibition on alcohol consumption – including consumption in religious service – in the name of public order. After all, the considerations that support such consumption include considerations of

consensus on reasons supports the freedom of religion; if a law supported by public reasons prevents a person from fulfilling his fundamental commitments, then those public reasons have no weight in that case.

Cohen makes a similar argument for the fulfillment of other, non-religious fundamental commitments. For instance, on freedom of expression, he argues that some non-political expression fulfills fundamental commitments, so they must be protected from laws that would prohibit them: Bearing religious witness with no intention of persuading; artistic expression; professional advice given out of a sense of duty with no expectation of convincing; and speaking on matters of justice, regardless of how the message is received.¹⁶⁰ If the law interferes with these activities, it gives weight to reasons that deny the weight of these fundamental commitments. Since reasonable people with such commitments cannot accept those reasons, the law cannot be based on reasons that all can accept and is therefore illegitimate. So, it turns out that interfering with these non-political expressions is actually a violation of the deliberative ideal, a “failure of democracy.”¹⁶¹

With this move, Cohen, like Rawls with his “overlapping consensus,” has added an element of a ‘convergence theory’ to his consensus theory. Convergence theories argue that acceptable laws or norms are those that would be acceptable to each person’s comprehensive doctrine. Similarly, with the requirement that Cohen has added, not only do reasonable people have to deliberate based on acceptable public reasons, but the decisions they reach must be acceptable to their *private* reasons, when they involve a fundamental commitment. Comprehensive doctrines must agree to both the reasons supporting a law *and* the law itself, meeting both consensus and convergence requirements. For Cohen, a “reason all can accept” is a reason that: 1) Has the support of all reasonable comprehensive views, and 2) Does not support a law that precludes the fulfillment of anyone’s fundamental commitments.¹⁶²

The problem with adding a convergence element to a consensus requirement is that it leads to indeterminacy or “stalemates.” If the only things that must be acceptable to all comprehensive doctrines are reasons, then no matter how many reasons were vetoed by private commitments, there would always remain some reasons that all doctrines could acknowledge as good reasons – public safety, efficiency, prosperity, and so on – and these would support some range of laws. But, if the laws themselves must be acceptable to all doctrines, then there will be cases in which having a norm precludes the fulfillment of a commitment and *not* having a norm precludes the fulfillment of a contrary commitment – a stalemate.

For example, on many issues, there are fundamental commitments to allow some conduct and fundamental commitments to prohibit that same conduct. Millions demand free

obligation, which will provide a suitable basis for rejecting a justification cast in terms of the value of public order, except in the most extreme conditions.” Ibid., p.235-236.

¹⁶⁰ Ibid., p.324.

¹⁶¹ Ibid., p.246.

¹⁶² Unless there is a very compelling reason that the individual herself can accept

speech because of their commitment to individuality or utility, while some religious groups have a commitment to prohibiting blasphemous speech. Liberal doctrines based in autonomy insist that all individuals get to choose their own destiny, while the Amish want to keep their children out of public schools to maintain their traditions. There are fundamental commitments on the side of liberty and liberals, but there are also commitments that oppose them. In these cases, it seems like reasonable people can neither accept a law that prohibits, nor a law that permits, these activities.

The obvious example is the abortion debate. Philosophical liberals have a fundamental commitment to the value of autonomy, which includes control over our bodies and our choices. That means that any public reason that would support a ban on abortions would lose its weight as a reason; it would preclude the fulfillment of a fundamental commitment, which liberals could not accept. But, it also seems that we cannot pass a law that permits abortions. The Catholic belief that life begins at conception – also a fundamental commitment – prevents them from accepting any public reason as a sufficient reason in support of a pro-choice law.

How can we resolve these stalemates? One option is just to acknowledge that we have reached the limits of public reason. Public reason has given us a framework to make decisions in the many cases that do not involve clashing fundamental commitments. If there are no commitments involved or the commitments are only on one side of the issue, then public reason applies. If not, then we need to find another way to resolve them. Perhaps we can defend a principled compromise in these cases or pursue a federalist solution. Maybe reasonable people could all accept a real, democratic decision-making procedure or set of institutions that treat the conflicting camps equally in their decisions.

However, Cohen argues that the representatives would all agree to resolve these issues in favor of those who seek more freedom rather than those who seek to limit it – in other words, in favor of liberalism. How can we defend this? Not through public reasons, it seems. Cohen has already said that any potential public reasons in this case – the value of a potential life, the health of the mother, the value of reproductive freedom – have little or no weight when they preclude the fulfillment of a fundamental commitment, which both having a law or not having a law would seem to do.

Maybe the solution, then, is to have no law on abortion at all, which would effectively allow it. This is similar to an argument of Gerald Gaus'. If two legitimate viewpoints cannot both accept any rule, then the default is to have no rule because of a "presumption in favor of liberty."¹⁶³ It does not matter if having no rule suits one viewpoint and frustrates the other; state coercion is what requires justification, not the lack of coercion. Many theorists have questioned this 'asymmetry,' as Gaus calls it. Nicholas Wolterstorff has argued that there is no reason to think that coercion stands uniquely in need of justification.¹⁶⁴ For example, the government can do many objectionable things without coercion, such as naming an official religion, which we would want a theory of

¹⁶³ Gerald Gaus, *The Order of Public Reason* (Cambridge: Cambridge University Press, 2011), p.341.

¹⁶⁴ Nicholas Wolterstorff, "The Paradoxical Role of Coercion in the Theory of Political Liberalism," *Journal of Law, Philosophy, and Culture* 1(1), pp.135-158.

legitimacy to have something to say about. Colin Bird has even argued that non-coercion stands equally in need of justification; those who call for coercion can have just as great a stake in the policy as those who they want coerced.¹⁶⁵

Cohen seems to vacillate between these two positions. At times, he seems to propose an independent norm that limits the democratic process to only those questions that need to be decided collectively. The case for majority rule, he argues, is stronger in cases where a collective decision is necessary, like in foreign policy, but weaker when we can make our own decision.¹⁶⁶ A norm like this would first decide which activities required a collective decision and, on those that do not require collective action, it would create a sphere of privacy, no matter who would fundamentally object.¹⁶⁷

At other times, however, Cohen clearly rejects such a norm. In this vein, he writes: “In particular, I cannot see how we could possibly identify the private arena with the family, or with the economy, or with any arena of social life, identified... prior to normative political argument... Thus the thesis that a decision is private, and ought not therefore to be regulated except for compelling reasons, is not best understood as a premise in political argument... but as a conclusion of such argument.”¹⁶⁸ If every sphere is, in part, a result of collective decisions, then it would seem that we need a collective decision to regulate or not to regulate that sphere – coercion and non-coercion stand in need of justification.

However we should decide this debate about asymmetry, none of the solutions to this dilemma is open to Cohen, as I have interpreted him. To propose a default position of non-coercion or a private sphere of rights, Cohen must argue one of the following: 1) There are norms of freedom or rights prior to deliberation. But, that possibility is ruled out by his procedural theory,¹⁶⁹ which denies that there are any norms prior to the value of the (in this case democratic) procedure.

Or, 2) Argue that these norms are inherent in the deliberative ideal, which Cohen describes as reasoning together among equals. But, reasoning together requires that reasonable people only offer each other views of the good that they reasonably expect

¹⁶⁵ Colin Bird, “Coercion and public justification,” *Politics, Philosophy and Economics* (3): 1470594-13496073.

¹⁶⁶ “I said earlier that the proponents of enforcement often appeal to the value of democracy, urging that the equality of citizens requires that majority values fix the moral environment. This point has considerable force when a collective choice is necessary, as, for example, in the area of security policy: there, we need to arrive at a community decision, so the majority may speak in the name of the community... But where regulation is unnecessary, as in the area of sexual morality, this rationale is unavailable.” Joshua Cohen, *Philosophy Politics Democracy*, (Cambridge: Harvard University Press, 2009), p.263.

¹⁶⁷ Cohen also writes: “Moral liberty – like religious and expressive liberty – is an ingredient in the democratic idea of collective choice by free and equal citizens. The decision to regulate deep interests cannot be collective; it cannot arise from free reasoning among equals.” *Ibid.*, p.264. But, what about the decision not to regulate when deep interests call for regulation?

¹⁶⁸ *Ibid.*, p.312-313.

¹⁶⁹ “Participants do not regard themselves as morally bound by the existing system of rights except insofar as that system establishes the framework of free deliberation among equals. That system is an object of deliberation.” *Ibid.*, p.233.

could persuade. If “religious people cannot accept...considerations that would preclude their fulfilling those commitments,” then pro-lifers cannot accept the considerations that support the principle: ‘In case of conflicting commitments, presume freedom and default to no law.’ Accepting such a principle would preclude them from fulfilling their commitment. In Cohen’s theory, freedom and equality are merely procedural values; they are about how we situate parties in deliberation, not what rights they have prior to it.

Or, 3) Claim that any view of the good that sought to regulate another’s actions would not survive rational scrutiny, while views of the good that supported liberty would survive. This argument, however, is implausible and contrary to the interpretation of rational scrutiny that deemed a variety of comprehensive doctrines “fully reasonable.” These three possible answers seem to exhaust the resources of a procedural theory.

Cohen offers one more possibility to draw liberal rights out of his diverse deliberators. As we saw, the abortion debate, and many others like it, results in a deadlock between opposing fundamental commitments. The way out of this, he argues, is to say that not all fundamental commitments can be vetoed after all. We only need to take a person’s commitments into account if they are being “burdened” by a regulation. So, in the case of abortion, pro-life Catholics and pro-choice Liberals are equally irrelevant to what count as acceptable public reasons, since neither is (necessarily) directly affected. The only views that matter are those of the people burdened, which, in this case, are those of women.

Their fundamental commitments are burdened in three ways: 1) There is a burden on women’s equality: the ability to participate equally in social and economic life; 2) A burden on women’s liberty: restricting a deeply personal choice; and, 3) A burden on judgment: Preventing women from exercising their own judgment on a weighty issue.¹⁷⁰ Since only the commitments of burdened individuals matter in public reasoning, we no longer have to take into account those of our Liberal and Catholic objectors, thus avoiding the deadlock.

How can we derive the principle that only the commitments of those burdened can serve as vetoes on public reasons – that only the burdened are entitled to reasons they cannot reject? Again, it does not seem to come directly from the democratic ideal, nor from its interpretation as reasoning together among equals. Nor does it seem to be a result of the idealizing process, whether rational scrutiny or any pressure toward consensus on views of the good.

