Essays on Neutrality, Perfectionism, and Public Reason Liberalism

Collis Tahzib

St John's College, University of Oxford

Thesis submitted in partial fulfilment of the requirements of the degree of Doctor of Philosophy in Philosophy

October 2019

73,500 words

Abstract

The six essays in this thesis all focus on the neutrality-perfectionism debate within contemporary liberal political philosophy. In the first paper, "Pluralist Neutrality", I develop a novel account of the principle of state neutrality, according to which neutrality is a pluralist notion with consequential, justificatory, and intentional dimensions. In the second paper, "Deontic Perfectionism", I distinguish between the claim that the state *may* promote the good life and the claim that the state *must* promote the good life, and I argue in favour of the latter, stronger thesis. In the third paper, "Does Edificatory Perfectionism Express a Quidnunc Mentality?", I defend perfectionism against an intriguing and highly original objection that has recently been developed by Matthew Kramer, according to which perfectionist regimes that seek the edification of the citizenry are guilty of the mentality of a village busybody who meddles in matters that are none of her business. The fourth paper, "Perfectionism: Political not Metaphysical", explores the idea of a "political perfectionism" — a view which combines the contractualist commitment to public justification favoured by political liberals with the perfectionist claim that the state may promote the good life. In the fifth paper, "Do the Reactive Attitudes Justify Public Reason?", I argue against Gerald Gaus's attempt to ground public reason in the kind of everyday reactive attitudes of resentment and indignation that seem impossible for us to renounce. In the sixth paper, "Does Social Trust Justify the Public Justification Principle?", I criticize Kevin Vallier's recent argument according to which adherence to the public justification principle is justified on the grounds that such adherence is necessary for sustaining a system of social trust within diverse and large-scale societies.

Acknowledgements

I am deeply indebted to my supervisors, Jeff McMahan and David Miller. Jeff and David have been models of kindness, encouragement and insight; their penetrating and illuminating comments have greatly improved the essays in this thesis.

I am also most grateful to the following individuals for helpful conversations or written feedback on various parts of this thesis: Max Afnan, Paul Billingham, Matthias Brinkmann, Samuel Bruce, Simon-Pierre Chevarie-Cossette, Henrik Dahlquist, Sarah-Lea Effert, Matthew Kramer, CM Lim, Andrew Lister, Simon Cabulea May, Yuuki Ohta, Fergus Peace, Jonathan Quong, Tom Sinclair, Anthony Taylor, and Steven Wall. For very helpful comments, I owe thanks to anonymous reviewers from the *European Journal of Political Theory,* the *Journal of Political Philosophy*, and *Philosophy & Public Affairs*. I also thank participants at various workshops and conferences for their feedback: the Nuffield Political Theory Workshop at which I presented Chapter 4 in October 2018; the conference in Oxford on Themes from the Political Philosophy of Gerald Gaus at which I presented Chapter 5 in March 2019; the Workshop in Honour of Matthew Kramer in Cambridge at which I presented Chapter 3 in July 2019; and the 7th Annual Workshop for Oxford Studies in Political Philosophy, at Syracuse University, at which I presented Chapter 2 in August 2019. Acknowledgements of particular points are also made in footnotes.

I am grateful for the enormously generous support of the Clarendon Fund and St John's College. I thank Alison Hills, my college advisor, for many enlightening conversations.

Finally, I thank my parents for their constant love and support, and Raven for her affection and encouragement. I dedicate this thesis to them, with love.

Contents

Introduction

The six essays in this thesis all focus on the neutrality-perfectionism debate within contemporary liberal theory. This debate is about one of the most important questions in political philosophy: namely, what the state can permissibly do, and in particular whether the state can permissibly seek to promote particular ideals of the good or flourishing human life.

According to proponents of liberal neutrality, the state should remain, as far as possible, neutral among the competing conceptions of the good life. The state, that is, should not seek to impose or even promote any particular ethical or religious doctrine. Rather, the state should maintain a fair framework of laws and institutions within which each citizen can freely pursue her own conception of the good or flourishing life. By contrast, perfectionists hold that the state may legitimately seek to promote valuable, meaningful or worthwhile ways of life and to inhibit the pursuit of degrading, worthless or hollow ways of life. For perfectionists, then, the state "should act with discrimination to encourage the good and the valuable and to discourage the worthless and the bad".[1]

There are various arguments in favour of both neutrality and perfectionism, but one of the most important sources of resistance to perfectionism and of support for neutrality has come from public reason liberalism. Public reason liberals are struck by what Rawls calls "the fact of reasonable pluralism" — that is, by the fact that, in free societies, citizens exercising in good faith their powers of reason will arrive at highly divergent yet reasonable views on a range of important matters.[2] In light of this, public reason liberals tend to argue that the liberal state should refrain from appealing to controversial perfectionist judgements about what does and does not constitute a flourishing life and should instead justify state action in terms of values and reasons that are in some sense acceptable to all reasonable citizens. Thus, the debate between neutralists and perfectionists connects in important ways with the debate between public reason liberals and their critics.

---

[1] Raz, J., "Liberalism, Skepticism, and Democracy", *Iowa Law Review* 74 (1989), p. 785.
[2] Rawls, J., *Political Liberalism* (New York, NY: Columbia University Press, 1996), *passim*.

This topic is significant both practically and theoretically. On the practical side, whether perfectionism or neutrality is true is generally thought to have important implications for the legitimacy of a range of concrete policies such as the level of state support for the arts, "sin taxes" on gambling and drugs, and other such measures and interventions intended to refine, elevate and edify aspects of individual and collective life.

On the theoretical side, it has often been said that the neutrality principle and the public justification principle are the definitive principles of liberal political philosophy and that these principles carry to a higher order of generality earlier liberal principles such as the separation of church and state. For Dworkin, neutrality between competing conceptions of the good life is the "nerve of liberalism", while for Waldron state neutrality has a "rich heritage" within the liberal tradition, stretching back through John Stuart Mill's *On Liberty* to John Locke's *Letter Concerning Toleration*.[3] The public justification principle has received similarly high praise, with Macedo calling it the "moral lodestar of liberalism", and Ackerman referring to it as "the organizing principle of liberal thought".[4] The principles that I examine in this thesis thus occupy a cherished position within liberal theory, and so the fate of these principles may well have broader implications for the tenability of liberalism as a theory of legitimate state power.

*

The first three essays of this thesis focus more explicitly on the debate over the neutrality principle and the final three essays of the thesis focus more explicitly on the debate over the public justification principle — though, as noted above, there are important connections between these two debates.

The debate about state neutrality sometimes suffers from a lack of conceptual clarity. For instance, critics of neutrality often object to liberal neutrality on the grounds that neutrality is an illusion and that any political arrangement is

[3] Dworkin, R., "Liberalism" in *A Matter of Principle* (Oxford: Clarendon Press, 1985), p. 183; Waldron, J., "Legislation and Moral Neutrality" in his *Liberal Rights: Collected Papers 1981 – 1991* (Cambridge: Cambridge University Press, 1993), p. 143.
[4] Macedo, S., *Liberal Virtues* (Oxford: Clarendon Press, 1990), p. 78; Ackerman, B., *Social Justice and the Liberal State*, p. 10.

bound to advantage some ways of life over others — to which neutralists often reply that they only ever intended to defend neutrality of *justifications*, not neutrality of *consequences*. This leads to a situation in which, as Arneson puts it, "what some authors defend under the heading of neutrality is not the same as what others attack".[5] The first paper, "Pluralist Neutrality", seeks to integrate these discussions and to address these conceptual unclarities by articulating a novel account of the principle of state neutrality. On this view, neutrality is a pluralist notion that has consequential, justificatory, and intentional dimensions. Though I do not endorse neutrality as an ideal (as the second and fourth and, to some extent, third papers make clear, I side with perfectionists), I take this to be the most plausible and sympathetic analysis of the concept of neutrality — an analysis that gives neutrality its best hearing and that could contribute towards curbing the tendency of defenders and opponents of liberal neutrality simply to talk past each other.

In the second paper, "Deontic Perfectionism", I point out a further unclarity in the neutrality-perfectionism debate — namely, that perfectionists are not always clear whether they hold that the state *may* promote the good life or that the state *must* promote the good life. I argue in favour of the latter, stronger thesis, according to which states have perfectionist duties. I then show how this deontic form of perfectionism has the resources to block a number of powerful recent objections to perfectionism, and I defend deontic perfectionism against the concern that it violates the public justification principle.

In the third paper, "Does Edificatory Perfectionism Express a Quidnunc Mentality?", which is more narrowly focused than the previous two, I defend perfectionism against an intriguing and highly original objection that has recently been developed by Matthew Kramer. This objection is that perfectionist regimes that seek the edification of the citizenry are guilty of a "quidnunc" mentality, one akin in important respects to that of a village busybody who distributes "boxes of delicious homemade fudge" to "inspire her fellow inhabitants to lead lives of propriety" and who in general meddles in matters that are "none of [her] business".[6] I argue that this objection begs the

[5] Arneson, R., "Liberal Neutrality on the Good: An Autopsy", in S. Wall and G. Klosko (eds), *Perfectionism and Neutrality* (Lanham: Rowman & Littlefield, 2003), p. 192.
[6] Kramer, M., *Liberalism with Excellence* (Oxford: Oxford University Press, 2017), pp. 251-96, at p. 282, 287.

question against those perfectionists (such as Raz, Wall and myself) who believe that states have duties of edification.