Moreover, some of these particular burdens do not seem to be ones that all the representatives could accept. The values of making deeply personal choices for oneself or exercising one’s judgment on a weighty issue are not values that all share, irrespective of the choices and issues involved. What is at issue in the abortion debate is precisely whether terminating a pregnancy is solely a personal choice or not, so the value of making a choice is in question. As Cohen rightly argues, what is personal and what is public must be the conclusion of an argument, not its premise. And, exercising one’s

¹⁷⁰ Ibid., p.314-315.

judgment on a weighty issue may be a shared value, but it is not a shared value to think that only the judgments of the burdened should carry any weight. It is hard to see why these self-regarding commitments would carry the day unless we were already committed to something like individual autonomy – unless the reasonable democrats were already liberals.

Instead of committing reasonable people to the principle that only the commitments of the burdened matter or the acceptance of these liberal-seeming burdens, perhaps we can reframe what reasonable people agree to in a way that is more congenial to all comprehensive doctrines. Perhaps we can argue that all doctrines could agree that there is a distinction between sins of omission and sins of commission. When one violates a fundamental commitment that one has full control over the fulfillment of, that is a sin of commission and all can agree that that is a more significant offense against one's own doctrine. When one violates a fundamental commitment that requires the compliance of others, however, that is a sin of omission and all can agree that this is a less significant offense against one's doctrine. In other words, comprehensive doctrines can all accept that, loosely speaking, ought implies can. Unlike the burdens distinction, this one would take comprehensive commitments seriously and give them weight, while acknowledging the limits of one's control.

So, while both are sins against one's comprehensive doctrine, sins of commission are greater sins and any law that forced us into such sins would be unacceptable. Smith could not accept a law that prevented him from using peyote in religious ritual because he has full control over his fulfillment of that commitment. But, if there were a religious group that had a commitment to ban all recreational drugs, the failure to do so would be less of a sin, from their perspective, because they can acknowledge that it is more difficult to force others to comply. A law that allowed peyote for religious use would therefore be more acceptable to them. If we can make a distinction like this, we can then say that only those laws that preclude the fulfillment of self-regarding commitments would be unacceptable to reasonable people.

Even if all comprehensive doctrines could agree to such a distinction in many cases, however, two problems remain. First, there may be issues on which some reasonable people will not be able to agree that sins of omission are less of a violation of one's doctrine than sins of commission. Some commitments may just be too important to excuse ourselves from our responsibility to pursue them, even if it requires others' compliance. The abortion issue is a prime example. For ardent Christians, the prohibition on abortion is not just a requirement that they themselves do not terminate pregnancies but that no one should. From their perspective, the fundamental commitment seems to be to prevent what they see as infanticide, not just to keep their conscience clean. So, a pro-choice law may be just as unacceptable to them as a pro-life law would be to others.

The second problem is that giving a veto only to commitments that one can personally fulfill and not to ones that require the compliance of others would not always produce outcomes that liberals would want. Take the 2015 Religious Freedom Restoration Act (RFRA) in Indiana, for instance. The issue there is that some cake-makers and wedding

photographers feel that they have a fundamental Christian obligation not to participate in gay weddings. Liberals want the state to regulate their conduct to prevent them from discriminating against gay people in what is plausibly a fundamental interest of all people – the ability to participate equally in social and economic life. So, this is again a clash of commitments, which means that public reasons such as equality of opportunity or equal protection carry no weight.

However, under either the burdens or the commission distinctions, if the only commitments that do have weight are those of the people being regulated, then only the photographers and bakers will have a veto on the laws. They are the ones being required to act in a manner that they oppose. So, they can veto any anti-discrimination laws as based on reasons they cannot accept, while the opposing commitment that calls for equality would have no veto. This would be a problem anytime the person burdened had a commitment to a result that liberals would oppose, such as, arguably cases of Christian Scientists, those opposed to vaccinations, or the Amish seeking to keep their children out of school.¹⁷¹ Perhaps liberals persuaded by Cohen's theory would simply accept these results, but these two problems suggest that there will be a number of issues on which public reason, on its own, will not lead to traditionally liberal results.

Conclusion

For public reason liberals like Rawls and Quong, this would be a problem. As they see it, the issue is that a procedural theory does not begin with a specifically liberal group of reasonable people, so there will never be a guarantee of a liberal output. And, I think they are right to worry. By validating all (fully reasonable) fundamental commitments, both liberal and non-liberal, Cohen's theory removes two defenses of liberal results: 1) most liberal values cannot be public reasons because they would not be shared by all comprehensive doctrines, and 2) there will be stalemates between liberal and non-liberal commitments that his procedural theory does not have the resources to resolve on the liberal side.

However, as I have already argued, Rawls and Quong are able to guarantee the justification liberal norms only by transforming reasonable people in ways that are inconsistent with the justifications of their theories. So, ultimately, I must conclude that public reason theories must lower their sights. For these theories to remain consistent with their justifications – which require only a modest idealization – they cannot aim at a robust liberalism. Nevertheless, I will argue in the final chapter that there may be values that reasonable-but-realistic people could accept, such as, as Cohen suggests, the value of deliberative procedures.

¹⁷¹ Perhaps one could even argue that libertarians' equation of taxes with slavery would constitute a fundamental commitment and allow them to veto redistributive laws.

Chapter Six – Public Reason and Private Bias

In this chapter, I argue that the problem of reasonable pluralism is even deeper than public reason theories have conceived of it. Public reason's diversity arguments explain why it is important to take into account reasonable pluralism, such as the diversity of religious views. We have seen several possible arguments for that; for instance, religious diversity is inevitable under liberal conditions or is the most we can reasonably expect of ordinary reasoners. Those same arguments, however, will require us to accommodate some beliefs that are the product of irrational influences as well. If so, then reasonable people will be, on some issues, irrational people, and their reasoning will no longer be mutually justifiable. This is a problem whether we are thinking of public reasons as those used in real-world deliberative processes or in a hypothetical model to decide political norms. We therefore need a new way of thinking about public reasoning that is consistent with some of our irrationalities but can still justify appealing norms. I close by sketching this new direction.

Introduction

Public reason theories begin with the thought that, to be legitimate, coercive state power should only be used on the basis of norms – laws, constitutional principles, or social rules – that all reasonable citizens could accept. And, all reasonable citizens “could accept” a norm when the norm is based on the right sort of reasons – “public reasons.” Theorists have offered two general conceptions of “public” reasons. “Shareable” reasons are those that all reasonable citizens share based on their common identity as citizens, rather than their private identities as subscribers to particular religions or philosophies of life. “Intelligible” reasons, on the other hand, are those that follow from one's fundamental religious and ethical beliefs. Other citizens might not share your intelligible beliefs, but they can see how those beliefs are reasons for you. Whichever conception of public reasons we consider, however, their goal is the same: to legitimate coercion. If a political norm is based on the right sorts of reasons, then it is legitimate for the state to coerce you based on that norm, whether you think it is the best possible norm or not. Because the law, principle, or rule is based on shareable or intelligible reasons, it is, in some sense, a norm that you “could accept,” so the state is entitled to act on it.

This public reason requirement is applied in two different ways – as a requirement on political norms or as a requirement on citizens¹⁷². As a requirement on norms, the public reason requirement is meant to ensure that state coercion is based on reasons that the coerced can accept. For the same reason, some theorists have argued that citizens' *proposals* to use state coercive power should also be based on reasons that the coerced can accept. After all, in a democracy, the best way to ensure that political norms are based on public reasons is for all citizens to rely only on public reasons when they argue and vote for those norms. So, some theorists argue that citizens have an ethical (though

¹⁷² John Rawls calls this requirement on citizens “the duty of civility.” John Rawls, *Political liberalism* (New York: Columbia University Press, 1993) p.217.

not legal) duty to ground their proposals on public reasons, or at least to know that they could do so.

As a requirement on the reasons that citizens offer each other, however, public reason often does not serve its intended purpose, which is to make our coercive proposals more acceptable to each other. Three decades of social psychology research have suggested that we don't always acknowledge, or even necessarily know, the true bases of our conclusions. So, we may think we are reasoning based solely on public reasons, while actually basing our conclusions on some factor that others can't accept, such as an implicit racial bias. For issues that we have good reason to worry are affected by such biases, the justificatory value of public reasons is undermined, making our conclusions no more acceptable to others than if we did not offer public reasons at all.

The requirement on norms seems to be able to avoid this problem of "objectionable influences." After all, even if citizens *actually* arrive at norms in objectionable ways, we can still ask whether they *could* have arrived at those norms in acceptable ways. However, I will argue that the problem of objectionable influences persists, even for the requirement on norms.

To claim that a norm *could be* based on acceptable reasons, theorists must tell us which reasons are ones that all citizens could accept. We have already seen that "reasons that all citizens could accept" are not necessarily reasons that they do accept, since the reasoning of actual citizens is flawed in various ways. Instead, theorists identify these public reasons by first conceiving of citizens without some of these flaws – realistic enough that their acceptance still matters, but idealized enough that their reasoning is acceptable to each other. For instance, all public reason theorists claim that suitably idealized citizens would retain their diverse religious beliefs, but in more coherent forms.

As I will argue, however, the rationales given for retaining religious beliefs also apply to certain objectionable influences on reasoning, such as implicit bias and motivated reasoning. So, for the same reasons that suitably idealized citizens must retain their religious beliefs, suitably idealized citizens also retain certain biases.¹⁷³ This means that some of the conclusions that could be accepted by all suitably idealized reasoners are arrived at based on biases, which undermines the justificatory value of those conclusions in the same ways that actually biased reasoning undermines our political proposals to other citizens.

If this is right, then it seems impossible to meet the public reason requirement. If we must justify norms to biased people, and their biases are not justifiable to others, then it is hard to see how we can find norms that are justifiable to all. The "public's" reasons appear to

¹⁷³ In the first chapter, I argue a weaker version of this claim. There, I argue that the justification for accommodating pluralism requires us to attribute certain realistic qualities, and perhaps even cognitive limitations, to suitably idealized people. Here, I argue that the problem is even deeper. The justification for accommodating diverse religions also requires us to accommodate irrationalities, rather than just limits on our rationality. This irrationality makes it unreasonable to expect that all suitably idealized people would accept liberal conclusions, as public reason theorists claim they would.

be too unreasonable. However, I will conclude by suggesting that, while public reason theories are committed to justifying norms to biased people, they are not committed to justifying norms to people who are unaware of these biases. Put positively, suitably idealized citizens are biased but aware of it. So, public reasons are not just considerations like common defense or general prosperity, but a shared awareness of our cognitive tendencies and limitations. Reasonable (or idealized) people are aware of the conditions and issues on which they are likely to be *unreasonable*.

So, when idealized citizens seek to justify political proposals to others, they include measures to counteract, or at least compensate for, these conditions. Public justification is therefore not just a matter of relying on the right sorts of reasons but on supporting a system under which our errors are corrected or mitigated. This more holistic conception of public justification, I suggest, has greater potential to defend liberal rights (and important values) than the exclusive focus on individuals' reasons.

The Public Reason Requirement on Citizens

John Rawls initially¹⁷⁴ argued that, as part of the public reason requirement, individual citizens (and lawmakers) have a “duty of civility” to rely only on public reasons in arguing and voting for their political proposals. When a citizen offers others only public reasons, her proposal is supposed to be acceptable to them, even if they still disagree with the proposal. To see how this might work, we must first consider what makes a reason public and how its public-ness contributes to mutual acceptability.

There are at least three different conceptions of “public reasons.” In the colloquial version of the theory, public reasons are shared reasons – beliefs and values that all or most liberal-democratic citizens share. John Rawls referred to this as the “public culture,” which contains values such as public safety, economic growth, and equality, but not reasons that we disagree about, such as religious and metaethical claims. To make an already abstract theory a little easier to understand, I will mostly refer to public reasons in this sense – as broadly shared reasons – although that is often less precise than what theorists have in mind.

More precisely, Rawls often conceives of public reasons as *shareable*, rather than actually shared. This leads to the more awkward formulation of: ‘reasons that you could reasonably expect other reasonable people to share.’ In other words, public reasons may not *actually* be broadly shared – perhaps due to ignorance or indifference – but it is reasonable to expect that others would accept those reasons, if others were exposed to them and were thinking reasonably well.

The third sense of “public” reasons is, by design, much weaker. In “convergence” public reason theories, public reasons do not have to be shareable by all, but they do have to be “intelligible” to all. That is, as a Hindu, I do not have to share your Christian beliefs, but I do have to see how, if one had your reasons, one would arrive at your conclusions. If you

¹⁷⁴ As we will see in the next section, Rawls later relaxed that requirement somewhat but, as I will argue, I'm not convinced that this was an improvement, so I will evaluate both versions.

and I both have intelligible reasons to support the same law – if we “converge” on that law – then that law is acceptable to both of us.

As I said, I will mostly discuss public reasons as broadly shared reasons, but I believe my argument would apply to all three senses. All three consider political norms to be acceptable to all when they are based on the “right sort” of reasons – shared, shareable, or intelligible. So, an argument that undermines the justificatory value of our reasons would be a challenge to all three.

Before I turn to such an argument, let me first outline how public reasons are supposed to contribute to mutual acceptability. In other words, how does the fact that I offer you only public reasons to support my proposal make that proposal – which you disagree with – nevertheless acceptable to you? There are three ways in which public reasons might make a proposal more acceptable. First, for shared reasons conceptions, the public reasons themselves are reasons you find valid – you share them. So, at least the content of my reasoning is acceptable to you, even if you disagree with how I weigh my reasons or the conclusion I draw from them. Second, for all conceptions, the fact that I rely only on public reasons in arriving at my conclusion shows you that I arrived at it in a way that you can accept – I followed an appropriate reasoning procedure, which gives you more confidence in my conclusion. There were no known distorting factors or other reasons for you to believe that I was not reasoning clearly or in good faith. And third, offering you public reasons shows you that I am a reasonable person – that I come to conclusions based on reasons and evidence, rather than my own wishes, and that I will therefore be responsive to further evidence and better arguments.

In some or all of these three ways, public reason theorists argue, requiring citizens to offer each other only public reasons makes their proposals mutually acceptable. However, recent work in social and political psychology has suggested at least two ways in which the justificatory value of public reasons is undermined. First, there is good reason to worry that, while we might claim to base our proposals solely on public reasons, we are actually influenced by reasons that others would reject. For instance, Martin Gilens has argued that the strongest predictor of anti-welfare views among survey respondents is a respondent’s belief that welfare recipients are undeserving of government assistance.¹⁷⁵ So far, that seems like a perfectly justifiable reason. Most Americans believe that merit or desert¹⁷⁶ is a valid consideration in deciding how to redistribute wealth.

However, Gilens also found that the strongest predictor of respondents’ beliefs that welfare recipients are undeserving is the belief that African-Americans are lazy¹⁷⁷, which

¹⁷⁵ Martin Gilens, *Why Americans Hate Welfare* (Chicago: The University of Chicago Press, 1999), p.63-65.

¹⁷⁶ Rawls would argue that “pre-institutional” desert is not a public reason, but here we can think of “post-institutional” desert, since people could be basing their judgments on the way institutions operate and what they expect from people. It would then be a public reason to say that: If people conform to those expectations, they “deserve” the rewards.

¹⁷⁷ *Ibid.*, p.71.

is *not* – at least for most¹⁷⁸ public reason theories – a reason that is acceptable to all. Nevertheless, because these respondents do not acknowledge – or perhaps even know – that their anti-welfare views are based on this stereotype, they can “justify” their views to others, despite its objectionable basis. Cases like these pose a serious problem for a public reason requirement that applies to individual citizens. If the public reasons one presents to others are selected or emphasized, knowingly or unknowingly, in whole or in part, because of an objectionable influence, then that undercuts the justificatory value of those public reasons. The public reasons no longer make the political norms they support justifiable to all.

Now, Rawlsians might try to exclude cases like these by adding a sincerity requirement to the requirement to rely on public reasons: each citizen must sincerely believe that the reasons they offer others are valid reasons and that they are sufficient to support their conclusion. This sincerity requirement seems like it could solve the problem of objectionable influences by weeding out those that are using their reasons as a cover for some deeper desire. If the public reasons were a mere cover, that person would not regard them as sufficient to support the conclusion on their own and would therefore be offering them insincerely. For instance, I might claim that expanding offshore drilling will serve the public interest in energy independence, when I actually support it because I own a refinery. That claim would fail a sincerity requirement.

But, adding a sincerity requirement cannot address a second way in which our psychology may undermine the justificatory value of public reasons. In many cases, we are not intentionally hiding our prejudice behind a public reason; we are simply unaware of it. Many social and political psychologists have argued that, since implicit biases like these merely reveal the norms and realities of the environment a person grew up in, nearly everyone has them to some degree. Implicit biases are the automatic and instantaneous connections we make between groups and associated ideas, traits, or emotions. These associations happen so quickly that we are not even aware of them or how they affect our thinking. The challenge of implicit biases is that we can be fully sincere in invoking public reasons and supporting conclusions based on them. And, even if we are unconvinced by the implicit bias literature, we can still posit some form of unconscious reasoning that relies on stereotypes in ways that we are unaware of. If these influences draw us to invoke this public reason rather than that, or to weigh it more heavily, then it is influencing our reasoning in ways that others could not accept.

For example, in the welfare case, only 44% of respondents reported holding the stereotype of African-Americans.¹⁷⁹ But, that does not mean that the remaining 56% did not have their reasoning influenced by this or other negative associations with African-Americans. Gilens argues that repeated associations in the media and culture between

¹⁷⁸ It is certainly not for Rawlsian “consensus” theories. But, even “convergence” theories would, I think, consider stereotypes to be unjustifiable. Gerald Gaus, for instance, argues that a person’s biases and heuristics-based reasoning are not justifiable to others, and stereotypes would fall into one or both of those categories. But, I am not certain what they would say, so I’ll leave the claim at “most public reason theories.”

¹⁷⁹ Ibid., p.68.

poverty, welfare, African-Americans, and laziness work together to produce explicit biases against African-Americans and therefore a negative view of welfare.¹⁸⁰ But, that same mechanism can also produce an *implicit* association between those things, such that, when welfare is mentioned, the associated group and trait immediately come to mind, without one's awareness, which is why implicit biases would never show up in Gilens' survey. So, even when one thinks one is considering exclusively public reasons such as whether the poor are deserving of government assistance, one may already have an implicit image of what the poor look like and an implicit feeling about their merit. This "information" comes to mind *before* one considers new evidence, seeks out reasons, and weighs the considerations, so it inevitably affects every step of the reasoning process. This is what Milton Lodge and Charles Taber call "affective contagion" because one's implicit "affect" (or feeling) not only changes one's conclusion about welfare, but the course of one's reasoning as well.¹⁸¹

To take an example from a different political issue, Efrén Pérez has argued that much of the opposition to immigration is driven by implicit biases against Latinos.¹⁸² Like Gilens, he cites a longstanding pattern of associations in the media between illegal immigration and Latinos, despite the fact that a substantial portion of illegal immigrants are not Latino and most Latino immigrants are not here illegally. This near-exclusive focus by the media has led to an automatic association in many people's minds between immigration and Latinos and Latinos and something negative (illegality), which Pérez demonstrates through both survey work and a test for implicit bias, the Implicit Association Test (IAT). He argues that those who possess these implicit biases are already predisposed to favor reasons against immigration, before they even consider the reasons in the immigration debate:

...broaching the issue of immigration will spontaneously activate people's implicit attitude toward Latino immigrants, thus making these evaluations mentally accessible to people...Lodge and Taber (2013) have established that affective responses like these often bias one's retrieval of considerations from memory, in particular, by recruiting considerations that are congruent with one's initial affective response. I extend this *affective contagion* mechanism to one's processing of political information. Specifically, I hypothesize that people will judge political issues according to the direction and intensity of their implicit attitudes, even when these are contradicted by explicit political information immediately before them. In other words, individuals will interpret political information in a manner that is consistent with the direction of their implicit attitudes. This takes place, I maintain, because the interplay between implicit attitudes and explicit political reasoning is subconscious

¹⁸⁰ Ibid., pp.102-153.

¹⁸¹ Cengiz Erisen, Milton Lodge, and Charles Taber, "Affective contagion in effortful political thinking," *Political Psychology*, 35(2), 187-206.

¹⁸² Efrén Pérez, *Unspoken Politics: Implicit Attitudes and Political Thinking* (New York: Cambridge University Press).