The fourth paper, "Perfectionism: Political not Metaphysical", explores the idea of a "political perfectionism" — a greatly neglected position in the conceptual space of these debates. Political perfectionism combines the contractualist commitment to public justification favoured by political liberals with the perfectionist claim that the state may promote the good life. By taking a "political turn" of this kind, perfectionists can incorporate many of the central insights of political liberalism — that political philosophers must show greater sensitivity to reasonable pluralism, and that the terms of our common political life should as far as possible be justifiable to all in order to realize a valuable form of civic friendship amongst the citizenry — whilst also avoiding its rather less appealing requirement that we "cordon off political morality from our best understanding of human flourishing".[7]

One question that has, I think, received insufficient attention relative to its significance for public reason liberalism is that of why we should care about justifiability to all reasonable citizens. That is: Why reason publicly? Why make politics hostage to what a specific set of idealized citizens would or would not accept? Why not simply implement the political rules and institutions that are *correct* — or at least that are most likely to be correct — regardless of whether those rules and institutions could be reasonably rejected by some citizens? The fifth and sixth papers of my thesis criticize two recent attempts, by Gerald Gaus and Kevin Vallier respectively, to answer these questions about the foundations of public reason liberalism. But, to be clear, while I reject Gaus's and Vallier's particular arguments for the public justification principle, I do support that principle on other grounds (namely, on grounds of civic friendship), as the fourth paper makes clear.

In "Do the Reactive Attitudes Justify Public Reason?", then, I argue against Gaus's recent attempt to ground public reason in the kind of everyday reactive attitudes of resentment and indignation that seem impossible for us to renounce. I object to Gaus's argument at three levels: (1) the reactive attitudes

[7] Wall, S., *Liberalism, Perfectionism and Restraint* (Cambridge: Cambridge University Press, 1998), p. 123.

do not presuppose public reason because we often seem to appropriately resent individuals for violating moral rules that they have no sufficient reason to endorse; (2) even if the reactive attitudes do presuppose public reason, these attitudes can in fact be renounced without great social cost, as Pereboom and others have recently argued; (3) even if the reactive attitudes do presuppose public reason, and even if the reactive attitudes cannot be renounced, this merely explains, rather than justifies, public reason because the reactive attitudes of resentment and indignation lack the right kind of normative status to tell us why public justification matters.

The sixth and final paper, "Does Social Trust Justify the Public Justification Principle?", also engages with the question of what grounds public reason liberalism. Recently, Vallier has argued that sustaining a system of social trust within diverse and large-scale societies requires adherence to the public justification principle. I argue that Vallier's defence of public justification does not succeed because there are alternative conceptions of democratic discourse and decision-making that lack a public justification principle yet that could still sustain social trust. I then defend this objection against two potential responses: that alternatives to public justification would produce *less* social trust, and that such alternatives would not produce social trust *in the right way*.

*

In conclusion, this thesis makes both critical and constructive contributions to the literature. On the critical side, I rebut three important recent arguments of Kramer, Gaus and Vallier (papers 3, 5 and 6). On the constructive side, I propose a novel analysis of neutrality (paper 1) and develop a view that, whilst being perfectionist, is intended to incorporate various insights of neutralists and public reason liberals (papers 2 and 4). More precisely, I defend a moderate form of perfectionism, one that is sensitive to the fact of reasonable pluralism and that satisfies a version of the public justification requirement. Though my arguments are pitched at a fairly abstract level, and address questions of political morality rather than specific matters of public policy, I believe this moderately perfectionist view would have important implications in a range of areas, as I suggest in the second and fourth papers.

Pluralist Neutrality

It is a central tenet of much contemporary liberal political philosophy that the state should be, in some sense, neutral among the competing worldviews and conceptions of the good life that exist within society. A neutral state, it is often said, is one that does not impose or even promote any particular ethical, religious or metaphysical doctrine.

A contrast is often drawn between liberal neutrality and perfectionism. According to perfectionism, the role of the state is to encourage citizens to pursue valuable, meaningful or worthwhile ways of life and to discourage citizens from pursuing degrading or hollow ways of life. By contrast, according to proponents of liberal neutrality, the state should not attempt either to promote or to inhibit any particular way of life that it judges to be better or worse. Instead, the state should simply uphold a fair and neutral framework of laws, rights, opportunities and institutions within which individuals can frame and pursue their own freely chosen conception of the good life.