– that is, citizens are unaware of how their implicit attitudes shape the more deliberative aspects of their decision making...¹⁸³

So, between explicit and implicit biases, there are many potential influences on our public reasoning that are not acceptable to all. This would seem to be a serious problem for a public reason requirement that required citizens to justify their proposals to each other. Rawlsian public reason theories argue that, if I present you with reasons that we both share (or could reasonably be expected to share) and those reasons support my conclusion, then that is sufficient to make my conclusion acceptable to you. But, if there is good reason to worry that I unknowingly selected, emphasized, or interpreted those reasons, even in part, because of an unshared further reason or influence, then my conclusion seems no more acceptable to you now than before I offered you any public reasons.¹⁸⁴

If you have good reason to worry that my reasoning is influenced in objectionable ways, it undermines all three ways in which public reasons were supposed to make my proposals justificatory. First, if the principle behind public reason theories is that a conclusion is acceptable to others when one arrives at it in an acceptable way, then the fact that a major factor in my conclusion is one that you actively reject means my conclusion has violated that principle. Second, the reasons that I presented to you, although “public,” now appear to be a mere cover for a deeper influence, which undercuts their justificatory value as reasons. And third, if offering you public reasons was supposed to show you that I am a reasonable person – that I come to conclusions based on reasons and evidence, rather than my own wishes, and that I will therefore be responsive to further evidence and better arguments – then the fact that some part of my mental process is immune to reason is a problem. It is a problem that, even when my implicit attitudes “are contradicted by explicit political information immediately before” me, my reasoning is still guided by unconscious biases rather than reason and evidence. This shows that, on this issue at least, I am an *unreasonable* person.¹⁸⁵

¹⁸³ Ibid., p.112.

¹⁸⁴ This would also apply to the convergence version of the theory. If you and I can both accept the same norm based on intelligible reasons – reasons that follow rationally from our core beliefs – then that norm is justified to both of us. But again, the presence of an objectionable influence, such as an implicit racial bias, skewing our reasoning seems to make our reasoning irrational and therefore not intelligible.

¹⁸⁵ A Rawlsian might respond by citing his “proviso” to the requirement on citizens. Instead of requiring that citizens arrive at their conclusions in a way that others can accept, the proviso insists only that they know that they *could* have arrived at their conclusions in an acceptable way, even if they did not. It is meant to allow for public arguments that invoke private values, such as, for instance, the civil rights protests that invoked the invoked Biblical claims, such as the equality of all people in God’s eyes. As long as the proponents of civil rights knew that they could support the same civil rights based only on public reasons, their actual reasoning was acceptable to all, even if it were influenced by claims that others could not accept.

However, in addition to the fact that the proviso is intuitively problematic (how could a racist saying that he *could* provide a non-racist defense of his proposal justify it to others?), it does not solve the problem of objectionable influences. After all, every version of the public reason requirement allows for the possibility that citizens think there’s a better, non-public, argument for the same conclusion. It simply requires them to offer a public defense that is sufficient to justify the conclusion. The proviso just makes this explicit by allowing citizens to cite both the argument they consider best, and a sufficient public argument. However,

Requirement on Political Norms

A public reason requirement on citizens states that political norms are publicly justified when citizens arrive at those norms in mutually acceptable ways. However, as I have just argued, we often have cause to worry that what appears to be mutually acceptable reasoning – the exclusive use of public reasons – is actually influenced in unacceptable ways, such as by explicit or implicit biases. This undermines the justificatory value of relying on public reasons and therefore the value of the public reason requirement on citizens.

Perhaps for that reason, many theorists have argued that the public reason requirement should apply only to political norms – laws, political principles, or social rules – and not to citizens’ proposals for those norms. As a requirement on norms, the public reason requirement states that norms are publicly justified when citizens *could* arrive at those norms in a mutually acceptable way, even if most citizens did not arrive at them in this way. (Indeed, even if most citizens did not arrive at them at all.) If citizens could arrive at a norm in a mutually acceptable way, then it is legitimate for the state to coerce citizens based on that norm. By ignoring the way that actual citizens arrive at their proposed norms, a public reason requirement on the norms themselves may be able to avoid the problem of objectionable influences.

Now, how do we know if citizens “could arrive” at a norm in an acceptable way? One possibility is that we could simply ask whether a proposed political norm – let us say, the rejection of welfare – could be supported by only public reasons. Can we make a case against welfare without relying on non-public (non-shareable or unintelligible) reasons? It seems like we can. Merit (or desert) is a reason that all or most citizens recognize as a point in favor of a norm, and it is certainly an intelligible one. So, we could argue that people should not get what they do not earn, and what they earn should be determined by the market for their talents and industry. If the market does not reward them, it is because they did not earn it. Whether this argument is right or wrong, it seems like it is made solely on the basis of public reasons, which meets the requirement for a legitimate norm.

However, this skips a crucial step in the public reason argument. How do we know that merit is a public reason? We need to know whether the public does, or could, accept it as a reason. But, as I have argued, that cannot mean just looking at whether most actual citizens accept it as a reason, since there is reason to worry that many citizens accept it based on unacceptable influences, such as implicit bias. Instead, we must ask whether citizens “could accept” merit as a reason. But, again, what does “could accept” mean?

since both arguments are coming from the same person, if we have reason to worry that their reasoning is tainted by objectionable influences, then both the private and public arguments are equally worrisome, no matter which they regard as best. If they are implicitly biased against Latinos, for instance, we still have reason to worry that the public values they consider sufficient to justify their conclusion on immigration were actually chosen, interpreted, or ranked based on bias and not on their merits. So, the public reason requirement on citizens with the proviso is susceptible to the same problem: on issues where we have good reason to worry that public reasons are influenced in ways that others would reject, invoking public reasons is undermined.

Public reason theorists answer this question by considering whether citizens could be expected to support a reason *if* they were reasoning fairly well. And this makes some intuitive sense. Why should we care about which political norms actual citizens accept, if actual citizens are often grossly deficient reasoners? Instead, public reason should only require that political norms be acceptable to *reasonable* people – decent citizens who are reasoning fairly well.

While the concept of reasonable acceptance is intuitive, however, the argument to demonstrate it is anything but. We can think of it as a three step process. The first step in this process is to define a “reasonable” citizen. How would citizens think if they were more decent and reasoning fairly well? Obviously, we cannot imagine them all as perfect reasoners who have the whole truth, or the idea of acceptability to a diverse citizenry is completely lost. Instead, public reason theorists must *partially idealize* citizens – improving some aspects of their actual attitudes, while retaining others just as they are. And, crucially, theorists must justify this degree to which they partially idealize citizens. These justifications, I will soon argue, revive the problem of objectionable influences, even after idealization. However, for now, it may be helpful to simply note that most public reason theorists retain each citizen’s fundamental religious and ethical views, while improving at least the coherence of those views, and sometimes much more.

Once theorists have conceptualized these (partially) idealized citizens, the second step in identifying public reasons is to see which reasons the idealized citizens share (or which are shareable or intelligible) – these are deemed “public reasons.” Now, unlike the reasons that actual citizens share, the public reasons of these idealized citizens are not susceptible to the problem of objectionable influences. Since these idealized citizens can be defined without the kinds of distorting backgrounds and biases that real citizens suffer from, the way that idealized citizens select, interpret, and rank public reasons cannot be influenced in objectionable ways. So, these idealized citizens are defined to arrive at their public reasons in a mutually acceptable way – that is, in a way that other idealized citizens would accept.

In the third step, theorists see which norms could be supported based only on these public reasons. If a norm could be supported based only on the public reasons of idealized citizens, then the norm has met the test of public justification. By showing that suitably idealized citizens can arrive at that norm in a mutually acceptable way, this shows that their real-life counterparts could also have arrived at that norm in a mutually acceptable way, if they were reasoning well. And if actual citizens *could* have arrived at the norm in a mutually acceptable way, then it is legitimate to coerce those citizens based on that norm, regardless of how (or whether) actual citizens arrived at it. So, to find out which norms are acceptable to reasonable people, we should imagine the attitudes of citizens if they were being reasonable, and then consider what people with such (improved) attitudes could reasonably be expected to accept.

Idealizing away people’s biases is the simplest solution to the problem of non-public influences, and is likely the reason that most public reason theorists do not worry about

any objectionable influences on public reasoning. Because we have abstracted from the experiences and incoherence of real people, we can now be confident that, when an idealized person holds a belief in the importance of individual merit, this is not a cover for a stereotype about African-Americans or the result of an automatic association with negative feelings. In other words, we can confidently claim that that norm is not based on objectionable influences, such as implicit bias. Gerald Gaus, for instance, explicitly argues that political norms should not be justified to actual citizens (which he accuses John Rawls of doing) because actual citizens are susceptible to all sorts of fallacious and biased reasoning.¹⁸⁶ Instead, we should represent actual citizens as their best selves by, essentially, making their reasons more coherent. We then ask which political norms could be justifiable to this more coherent set of reasons.

As appealing as this solution is, however, it is not available to public reason theories. Public reason theories are committed, on pain of internal inconsistency, to including certain human biases, even in the idealized model of citizens' reasoning. This is because they are committed to including influences that lead to diverse religious views in the idealized model. Depending on which rationale they give for including these influences, however, they will therefore have to include at least some forms of human bias or irrationality, such as implicit biases and motivated reasoning.

Religious and Non-Rational Influences

All public reason theories are committed to claiming that even idealized citizens will retain their diverse religious beliefs in some form, and that this claim is relevant to determining political norms. But, why is the fact that citizens hold diverse religious beliefs relevant? After all, since many religious claims are irreconcilable, some of them are necessarily false. So, why should false claims be relevant to determining legitimate political norms?

Part of the answer must involve the relevance of non-rational influences. If the only thing that influenced idealized citizens' reasoning were purely rational factors, there would be no reason to expect citizens to arrive at such different religious beliefs. Perhaps one could argue, as Gaus does¹⁸⁷, that even fully rational reasoners would arrive at different conclusions from the same premises. However, it would stretch the concept of rationality very thin to argue that exclusively rational reasoning would produce mutually irreconcilable and wildly different religions. It is only by introducing some non-rational element into human reasoning – different upbringings, for instance – that we should expect even idealized reasoners to arrive at such different belief-systems. So, public reason's commitment to religious diversity in the idealized model is also a commitment to the relevance of at least some non-rational influences.

In particular, public reason theories are committed to the relevance of non-rational influences that they claim cause religious diversity, even among idealized reasoners. So, by identifying the non-rational influences that make religious diversity relevant, I can

¹⁸⁶ Gerald Gaus, *Justificatory Liberalism* (Oxford: Oxford University Press, 1996) p.132-136.

¹⁸⁷ Gerald Gaus, *The Order of Public Reason* (Cambridge: Cambridge University Press, 2011), p.241.

then determine whether they would also make other forms of sub-optimal reasoning relevant, such as implicit biases. If, as I will argue, the same non-rational influences that lead to religious diversity also lead to *irrational* influences like implicit biases then, like religious diversity, biases are also relevant to determining political norms. They must therefore be included in the idealized model.¹⁸⁸

So, what is the rationale for including religious reasoning in the idealized model? Here are three possibilities that one can draw from the literature. Theorists have (or could) argue that we should represent religious reasoning in the idealized model because we should represent reasoning that is: 1) Reasonable to expect of ordinary people, 2) Reasonable to expect of highly competent reasoners, or 3) Coherent, given the inevitable effect of liberal rights. Let us consider them in order.

First, Rawls¹⁸⁹ and Gaus¹⁹⁰ argue, in their different ways, that we ought to represent the reasoning that we can reasonably expect from ordinary people or, in a slightly different formulation, that is the most we can expect from ordinary people. And, what we can reasonably expect from ordinary people is constrained by the non-rational influences of our circumstances and limitations. Gaus, for instance, argues that, since the purpose of the idealized model is to yield political norms that real citizens, at their best, could actually internalize, the reasoning for those norms cannot be beyond what we could expect (or the most we could expect) from an ordinary citizen under ordinary conditions. And ordinary conditions include the fact that ordinary people cannot spend a lot of time deliberating on moral issues, and that ordinary people only have access to easily available information.¹⁹¹ This is why we cannot idealize citizens to reject their religious beliefs.

¹⁸⁸ I respond to this worry in more detail later on, but in case it occurs to the reader now, I will preview my response. Public reason theorists might argue that, even if my argument succeeds – even if I can show that the non-rational influences that lead to religious diversity also lead to certain biases – that doesn’t mean we should include biases in the reasoning of idealized citizens. After all, idealized citizens are supposed to be the kinds of reasoners that we believe are *entitled* to determine political norms – a constituency that theorists design for that purpose. Theorists are therefore free to include the aspects of citizens that they intuitively think are relevant and exclude those that aren’t, without needing to say *why* they are relevant. And, intuitively, we think the influences that lead us to disagree about religion are relevant influences, while the influences that lead to implicit racial biases are not.

A major problem with this defense is that this intuition is not shared by public reason’s critics, so it cannot be used to defend the theory. The intuitive claim is that the (in at least some cases) false religious beliefs of citizens are relevant, in some way, to determining political norms. This may be an intuition shared by all public reason theorists, but it is decidedly opposed by all their critics. Critics of public reason deny that false beliefs should play any role in determining political norms; for them, the only reason that matters is right reason. Critics might further argue that, if theorists are entitled to idealize citizens based on their intuitions about who *should* decide political norms, then every theorist can endorse a different group; Kantians will argue that only Kantians ought to decide, and so on. To answer these critics, therefore, public reason theorists must offer a rationale for including the non-rational influences that lead to religious diversity, rather than simply an intuition. And, once they offer a rationale, it is fair to ask what other types of reasoning these non-rational influences produce.

¹⁸⁹ For instance: John Rawls, *Political liberalism* (New York: Columbia University Press, 1993) p. 224.

¹⁹⁰ For instance: Gerald Gaus, *The Order of Public Reason* (Cambridge: Cambridge University Press, 2011), p.266.

¹⁹¹ ¹⁹¹ “Given the realities of life, the demands on their time and resources, and differences in the sophistication of their reasoning, most simply would never understand the reasons this practice ascribes to

Such reasoning would exceed the ability of ordinary citizens, would take too much time, or would require information that is not close at hand. So, because we ought to represent what we can reasonably expect from ordinary people, and ordinary people are influenced in non-rational ways by certain limitations, we ought to represent their religious reasoning.

For other theorists, the reasoning of ordinary people is too low a bar. Instead, as Leland and Van Wietmarschen recommend, public reason theories should insist that political norms be justifiable to *highly competent* reasoners.¹⁹² To say that a norm is justifiable to a merely ordinary person, they argue, is a weak reason to accommodate sub-optimal reasoning. You do not, for example, think your child should have a say in major household decisions; he simply doesn't know enough. So, to convince you to take another person's views into account, you must not only believe in the moral equality of that person, you must believe in his epistemic merit. His epistemic competence would give you a good reason to accommodate his beliefs. If you were convinced, for instance, that a highly competent person held different religious views to yours, you would therefore have a strong epistemic reason to accommodate his religious views. For that reason, public reason theories must represent reasoning that is reasonable to expect of highly competent reasoners. And, since even highly competent reasoners hold a variety of religious faiths due to various non-rational influences that affect even the best of us, we ought to retain diverse religious beliefs in the idealized model.

Third, theorists might argue that internally coherent religious reasoning should be represented in the idealized model because it is an inevitable effect of valued liberal rights. For example, Jonathan Quong argues that we should begin by assuming the importance of basic liberal rights such as freedom of association, speech, and conscience. Given that an ideal liberal society would protect these rights, the liberal ideal should also accommodate the inevitable consequences of these rights, including its non-rational influences on our reasoning. One of the inevitable consequences of such freedoms is that people will form different communities, and those communities will develop different religious and ethical norms and practices. Since even the citizens of an ideal liberal democracy would have coherent but diverse religious beliefs, we should retain it in the idealized model.¹⁹³ This also implies that non-rational influences such as our idiosyncratic backgrounds and experiences are relevant to determining political norms.¹⁹⁴

them..." Ibid., p.255. Also see: "Even Members of the Public, who recognize their sufficient reasons, are of limited rationality, and as creatures of limited reasoning powers, when asked whether they have sufficient reason to endorse a rule, they must ask, "what are the alternatives we are deciding between, and what are the costs of refusing to endorse any of them?" Ibid., p.269.

¹⁹² R.J. Leland and Hans Van Wietmarschen, "Reasonableness, Intellectual Modesty, and Reciprocity in Political Justification," *Ethics*, 122(4), pp.721-747.

¹⁹³ Jonathan Quong, *Liberalism Without Perfection* (Oxford: Oxford University Press, 2011) pp.142-143.

¹⁹⁴ Rawls acknowledges this in the burdens of judgment. He argues that, even if we all somehow possessed the same evidence and values, our different backgrounds and experiences would lead us to assess that evidence and interpret those values differently, which leads to diverse conclusions that political norms must accommodate.

Asymmetric Responses and Motivated Reasoning

All three arguments for representing religious reasoning in the idealized model imply the relevance to idealized reasoning of certain non-rational influences, whether it is the conditions or limitations of human reasoning or our idiosyncratic backgrounds and experiences. These non-rational influences lead to religious diversity, even among idealized reasoners, which is why the idealized model reflects that diversity. And, if these non-rational influences lead to other sub-optimal types of reasoning, these too must be represented in the idealized model. Here, I argue that two types of sub-optimal reasoning – an asymmetric response to evidence and implicit bias – fit all three rationales for inclusion. Consider the following two stylized scenarios:

Scenario 1: Maria grows up in an ideologically liberal household in New York. Because of this background, she acquires both liberal ideological beliefs, as well as a strong identification with the Democratic Party. Her prior beliefs therefore align closely with those of liberal democrats. Like others with these backgrounds, her beliefs and party affiliation influence the way she acquires and responds to new evidence. For instance, one study showed that, when presented with evidence supporting the existence of man-made climate change, climate change “acknowledgers” tended to increase the strength of their belief that climate change was occurring. When presented with evidence that it was not occurring, however, their views barely changed. And the same was true of climate change deniers. Both sides responded to evidence asymmetrically.¹⁹⁵ A similar study showed that, when acquiring testimonial evidence, Democrats tended to give more credence when the same presenter was labeled a Democrat than when he or she was labeled a Republican.¹⁹⁶ So, Maria’s political background clearly influences her reasoning – particularly, the way she acquires and weighs new evidence.

This asymmetric response to evidence is a kind of reasoning that, in a public reason context, others would rightly object to. If you argued for a policy to combat climate change based solely on public reasons – on the science supporting the existence of man-made climate change and its likely consequences – the fact that I know you only give

¹⁹⁵ “We find that people who doubt that man-made climate change is occurring, and who do not favor an international agreement to reduce greenhouse gas emissions, show a form of asymmetrical updating: They change their beliefs in response to unexpected good news (suggesting that average temperature rise is likely to be less than previously thought) and fail to change their beliefs in response to unexpected bad news (suggesting that average temperature rise is likely to be greater than previously thought). By contrast, people who strongly believe that man-made climate change is occurring, and who favor an international agreement, show the opposite asymmetry: They change their beliefs far more in response to unexpected bad news (suggesting that average temperature rise is likely to be greater than previously thought) than in response to unexpected good news (suggesting that average temperature rise is likely to be smaller than previously thought). The results suggest that exposure to varied scientific evidence about climate change may increase polarization within a population due to asymmetrical updating.” Cass R. Sunstein, Sebastian Bobadilla-Suarez, Stephanie C. Lazzaro, and Tali Sharot, “How People Update Beliefs about Climate Change: Good News and Bad News,” *Cornell Law Review*, 102(6), p.1431.

¹⁹⁶ Jason Brennan, *Against Democracy* (Princeton: Princeton University Press, 2016), pp.39-40.

weight to evidence supporting that position would undercut the otherwise justifiable reasons you were giving me.

Of course, public reason theorists will deny that this kind of reasoning would be part of an idealized model of public reasoning; these kinds of flaws are exactly what they would idealize away since an idealized (or reasonable) person does not respond to evidence asymmetrically. However, compare the rationale for representing Maria as an asymmetrical reasoner to the three rationales for representing idealized citizens as religious reasoners.

Is asymmetrical reasoning typical of, or even the most we can reasonably expect from, an ordinary reasoner under ordinary constraints? Certainly. We have just seen that under artificial conditions, when presented with evidence and given time to consider, people engage in asymmetrical reasoning. That is even more likely to be true of ordinary reasoners until time constraints, where they tend to rely on heuristics (or shortcuts) in their reasoning. Rather than evaluate new evidence independently, which would take time, ability and effort, ordinary reasoners often only accept evidence that is in line with previous beliefs or presented by someone in their group. For example, Dan Kahan has argued that the distorting effects of party identity are the only way to explain the “clustering” of otherwise unrelated political positions that party members tend to take.¹⁹⁷ There is no ideology that explains the clustering of anti-gun control and anti-abortion views, for example; the best explanation is that people adopt the position of their party (or group) and then find reasons to justify it. In other words, rather than respond to the evidence, ordinary people often engage in motivated reasoning.