But it turns out that it is actually quite difficult to give this notion of neutrality more precise expression. It is quite difficult, that is, to specify the sense in which states are supposed to be neutral among different conceptions of the good life. By far most commonly, the ideal of state neutrality has been understood in terms of justifications. This tendency to reduce neutrality to some form of public justification is evident in the current popularity of public reason liberalism as well as in the slight shift in the recent literature away from analyses of neutrality and towards analyses of different varieties of public justification principle — their structure, scope, content, constituency, and moral foundations.

In this paper, I argue in favour of a pluralist account of the principle of state neutrality. On this view, neutrality is a complex, multifaceted, gradable concept with consequential, justificatory, and intentional dimensions. The upshot of this is that the justificatory focus of contemporary liberal theory is one-sided and constraining and, while it captures the justificatory and reason-giving aspect of neutrality, it ignores other important aspects of the liberal aspiration to neutrality.

To make this case, I first evaluate some of the main accounts of neutrality that have been given in the literature to date: neutrality of consequences (Section 1), neutrality of justification (Section 2), neutrality of intention (Section 3), and neutrality as equality of opportunity (Section 4), discussing, along the way, some recent proposals by Patten and Kramer. I argue that none of these accounts is adequate. Then, in light of the problems facing existing accounts, I spell out in Section 5 my own pluralist account of neutrality. In Section 6, I close by defending my account against four objections and by flagging up some further lines of inquiry.

1 Neutrality of consequences

Perhaps the most natural way of understanding neutrality is in terms of consequences (also known as "neutrality of effect"). On this view, for a state to be neutral its policies and actions must not have the effect of promoting certain activities, ideals, ways of life or conceptions of the good (for example by making them more successful, or more popular, or more easily realizable by citizens). For instance, a government that provided subsidies, exemptions and other special privileges to Christianity, but not to other religions or worldviews, would thereby be flouting neutrality of consequences because such action would have the effect of making Christianity do much better than other religions and worldviews.

However, except for a tiny minority, almost everyone in this literature roundly rejects consequential neutrality on the grounds that it is too strong and makes it impossibly difficult to attain neutrality.[1] Typically, neutrality of consequences is dismissed in one swift argument: (i) neutrality of consequences implies that if a state's laws, institutions and policies have the effect of advantaging one conception of the good over another, then that state violates neutrality; but (ii) *any* law, institution or policy is bound to have differential effects with respect to the range of conceptions of the good that citizens hold; so (iii) neutrality of

[1] Four exceptions are: Goodin, R. and Reeve, A., "Do Neutral Institutions Add Up to a Neutral State?" in R. Goodin and A. Reeve (eds), *Liberal Neutrality* (London: Routledge, 1989), pp. 193-210; Wilkins, B., "A Third Principle of Justice", *The Journal of Ethics* 1 (1997), pp. 355-74; Clarke, S., "Consequential Neutrality Revivified", in R. Merrill and D. Weinstock (eds), *Political Neutrality: A Re-evaluation* (New York: Palgrave Macmillan, 2014), pp. 109-121.

consequences is an impossibly strong conception of neutrality.[2] Here, for instance, is Rawls's rendition of that argument:

> It is surely impossible for [state action] not to have important effects and influences as to which comprehensive doctrines endure and gain adherents over time; and it is futile to try to counteract these effects and influences, or even to ascertain for political purposes how deep and pervasive they are. We must accept the facts of commonsense political sociology…Neutrality of effect or influence political liberalism abandons as impracticable.[3]

As I shall argue in the following sections, I believe that this dismissal of neutrality of consequences is a mistake. For while a strict equality-of-outcome understanding of consequences is indeed impossibly strong, there is, I think, no need to throw the consequential baby out with the strict egalitarian bathwater. Neutrality of consequences seems to me to capture something important about neutrality; and unless neutrality is taken to have at least some consequential aspect, we have no grounds for faulting the neutrality of a state whose policies and laws have the cumulative effect of making a society's background culture and public life highly aligned with one specific comprehensive worldview.

Alongside the worry about the impracticability of consequential neutrality, some have also wondered whether there is any coherent way of specifying the baseline that we use for deciding whether the state has done something that has the effect of promoting one conception of the good over another.[4] What is the baseline from which departures are deemed non-neutral? Suppose the baseline is equality. But then consider the UK, in which 60% of the population are Christian, 25% atheist, 5% Muslim, 2% Hindu and so on. A baseline of equality would seem to have the odd implication that the state ought to aim to

---

[2] For theorists who have made this rather curt argument, see, e.g., Galston, W., *Liberal Purposes: Goods, Virtues, and Diversity in the Liberal State* (Cambridge: Cambridge University Press, 1991), p. 100; Sher, G., *Beyond Neutrality: Perfectionism and Politics* (Cambridge: Cambridge University Press, 1997), pp. 3-4; de Marneffe, P., "The Slipperiness of Neutrality", *Social Theory and Practice* 32 (2006), p. 17; Colburn, B., *Autonomy and Liberalism* (Oxford: Routledge, 2010), p. 43; Merrill, R., "Introduction" in R. Merrill and D. Weinstock, *Political Neutrality*, p. 3.