So, we can consider the second rationale for inclusion in the idealized model – perhaps highly competent reasoners are not prone to asymmetric reasoning? On the contrary, however, some argue that highly competent reasoners are *more* likely to engage in politically motivated reasoning, perhaps because they have greater ability to further their motivation – such as the desire to affirm their party’s platform – by summoning reasons to support it.

High numeracy—a quantitative reasoning proficiency that strongly predicts the disposition to use System 2 [“conscious, effortful”] information processing—also magnifies politically motivated reasoning. In one study, subjects highest in Numeracy more accurately construed complex empirical data on the effectiveness of gun control laws but *only* when the data, properly interpreted, supported the position congruent with their political outlooks. When the data properly interpreted was inconsistent with their predispositions, they were more disposed than low numeracy subjects to dismiss it as flawed. If this is how people use their reasoning proficiencies to assess evidence about contested facts in the real world, then we would expect to see exactly what observational studies consistently find: namely, a progressive increase in political polarization

¹⁹⁷ Dan Kahan, “The Politically Motivated Reasoning Paradigm,” *Emerging Trends in the Social and Behavioral Sciences*, ed. by Robert Scott, Stephen Kosslyn (New York: Wiley) pp.1–15.

as individuals of opposing outlooks become even more proficient in critical reasoning...¹⁹⁸

So, if the test for inclusion in the idealized model is whether reasoning should be reasonably expected of highly competent reasoners, then motivated reasoning (and perhaps other sub-optimal influences) appears to pass.¹⁹⁹

The third rationale for representing sub-optimal reasoning begins with our diverse starting points, which are an inevitable result of liberal rights. And we can certainly agree that, just as liberal freedoms lead people of the same religion to form “in-groups,” those freedoms also lead people with similar ideologies to do so. Given this over-exposure to one ideology or party affiliation, one can further argue that coherence demands an asymmetric response to evidence. Internal consistency would require us to evaluate new evidence based on its coherence with our other beliefs. If our prior beliefs were sufficiently skewed by over-exposure to one ideology, then it would be coherent to doubt contrary evidence and endorse confirming evidence.

For instance, prior to hearing new evidence about climate change, let us say that the liberals believed that there was a very high probability that man-made climate change was occurring. So, when they heard the evidence supporting that view, their other beliefs suggested that the best explanation for that evidence was that climate change was, indeed, occurring. However, when they heard evidence against the occurrence of man-made climate change, their other beliefs could have suggested that the cause of this new evidence was not the absence of climate change, but some human factor: conservative propaganda or random variation in the climate.²⁰⁰ Given that reading of the evidence, it

¹⁹⁸ Ibid., p.12.

¹⁹⁹ More support from Christopher Achen and Larry Bartels: “In fact, the more information the voter has, often the better able she is to bolster her identities with rational-sounding reasons...But she may be just as impervious to evidence as anyone else...” Christopher Achen and Larry Bartels, *Democracy for Realists: Why Elections Do Not Produce Responsive Government* (Princeton: Princeton University Press, 2016), p.268

“...many observers assume that “rationalization is probably greater for less informed citizens. (Aldrich, Sullivan, and Borgida 1989, 132). However, Danielle Shani’s analysis of responses to a wide variety of factual questions produced a good deal of evidence to the contrary, leading her to conclude that “political knowledge does not correct for partisan bias in perception of objective conditions, nor does it mitigate bias. Instead, and unfortunately, it enhances the bias; party identification colors the perceptions of the most politically informed citizens far more than the relatively less informed citizens.” (Shani 2006, 31).” p. 279

“The “essentially random partisan shock” of Watergate tended to affect the best-informed citizens’ views on a range of unrelated issues: “government jobs, school busing, and criminal justice procedures...If anyone has asked these well-informed citizens to explain the changes in their thinking about school busing or government employment programs between 1972 and 1976, we suspect that they would have provided rationalizations of exactly the sort posited by Rahn, Krosnick, and Breuning (1994, 592), “mentioning the reasons that sound rational and systematic and that emphasize the object being evaluated, while overlooking more emotional reasons and factors other than the object’s qualities.” p.291-293

²⁰⁰ “...Brian Gaines and his colleagues (2007) found in a series of repeated interviews with college students in late 2003 and 2004 that “all partisan groups, strong Republicans included, held reasonably accurate beliefs” about the Iraq War and updated those beliefs as circumstances changed; however, Democratic and Republican students differed greatly in their *interpretations* of pertinent facts, causing the authors to

was more coherent for them to dismiss it. In that way, it could be more coherent to respond to evidence asymmetrically based on one's ideological background. As Achen and Bartels write:

Better-informed voters may be logically consistent and even “rational” in the thin technical sense of that term, but the truth of the matter is that they – and all of us, most of the time – exist...comfortably in their false but efficient pseudo-environments...Democratic competence requires not only logical consistency and cognitive efficiency, but also some modicum of accuracy in perception and receptiveness to new and, perhaps, disconfirming evidence. In Amartya Sen's (1977) phrase, it is perfectly possible to be a rational fool...But their consistency is often driven by group loyalties and partisan biases, even when it comes to straightforward matters of fact...²⁰¹

So, whatever the rationale for representing religious reasoning in the ideal model, there appears to be a parallel rationale for representing non-rational and objectionable influences such as motivated reasoning. Now, an idealized model in which people's reasoning is inevitably influenced by their epistemic starting points or their “total life experiences” may seem vague enough to be innocuous. After all, these influences are natural and universal and often unidentifiable. However, it is much harder to say the same about an asymmetric response to evidence or motivated reasoning, which are equally “natural.” An idealized model in which people's reasoning is inevitably influenced by these objectionable influences – by their skewed prior beliefs or group identity – does not tell us which norms would be acceptable to reasonable people. Quite the opposite; on these issues at least, such people are *unreasonable*.

Implicit Biases and Stereotyping

The inclusion of motivated reasoning in the idealized model of public reasoning would be troubling enough, but there is an even more objectionable case.

Scenario 2: Adam grows up in a mostly racially homogenous community, and socializes in mostly racially homogenous groups, as many other White Americans do.²⁰² As Elizabeth Anderson argues, these facts alone can lead

conclude that partisans “effectively used interpretations to rationalize their existing opinions” rather than to rationally revise those opinions.” *Ibid.*, p.278.

²⁰¹ *Ibid.*, p.295.

²⁰² Residential segregation by race has declined somewhat over the last 30 years, but remains very high, with most cities in the United States needing between 50% and 70% of African-Americans to move to achieve an even distribution with whites. William Frey, “Census Shows Modest Declines in Black-White Segregation,” *Brookings*, Dec 8 2015, <https://www.brookings.edu/blog/the-avenue/2015/12/08/census-shows-modest-declines-in-black-white-segregation>

- An analysis of people's residential moves shows between 1977-2005 shows, among other things, that 56.8% of white families moved to a predominantly white neighborhood, while only 2% moved to black neighborhoods. Residential segregation by income has increased over that time to the point where, in 2010, 46% of housing was either in exclusively high or low-income areas. Daniel Fowler,

him to hold various stereotypes and biases about non-White racial groups that unconsciously affect his reasoning, even on issues that are not about those groups and even when his reasoning seems to rely only on public reasons.

How might socializing in a racially homogenous community lead to this objectionable reasoning? First, the racial segregation may lead him to have a host of biases toward less familiar racial groups. Elizabeth Anderson argues that segregation does so, initially, by producing actual differences between groups: higher rates of poverty, unemployment, different dialects, different social norms, denies services and role models and concentrates social ills.²⁰³ Adam's unfamiliarity with these other groups then produces psychological effects that exaggerate these real differences and skew his understanding of their causes. The "attribution bias"²⁰⁴ leads him to attribute the cause of these disadvantages to other groups based on individual qualities of their members, while he attributes disadvantages within his own group to structural causes. The "shared reality" bias inclines him to "align perceptions and judgments with [his] group," which leads him to underestimate the prejudice that others experience and the effects it has on them.²⁰⁵ Finally, the "just world bias" inclines him to believe that, if other groups are not succeeding at the same rate as his, it must be because they are lacking in some way because, in his mind, the system is fair.²⁰⁶

Second, rather than merely exaggerate actual differences caused by segregation, segregation also leads to the formation and reinforcement of stereotypes about other groups with no basis in fact. Stereotyping, Anderson continues, is a natural response to objects or groups that one lacks frequent contact with, and therefore lacks more

"Study: Residential Segregation Still a Problem in U.S.," *American Sociological Association News*, May 29, 2012, http://www.asanet.org/sites/default/files/savvy/documents/press/pdfs/ASR_June_2012_Kyle_Crowder_News_Release.pdf

- The percentage of schools with predominantly minority and low-income students is rising. Jacqueline Nowicki, "K-12 Education: Better Use of Information Could Help Agencies Identify Disparities and Address Racial Discrimination," *Government Accountability Office*, May 17, 2016, <http://gao.gov/products/GAO-16-345>
- "Americans' core social networks tend to be dominated by people of the same race or ethnic background. However, the degree of racial and ethnic diversity in Americans' social networks varies somewhat according to their particular race or ethnicity. Among white Americans, 91% of people comprising their social networks are also white, while five percent are identified as some other race. Among black Americans, 83% of people in their social networks are composed of people who are also black, while eight percent are white and six percent are some other race. Among Hispanic Americans, approximately two-thirds (64%) of the people who comprise their core social networks are also Hispanic, while nearly 1-in-5 (19%) are white and nine percent are some other race." Daniel Cox, Juhem-Navarro Rivera, Robert Jones, "Race, Religion, and Political Affiliation of Americans' Core Social Networks," *PRRI*, Aug 3 2016, <https://www.prr.org/research/poll-race-religion-politics-americans-social-networks/>

²⁰³ Elizabeth Anderson, *The Imperative of Integration* (Princeton: Princeton University Press, 2010) p.44.

²⁰⁴ *Ibid.*, p.47.

²⁰⁵ *Ibid.*, p.46.

²⁰⁶ Matthew Hunt, "Status, Religion, and the "Belief in a Just World: Comparing African Americans, Latinos, and Whites," *Social Science Quarterly*, 81(1), 325-343.

individualized knowledge of. These stereotypes are reinforced by the knowledge one does have, which comes mostly from the media. The media feeds into prevailing stereotypes, in part through the “illusory correlation bias,” which “disposes people to form stereotypes about a group with which they have little contact on the basis of unusual events such as sensational crimes.”²⁰⁷

All of these effects, and undoubtedly others, lead to the kind of reasoning that we saw in the welfare and immigration cases. Based on unfamiliarity and media representations, poverty and therefore welfare are associated – in many people’s minds – with African-Americans, and illegal immigration with Latinos. Because of further media narratives, as well as the illusory correlation, attribution, shared reality, and just world biases, people tend to explain the disadvantages of African-Americans and the illegal immigration of Latinos as due to their own failings, rather than the system’s. So, when considering the issues of welfare and immigration, many people will automatically call to mind groups and traits that they perceive negatively, leading them to not only reject the policy but to summon public reasons, like merit, to justify it.

Do we have to represent citizens in the idealized model with these kinds of explicit and implicit prejudices? Consider, again, the three arguments for including religious reasoning. First, we cannot reasonably expect ordinary reasoners under ordinary conditions to avoid, at least, implicit biases, and perhaps not even explicit stereotypes. Ordinary conditions include only easily accessible information, which, for people like Adam, will not include much information about other racial groups. The response of ordinary reasoners is to rely on heuristics like stereotypes to fill the gaps in their understanding, so we can expect ordinary reasoners to hold some stereotypes as well. We also cannot expect ordinary reasoners to escape automatic associations that they are repeatedly exposed to through the media, which produce implicit biases about particular groups and issues. Nor does the situation improve if we raise our expectations to representing only highly competent viewpoints, which is the second rationale for inclusion. As Perez and Gilens both point out, highly educated people are equally likely to hold implicit biases, and are equally susceptible to believing media-reinforced stereotypes, such as that most poor people are African-American.

Finally, and for many of the same reasons, these biases are an inevitable effect of liberal values and institutions. Freedom of association, combined with the natural human tendency to form in-groups and out-groups, allows us to live in communities with only in-group members. Since one of the predominant lines along which we form in-groups is race, this leads most Americans to live, work, socialize, or be educated in largely racially homogenous communities. This segregation, in turn, contributes significantly to the explicit and implicit biases described above, many of which can be traced back to important liberal values. Given that this is Adam’s epistemic starting point, it is, arguably, coherent for him to hold stereotypes and implicit biases against other racial groups.

²⁰⁷ Elizabeth Anderson, *The Imperative of Integration* (Princeton: Princeton University Press, 2010) p.46.

However, there is one further response that public reason theorists could make. They could concede that there is no descriptive feature that separates the good kind of non-rational influences from the bad kind. Instead, the difference is normative; the good kinds are the influences we should tolerate and the bad kinds are the influences we should not. Jonathan Quong might offer a response like this. He argues that public reason theories should begin by assuming the truth of basic liberal values. We should then construct the idealized model with that liberal filter in mind. So, we would only represent idealized people whose influences are consistent with basic liberal values, while excluding all other influences. That would, at least arguably, filter out at least some of the objectionable influences I have cited. It may be inconsistent with the value of equality, for instance, to be influenced by implicit biases that produce a negative association with other racial groups.

I will not address this response in much depth because I have argued against Quong's theory in a previous chapter. But, briefly, I find three problems with this kind of normative response to the problem of objectionable influences. First, we might wonder whether this normative response would address all potential objectionable influences on idealized reasoning. Perhaps implicit racial bias is inconsistent with a belief in equality, but motivated reasoning and the group identities that produce it seem consistent with liberal values. Second, as I argued in a previous chapter, idealizing people to reason solely from liberal values requires a level of idealization that is inconsistent with even the minimal amount of realism that Quong's public reason theory requires.²⁰⁸ And third, saying that some influences are in and some are out for normative reasons is an arbitrary response. If we want to know which features of real people are relevant to political justification, the answer ought to be more than: "the features we intuitively think should be relevant." After all, public reason's critics do not think any features are relevant at all. For them, political justification is about truth, not what people think.

Without any defensible way to distinguish between good and bad influences on idealized reasoning, public reason theories must represent objectionable influences in the idealized model. But, doing so seems fatal to the project. If idealized citizens are susceptible to these objectionable influences then, even if they are idealized to rely only on public reasons, the resulting norms will not be justifiable to all. A hypothetical deliberation among potentially unreasonable citizens will not tell us which norms all reasonable people could accept.

And this holds no matter what conception of "public reasons" we consider. If public reasons are shareable reasons, but those reasons are chosen based on prior motivation and

²⁰⁸ To summarize my previous argument: Quong's justification for rejecting a Kantian type of liberalism and favoring public reason liberalism is to meet a weak realism requirement. That is, he seeks to identify a kind of liberalism that would be acceptable to most citizens in an ideal, but possible, liberal society. And, he argues, a Kantian liberalism would be impossible in even an ideal liberal society because citizens would reasonably disagree about autonomy as a foundational value. However, I argued that the beliefs that Quong must attribute to ideal liberal citizens would also be impossible in any liberal society. Quong's ideal citizens must accept substantive liberal values, prioritize them over all other values, and do all this because of Quong's philosophical argument. This is not different than attributing Kantian beliefs to all ideal liberals, and therefore violates his weak realism requirement.

implicit bias, then the conclusion of shareable reasons will still be unjustified. And, if public reasons are intelligible reasons, but the connection between people's reasons and their conclusions depends on non-rational influences, then their reasoning will not be intelligible to others. In short, if public reason theories require all idealized citizens to rely on the right sort of reasons, but idealized citizens must retain certain biases, then even an idealized reasoning process confers no justification on the result.

Public reason theorists began with the problem that even reasonable citizens disagree in their political conclusions. This problem could be solved, they argued, if reasonable citizens could at least consider the reasons behind the conclusions to be justifiable to all. In a sense, theorists shifted the focus of justification from the outcome of deliberation to the reasoning process, whether real or idealized. However, it now appears that the process is also not justifiable to all – even the seemingly public reasons are compromised by objectionable influences.

A Publicly Justified System

If we take both public reason and this problem seriously, what can we do? One thing would be to simply draw the conclusion that my objection suggests: there can be no legitimate norms on issues where public reasoning is likely to be influenced in objectionable ways. But, this would likely invalidate norms on far too many issues to maintain a workable government. If partisanship were a distorting factor, for instance, then we could not have legitimate laws on many of the issues that American citizens care most about right now.

Instead, we might say that distortions in reasoning do not invalidate the whole issue as a subject for legitimate law; they merely invalidate certain positions on those issues. Maybe I, as a reasonable-but-biased person, ought to reject my opinion on any topic on which social psychologists tell me I am susceptible to bias. For instance, imagine the science of implicit bias developed to the point where researchers agreed that all opposing positions on welfare were influenced in part by negative associations with African Americans. Moreover, this objectionable influence biased people's views against increasing welfare services, but not in favor of increasing services. So, perhaps the way to counter that bias is to reject the conclusion of the biased line of reasoning and declare the anti-welfare position to be unreasonable; that is, to be the product of reasoning that is susceptible to distortion. This may be a way to justify liberal norms of group equality by ruling out the opposition from the start.

There are two problems with drawing this conclusion from the worry about biases. First, social psychologists may themselves be biased. As a group, they are likely to come from a similar background, have similar educations, and potentially similar political beliefs. So, if we accept that they have identified genuine biases, then we must also accept that their conclusions about those biases may not be fully reliable. Second, in most cases, social science and psychology are not developed enough to make sweeping conclusions of the causes of our positions on particular issues. And, it is unlikely we will be able to trace the origins of our thinking with that kind of precision in the near future.

These problems make it difficult to rule out particular positions as the product of tainted reasoning, but they do not undermine the general worry about biased reasoning. For one thing, even if we doubt the social scientists due to their biases, we are still relying on the existence of these general biases to motivate our doubts, so we are accepting their conclusions to that extent. And second, there is a lot of evidence that we have general tendencies in our reasoning on certain issues, whether that is due to motivated reasoning, in-group favoritism, or unconscious bias. We may not be able to know the degree to which a line of reasoning has been influenced, nor the precise number of people who are affected, but there is enough evidence to worry that public reasons may be so affected, and that is enough for others to demand further justification.

If we cannot rule out entire issues or particular positions as distorted, then what can we do? In the remainder of this chapter, I will sketch a couple of potential directions for public reason theories that can acknowledge these irrationalities, while potentially still justifying appealing political norms. I plan to pursue these directions in future work, but for now, I can only outline the benefits and pitfalls of each.

One possibility is to acknowledge that reasonable people cannot all accept certain lines of public reasoning, but there is something they can agree on: their own limitations. Just as public reason's justification requires representing reasonable people as somewhat irrational, it also requires representing them as epistemically modest. As I have argued, idealized citizens see the religious reasoning of others as relevant to political norms because of some quality of that reasoning – it is coherent, perhaps, or the most we can expect of ordinary citizens. But, that is only half the story. Reasonable people appreciate the quality of other people's reasoning, but they also recognize the limitations of their own. In Rawls' account, for instance, reasonable people are supposed to accept the burdens of judgment, which means they accept the fact that their own life experiences inevitably influence their reasoning in ways they cannot identify. Similarly, in Gaus' account, reasonable people recognize that they are "interpreters" of morality, not diviners, which is why they cannot impose their version of the truth on others.²⁰⁹ It is therefore consistent with the purpose of public reason theories to idealize citizens to recognize the limits of their own reasoning or, for my purposes, to understand their own biases.

On this kind of public reason model – one that emphasizes, rather than idealizes, human limitations – our most significant public reasons are not substantive values but a shared awareness of our limited and idiosyncratic epistemic positions, and the consequences they have for our reasoning. This does not mean that all reasonable people doubt all of their moral views; if that were the basis of public reason, then it would undermine its own goal of defending reasonable pluralism. It does mean, however, that reasonable people recognize that, on certain types of factual questions, their reasoning is likely to be compromised due to their epistemic environment and natural tendencies. And they also recognize that some factual misconceptions, such as stereotypes, can taint some of their moral views, such as their positions on welfare.

²⁰⁹ Gerald Gaus, *The Order of Public Reason* (Cambridge: Cambridge University Press, 2011), p.15.

While reasonable people may not directly agree on substantive values in this model, however, it may still be able to justify important political values. In what follows, I sketch a few possible ways. Whether these specific proposals are right or not, however, the general point is this: For idealized but biased citizens to justify their reasons to each other, they must propose those reasons in the context of an epistemic and deliberative system – one that will correct for and mitigate any errors due to bias. Supporting that system is how biased reasoners can justify proposing their views, even when they are aware of their possible biases. To see how such an argument might work, let us consider a few potential conclusions of this model.