[3] Rawls, J., *Political Liberalism* (New York, NY: Columbia University Press, 1996), pp. 193-4.

[4] See, e.g., Raz, J., *The Morality of Freedom* (Oxford: Oxford University Press, 1986), p. 121ff. I thank a reviewer for pressing me on this point, and for some of the formulations in this paragraph.

equalize the number of adherents of each of the different faiths.[5] Other normative baselines such as equality of opportunity or fair opportunity for self-determination also seem problematic because they threaten to make the concept of neutrality itself redundant by having whatever value is referenced in the baseline do all the moral work. Perhaps instead proponents of consequential neutrality should adopt an empirical baseline and say that the state ought to enact policies that preserve the current proportion of adherents of each conception of the good. But the problem with this move is that the current patterns of adherence may have been shaped by clearly non-neutral policies in the past and so this would arbitrarily privilege the status quo.

2 Neutrality of justification

In an effort to "accept the facts of commonsense political sociology", and in response to the alleged incoherence of consequential neutrality, the majority of liberals now understand neutrality in justificatory terms.[6] Indeed, I take the current popularity of public reason liberalism as evidence that neutrality of justification has become the dominant conception of neutrality.

According to neutrality of justification, states may not base their policies on reasons and considerations drawn from some specific conception of the good or from some controversial comprehensive religious or ethical doctrine. The state cannot, for instance, justify civil law by reference to religious scripture. Rather, states must base their public policies on values and considerations that are in some sense publicly intelligible or accessible or shareable. In Rawls's rendition of this idea, there is a duty to "appeal only to presently accepted general beliefs and forms of reasoning found in common sense, and the methods and conclusions of science when these are not controversial", as well as to

[5] More precisely, taking a certain sort of *numerical* equality as the baseline has this implication. Perhaps other more plausible "currencies" of equality, such as resource equality, can serve as the baseline. In this sense I think the baseline problem is connected to the currency problem. I return to this point when defending my pluralist account in Section 6.

[6] See, e.g., Larmore, C., *Patterns of Moral Complexity*, p. 44; Dworkin, R., *A Matter of Principle*, p. 191; Rawls, J., *Political Liberalism*, p. 137; Barry, B., *Justice as Impartiality* (Oxford: Clarendon Press, 1995), p. 160; Gaus, G., "Liberal Neutrality: A Compelling and Radical Principle" in S. Wall and G. Klosko (eds), *Perfectionism and Neutrality* (Lanham: Rowman and Littlefield, 2003), p. 153; Quong, J., *Liberalism without Perfection* (Oxford: Oxford University Press, 2011), p. 18; Vallier, K., *Liberal Politics and Public Faith: Beyond Separation* (Oxford: Routledge, 2014), p. 46.

"fundamental political ideas viewed as implicit in the public political culture of a democratic society" — ideas such as freedom, equality and fairness.[7]

Neutrality of justification will subdivide into at least two separate views depending on how "the justification" of a law, policy or institution is understood. The first view — call it neutrality of *publicly stated* justification — focuses on those justifications that were actually advanced in public, either in the wording of the statute that ultimately gets enacted, or in the justification given by legislators in debate among themselves or in addresses to the citizens. This understanding of neutrality of justification has proven especially congenial to political liberals with deliberative-democratic leanings, such as Rawls himself. Of course, questions still remain here. What if various different reasons were given in public by separate legislators in support of the law? Must they all be neutral in order to satisfy justificatory neutrality? Must only one? But it should at least be clear that the focus here is on publicity. The second view — neutrality of *objective* justification — holds that all that matters for neutrality is that there in fact exists some good neutral justification for a law or policy that one *could* provide, even if that neutral justification was never actually advanced in public. And these two versions of neutrality of justification will subdivide once again depending on how heavily one idealizes the justificatory constituency and on whether the *same* neutral justification must be *shared* by all citizens (as consensus theorists hold) or whether (as convergence theorists hold) a law may be justified to different citizens on the basis of different reasons.[8]

However, an important problem with justificatory neutrality is that it is too weak in the sense that it judges to be neutral circumstances that are intuitively non-neutral. To see this, consider Kramer's case of Convivia.[9] In Convivia, the state establishes Buddhism as its national religion and officially bestows upon it special privileges. This seems non-neutral: the establishment of a state religion is a paradigm case of non-neutrality. And indeed, if the state justifies this policy

[7] Rawls, J., *Political Liberalism*, pp. 223-4.
[8] For good discussion, see Vallier, K., *Liberalism and Public Faith*, chapters 4 & 5.
[9] Kramer, M., *Liberalism with Excellence* (Oxford: Oxford University Press, 2017), pp. 23-4. This kind of "non-neutral policy for neutral reasons" argument against justificatory neutrality has also been made by Raz in *The Morality of Freedom,* p. 116; by Arneson in "Liberal Neutrality on the Good: An Autopsy" in Wall and Klosko (eds), *Perfectionism and Neutrality*, p. 194; by Patten in "Liberal Neutrality: A Reinterpretation and Defense", *The Journal of Political Philosophy* 20 (2012), p. 255; and by Clarke in "Consequential Neutrality Revivified", p. 111.

of establishment by saying that Buddhism is the true religion, or that Buddhism is superior to other religions, then justificatory neutrality would have no difficulty in judging this to be a case of non-neutrality. Suppose, however, that the state publicly justifies the establishment policy on the grounds that the spread of the Buddhist creed will enhance domestic peace and public-spiritedness and suppose that in the circumstances these considerations do in fact generate an all-things-considered neutral justification for religious establishment. Then, since this justification does not invoke any particular conception of the good — since the policy is grounded in the neutral value of civic peace and public-spiritedness — justificatory neutrality delivers the odd verdict that the establishment of Buddhism as state religion does not violate neutrality.[10]

This is of course a rather contrived example, but there are many more realistic ways of making the same point. For instance, a number of theorists have argued that a liberal society defined by principles of neutrality and public reason can actually over time have the effect of creating a highly secular public culture that is inhospitable to religion and that forces religious citizens to divorce themselves in public life from their most powerful and most deeply held convictions and thus to "become, in and for politics, people we cannot recognize as ourselves".[11]

---

[10] A reviewer has suggested that perhaps the Convivia case really is neutral since the policy expresses a preference for Buddhism qua effective means for promoting public-spiritedness but not qua comprehensive conception of the good. Consider an analogy. Alf is locked in a room with only a banana and a pear. Because Alf is in desperate need of potassium, he chooses to eat the banana. In doing so, he does not express a preference for bananas over pears: he only expresses a preference for potassium over non-potassium. However, the problem with this argument is that it implicitly defines non-neutrality as a preference for a comprehensive conception of the good (or a fruit) *on the grounds that it is* a comprehensive conception of the good (fruit). The "qua" condition thus smuggles in justificatory neutrality. But this just defines consequential neutrality out of existence. If we keep the spirit of this preference-based definition of neutrality but excise its tendentious qua-condition — if, that is, we understand neutrality in terms of a preference for or against some conception of the good — then it becomes clear that there is a (consequential) sense in which the policy expresses a preference for Buddhism, thus rendering it non-neutral in that respect, while there is also a (justificatory and intentional) sense in which the policy does not express a preference for Buddhism, thus rendering it neutral in those respects. Similarly, when Alf chooses the banana, there is clearly a sense in which he does express a preference for the banana, even if that preference is grounded in its being a potassium source. Careful consideration of these cases, then, shows neutrality to be a complex notion with distinct internal dimensions and thus actually vindicates the pluralist account that I develop in Section 5.

[11] Dworkin, R., "Foundations of Liberal Equality", in S. Darwall (ed.), *Equal Freedom: Selected Tanner Lectures on Human Values* (Ann Arbor: University of Michigan Press, 1995), p. 202.

Another real-life example of non-neutral policy for neutral reasons would be the policies of nation-building pursued by many Western states.[12] These policies — which include the promotion of an official state language to be used in government employment, courts, legislatures, health services and street signs; decisions regarding core curricula in education; the specification of the duties and responsibilities of citizenship; the adoption of national holidays and symbols; the renaming of streets, towns, rivers and hills in order to memorialize heroes or events[13] — are often adopted on neutral grounds (such as civic peace and social cohesion) as opposed to perfectionist grounds (such as a totalizing and romantic nationalism). Yet such policies are hardly neutral in that they actively promote a particular common culture and national way of life. Of course, the "culture" that these nation-building policies promote is not as ethnographically thick as, say, Amish culture; but, as Kymlicka has argued, it is still "far from trivial" and thick enough that attempts to integrate individuals into this common culture have been met with "serious resistance".[14]

The point, in other words, should by now be clear: neutral reasons and justifications do not necessarily add up to a neutral politics in the broader sense. So while consequential neutrality is too strong in the sense that by its lights too few systems would qualify as neutral, justificatory neutrality is too weak in the sense that by its lights too many systems would qualify as neutral.