I have said that reasonable people are aware of the conditions under, and issues on, which they are unreasonable.²¹⁰ If reasonable people recognize the issues on which they are unreasonable, but they still seek to justify their proposals on those issues to others, how can they do so? The answer depends on *why* they are unreasonable. As I have described above, some unreasonableness is just the product of our limited epistemic positions. Each person is just one person, living in one place, at one time, often among one group. And this leads to various limitations of evidence, idiosyncrasies of vantage point, and many variables that we cannot control for. These limitations, in turn, produce generalizations, implicit biases, and prior probabilities that make us poorly responsive to new evidence and therefore unreasonable.

We can think of this kind of unreasonableness as the kind that many scientists are increasingly worried about. For instance, John Ioannidis has argued that there are certain conditions under which medical studies are less likely to get accurate results:

...a research finding is less likely to be true when the studies conducted in a field are smaller; when effect sizes are smaller; when there is a greater number and lesser preselection of tested relationships; where there is greater flexibility in designs, definitions, outcomes, and analytical modes; when there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance.²¹¹

At least some of these conditions are analogous to those I have described above. Like individual scientists, individual citizens reason based on small sample sizes, with great flexibility in their “analyses.” These conditions, Ioannidis argues, make any one scientific finding unlikely to be true. Even if we don’t want to go that far, we can at least agree that these conditions make individual scientific findings provisional. However, this limitation

²¹⁰ That doesn’t mean they think that their beliefs on that issue are wrong; they simply recognize that there is, at least from other people’s perspective, a reason to worry that something other than public reasons is influencing their conclusion, and therefore that their conclusion is not justifiable to others. They may be perfectly content with that outside influence, as they would be, for instance, if they endorsed their own religious background. But, they still recognize that the potential for objectionable influence makes their views unjustifiable to others.

²¹¹ John Ioannidis, “Why most published research findings are false.” *PLoS Med.* 2(8):e124, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1182327/>

on individual results does not mean that a scientist should not publish her results or that other scientists should disregard them.

Instead, it means that she should do two things. First, to the extent possible, she should try to correct for the conditions that make her prone to error and place herself under better conditions: conduct larger studies, avoid questions with a financial interest, and so on. This will make her findings more justifiable to others. And second, she should offer her findings, not as the final truth, but as a contribution to a larger epistemic system that she also supports – one that is designed to correct for any errors. She should publish her data and methods, as well as her results, to allow other scientists to try to replicate her work. So, even if others believe her finding has a lower probability of being correct, it is justifiable to others as an input into a system that will increase its justifiability over time. In short, the solution to her parochial epistemic conditions is more and better science.

Public justification can work the same way. Despite recognizing the limitations of their epistemic positions, reasonable people can still offer others public reasons, even though they may be tainted by objectionable influences. To make those reasons more justifiable to others, however, reasonable citizens must, first, support conditions that improve the epistemic position in which they form those reasons. Since the problem with their epistemic positions comes from the skewed nature of the evidence they acquire, the solution is to support conditions that expose them to more and better evidence. Rather than merely the testimony of one's friends, neighbors and colleagues, then, reasonable people will seek out the testimony of people unlike them, who have very different views. In other words, reasonable people can improve the justifiability of their reasons by supporting broader and higher quality democratic deliberation.

For instance, if the problem with one's epistemic position is that one's prior probabilities about the existence of climate change are shaped only by the testimony of Democrats, then testimony from non-Democrats may update those probabilities so that they are less skewed. This is not to say that one will therefore reject the existence of man-made climate change, but merely that deliberation can make one more reasonable – better able to respond to evidence on either side. This will decrease the chances that one's empirical – and, by extension, certain political – beliefs will be influenced in objectionable ways, which therefore makes one's views more justifiable to others.

Second, on issues on which reasonable people may be unreasonable, they can further justify their views by offering their proposals, not as the final truth, but as a contribution to an ongoing deliberative process that can correct for any errors. Science corrects for errors by narrowing the variables under consideration and replicating the experimental process. Similarly, reasonable people can potentially improve the justifiability of their proposals by narrowing their scope and impact, seeing what worked, and then revisiting the decision. This suggests a politics that values incremental change and federalism, at least on partly empirical questions, to encourage experimentation and replication, along with features like sunset clauses to force reevaluation.

However, I have also described a second source of unreasonableness – one that cannot be solved by more and better deliberation. Cass Sunstein points out this second source when he tries to explain the results of his survey on climate change. In that survey, you will recall, liberals tended to discount evidence against man-made climate change, while responding to evidence in favor of it, while conservatives did the reverse. One explanation for this, as I mentioned, is that people’s prior beliefs about climate change were so extreme – based on the testimony from their own group – that they found the non-existence of man-made climate change to be too preposterous to take seriously. The solution to that, I argued, was providing such people with more and broader evidence.

But, Sunstein points out a second possible explanation for their unreasonableness – identity politics – that demands a different response from reasonable people. In discounting the evidence against man-made climate change, perhaps they were simply supporting Team Liberal over Team Conservative. This kind of unreasonableness cannot be solved by more deliberation because, rather than responding to the new testimony, people motivated solely by identity will either ignore it or explain it away. The same can be true of people under the influence of implicit bias, as we have seen. Implicit bias can cause people to discount the testimony of groups they are implicitly biased against, so being exposed to more diverse testimony may not make people more responsive to evidence.²¹²

Reasonable people must therefore account for the possibility that deliberation will not correct their biases, and may even make them worse. So, how can a reasonable person counter the possibility of an objectionable *and* uncorrectable bias? By supporting a system that counters that bias. For example, we may not know precisely how in-group bias affects our political proposals, but we can say that there is a general tendency to favor one’s own group. And the best system we have for combating group favoritism, whether one’s own or the majority’s, is a liberal, or at least a multicultural, one. Allowing individuals or groups freedom from state interference addresses the worry that reasonable

²¹² As Jennifer Saul argues, implicit bias can affect our views, not only through implicit stereotypes, but through the credence we give (or deny) to testimonial evidence from different groups: “The problem starts to become vivid when we ask ourselves *when* we should be worried about implicit bias influencing our judgments. The answer is that we should be worried about it whenever we consider a claim, an argument, a suggestion, a question, etc. from a person whose apparent social group we’re in a position to recognize. Whenever that’s the case, there will be room for our unconscious biases to perniciously affect us. Most discussed in the literature so far (see Fricker 2007), we might make a mistaken judgment of credibility when assessing testimony. But we also might fail to listen properly to a contribution; fail to carefully consider a question; judge an argument to be less compelling or original than it is; think the evidence presented is worse than it is... And *this* is going to happen a great deal. It happens whenever we are dealing with the social world in a non-anonymised manner. Since the world is only rarely anonymised for us, this will happen nearly all the time. Much of our knowledge comes from testimony, or from arguments or evidence that we are presented with. Those testifying, or presenting the arguments or evidence, are usually people. And people are generally (though not always) perceived by us as members of social groups. Moreover, much of the knowledge we already have has come to us in this way. Our acceptance or rejection of testimony, arguments, evidence and the like has shaped the worldviews we have now. And this acceptance or rejection was, we can be fairly certain, distorted by the perceived social groups of those presenting the testimony, arguments or evidence. Worse yet, we cannot even go back and attempt to consider or correct errors that we might have made — we are very unlikely to remember the sources of these beliefs of ours.” Jennifer Saul, “Scepticism and Implicit Bias,” *Disputatio*, 5(37), pp.250-251.

people tend to unreasonably favor their own or the majority's group, at the expense of others. It also addresses the worry that we do not always know which biases are correctable by deliberation and which are not. As J.S. Mill argued, preserving a diversity of beliefs and lifestyles is the best way to both facilitate future deliberation *and* guard against error.

Ultimately, the public reason project is an attempt to show that liberalism is internally consistent – that advocating tolerance for others is not itself based on an intolerant philosophy. Public reason theories therefore defend a slightly weaker liberalism that they argue is justifiable to all reasonable citizens, and therefore tolerant of their different beliefs. The challenge has always been to understand how such diverse people end up agreeing on substantive liberal values. I would suggest that this version of the public reason model is a more promising route to defending liberalism. Rather than defending liberal rights as the shared reasons of diverse but reasonable citizens, we may be able to defend them as the only solution to shared unreasonableness. Here, liberal rights are the only proposals that reasonable people can justify to each other, when their actual views may be unreasonable.

There is one further and important objection that this kind of solution must address. If I am arguing that even an idealized model of citizen reasoning must represent our biases, then shouldn't the model also represent our well-known bias to discount our own biases?²¹³ That is, one of our strongest biases is to think that we have the truth, while others are biased. If we must represent reasonable people with this particular bias, then reasonable people would be, for public reason's purposes, doubly useless. They would be too biased for their current views to be justifiable to others, which, I have argued, is the problem with current public reason theories. But, they would also be too unaware of their own biases – too epistemically immodest – to compensate for their errors, which would negate this potential solution.

The force of this objection depends on offering a definitive account of the kinds of biases that belong in the idealized model, rather than the survey that I have presented here. By some rationales for including biases in the idealized model, the bias to ignore one's own biases appears to belong. For example, it does seem typical of an ordinary reasoner under ordinary constraints to be unaware of his own biases. On the other hand, other rationales for inclusion would seem to exclude the bias to ignore one's biases. I have suggested, for instance, that highly competent reasoners such as scientists are able to recognize their own biases to the extent that they offer only provisional hypotheses into a system designed to correct their errors, which is the extent of the humility I am attributing to reasonable people.

However, even if we must represent the reasonable person as too biased to see his own biases, there may still be good reason to accept the systemic counter-measures I have just described. This is because, while he may not recognize his own biases, he is perfectly capable of seeing others'. And, in recognizing others' biases, the reasonable person

²¹³ Nathan Ballantyne, "Debunking Biased Thinkers," *Journal of the American Philosophical Association* 1(1) pp.141--162.

knows that, due to their biases, others see him as biased. If others see him as biased, then they do not see his proposals as publicly justified. So, just as before, the only way for others to see his proposals as justified is to counter and mitigate their biases, to whatever extent possible.

A full defense of this approach would doubtless face many more objections, which I hope to consider in future work. However, given the challenges of the tradition public reason view, an alternative that can both acknowledge realistic qualities and still potentially justify important norms seems well worth exploring. If public justification is required for legitimacy, and made difficult by the qualities of reasonable people, then the preconditions for public justification are conditions that mitigate those qualities. The norms justified by public justification, then, are the institutions and norms that make public justification possible. It is not all public reason liberals could hope for, but perhaps it could serve as a more modest ideal for a more modest people.