3 Neutrality of intention

According to neutrality of intention (or "neutrality of aim"), the state should not act with the *aim or intention* of promoting one conception of the good life over others.[15] On this view, neutrality is not a constraint on consequences (as in

[12] For a very interesting treatment of this, see Kymlicka, W., "Western Political Theory and Ethnic Relations in Eastern Europe" in W. Kymlicka and M. Opalski (eds), *Can Liberal Pluralism be Exported?* (Oxford: Oxford University Press, 2001), pp. 13-105.
[13] This list comes from ibid., p. 19.
[14] Ibid.
[15] See, e.g., Brighouse, H., "Neutrality, Publicity, and State Funding of the Arts", *Philosophy & Public Affairs* 24 (1995), pp. 35-63; Kramer, M., *Liberalism with Excellence*, p. 18ff, 78 fn. 15. Authors sometimes run together neutrality of intention and neutrality of justification by using ambiguous language about the "reasons" on the basis of which the state may or may not act. (See, e.g., Waldron, J., *Liberal Rights: Collected Papers 1981 – 1991* (Cambridge: Cambridge University Press, 1993), p. 150.) I take it that intentional neutrality concerns "motivating reasons" whereas justificatory neutrality concerns "normative reasons". For this distinction see, e.g., Dancy, J., *Practical Reality* (Oxford: Clarendon Press, 2000)

consequential neutrality), nor a constraint on reasons and justifications (as in justificatory neutrality), but rather a constraint on aims or intentions.

Some immediate and obvious objections that have been leveled against intentional neutrality can be quickly dispatched. For instance, some have raised metaphysical worries about whether "the state" is even the sort of thing that can have an intention.[16] Others have suggested that, even if we could explain the intentions of the state in terms of individual intentions, it would be overly "inward-looking" to think that the legitimacy and neutrality of laws and policies should depend on the intentions of government officials. As Sher has made this point, how, in the case of well-entrenched law, could the legitimacy of that law or constitution turn on the "mental processes of long-dead persons"?[17] And, as he also asks, wouldn't neutrality of intention have the absurd implication that we ought to repeal laws that were enacted by legislators on the basis of non-neutral intentions, and then re-enact those laws with minds cleansed of all impure perfectionist thoughts?[18]

As Kramer has recently argued, however, these objections stem from an overly psychologistic construal of state intentions. The task of imputing an intention to a law or policy is "not an exercise in mind-reading"; it does not involve peering into the souls of legislators and government officials, or trying to divine the mental processes of long-dead persons.[19] Rather, argues Kramer, in line with the jurisprudential literature on legal interpretation and legislative intent, the attribution of intention to a law involves trying, in light of a range of considerations, to make sense of a law and the objectives that most plausibly account for its having been adopted and sustained.[20] The relevant considerations will include the terms of the law itself, the various public statements made about the law by those officials responsible for its enactment or sustainment, and the general social context within which such statements are made. Of course, this approach is not without its own complications — for instance, it involves difficult judgements about the sincerity of public statements, and in cases in which more than one intention can reasonably be

[16] Merrill, R., "Introduction", p. 4.
[17] Sher, G., *Beyond Neutrality*, p. 24
[18] Ibid.
[19] Kramer, M., *Liberalism with Excellence*, p. 81.
[20] Ibid., pp. 80-3.

attributed to a law it involves difficult judgements about which of these intentions is predominant — and "there is no guarantee that any given law will lend itself to the imputation of an intention".[21] But placing intentional neutrality in the context of legal interpretation and sense-making does seem to extricate it from the kind of worries about the metaphysics and inwardness of intention that were canvassed in the previous paragraph.

However, one significant objection to the intentional account of neutrality is that it is still too weak. Return to the examples that show that justificatory neutrality is too weak — namely, the legal establishment of Buddhism as state religion, the development over time of a highly secular public culture that leaves religious citizens feeling alienated from their deepest convictions, and the adoption of nation-building policies that actively inculcate a particular common culture and way of life. The problem is that while some of these neutrally-justified non-neutral policies might be adopted on the basis of (hidden) non-neutral intentions, and thus be in breach of intentional neutrality, it is at least possible, and in some cases quite plausible, that these and other non-neutral policies *really were* adopted on the basis of neutral intentions. In such a case, the intentional account of neutrality would have to judge these policies as impeccably neutral.

In his recent book *Liberalism With Excellence*, Kramer seeks to defend his intentional account against this objection from neutrally intended yet non-neutral policies. Kramer's account of neutrality requires that *both* the end (such as social stability) and the means (such as religious establishment) be neutral. So on this view seeking the end of social stability via the means of religious establishment would not be considered neutral because the means chosen is non-neutral.

But what makes the means non-neutral? In what sense is religious establishment a non-neutral means? This is the crucial question that accounts of neutrality have been trying to answer. And for Kramer the neutrality of a means seems to depend on the intentions behind its adoption. Considering a structurally similar case, he says:

[21] Ibid., p. 81.

> The intention of a law pertains not only to the ends which the law is designed to serve, but also to the means selected for the advancement of those ends. A law which requires the recitation of prayers in public schools is non-neutrally intended to favor religious conceptions of the good over non-religious conceptions, even if the favoring of those former conceptions is based not on any ascription of truth to them but solely on instrumental considerations.[22]

The key claim here is that the means (namely, the recitation of prayers in public schools) is "is non-neutrally intended to favor religious conceptions of the good". Elsewhere, in a similar case involving religious establishment, he makes a similar claim: "legal-governmental officials have intentionally elevated one reasonable conception of the good over others as a means for the attainment of a neutral end".[23] As I read it, Kramer's argument here is this. A state intends the recitation of prayers in school as a means; the recitation of prayers in school has the effect of securing social peace as well as the effect of advantaging religious over non-religious conceptions of the good; so a state that intends the recitation of prayers in school as a means also intends to advantage religious over non-religious conceptions of the good; so such a state is non-neutral in intention.

The problem with this argument is that its implicit premise — namely that every effect of an intended action is an intended effect — is false. It is possible to intend a means (the prescription of prayer in school) without intending a causal effect of that means (that religion be advantaged). Suppose, for instance, that the state passes a law requiring the recitation of prayers in school and that, in addition to successfully achieving its intended end of social stability, it also, contrary to what many might have predicted, has the side effect of *disadvantaging* religion by putting a generation of young people off religion and promoting a widespread culture of atheism. Surely no intention of the legislators has been thwarted: though they might have *expected* that religion would prosper, they did not in any way *intend* for it to do so as a result of their policy. What this shows is that the legislators do *not* "intend to favor religious

[22] Ibid., p. 77.
[23] Ibid., p. 24. See also p. 19.

conceptions of the good", either as an end or as a means, and so non-neutral intention cannot explain the non-neutrality of the intended means.

There is, however, an alternative reading of Kramer's argument.[24] On this reading, it is not that, by intending to prescribe prayer, one intends to *cause* religion to do better — and in this sense one has a non-neutral intention. For clearly one can intend to prescribe prayer without intending to cause religion to do better. Rather, prescribing prayer itself *constitutes* a state of affairs in which religious conceptions of the good are favored and elevated over non-religious conceptions of the good (roughly in the sense that this state of affairs is logically entailed by the prescription of prayer in school). So perhaps Kramer's point is that, by intending to prescribe prayer, one intends an outcome that constitutes a state of affairs in which religious conceptions of the good are favoured and elevated over non-religious conceptions of the good — and in this sense one has a non-neutral intention. In other words: even if one can intend a means without intending its caused effects (as I argued above), one cannot intend a means without intending its constituted effects.

However, there are arguably many cases in which one can intend an action without intending the state of affairs of which that action is constitutive. For instance, if I have the intention of thanking three people I need not also thereby have the intention of thanking a prime number of people — despite the fact that my thanking a prime number of people is a state of affairs that is logically entailed and constituted by my thanking three people. Likewise, there seems no reason to insist that someone who has the intention of prescribing prayer in school must also thereby have the non-neutral intention of favouring and elevating religious conceptions of the good over non-religious ones — despite the fact that the favouring and elevating of religious conceptions of the good over non-religious conceptions of the good is logically entailed and constituted by the prescription of prayer in school. The general point here, familiar from the closeness problem, is that "intentions are intensional: however 'close' X is to Y, even if X is [constitutive of or] identical with Y, it is possible for an agent to intend that X and not to intend that Y".[25] So since one can intend a means without intending its constituted effects, it remains unclear

[24] I thank a reviewer for suggesting this.
[25] Hills, A., "Intentions, Foreseen Consequences and the Doctrine of Double Effect", *Philosophical Studies* 133 (2007), p. 265.

after all what grounds there are for saying that the prescription of prayer in school "is non-neutrally intended to favor religious conceptions of the good" or that it involves religious conceptions of the good being "intentionally elevated" over others.

4 Neutrality as equality of opportunity

In an attempt to avoid these problems, and to show greater sensitivity to effects, some theorists have argued that neutrality is best understood in terms of equality of opportunity. Raz, for instance, suggests that a neutral state is most plausibly understood as one that "ensure[s] for all persons an equal ability to pursue in their lives and promote in their societies any ideal of the good of their choosing".[26] And in a recent article Clarke argues that neutrality requires that state action "not have the effect of giving some people greater opportunities to pursue their ideals relative to other people's opportunities".[27]

Patten's much-discussed "neutrality of treatment", developed in his recent book *Equal Recognition,* can also be understood as part of this family of views of neutrality as equality of opportunity.[28] According to Patten, a state is neutral when, relative to a baseline of fair opportunity for self-determination, its policies are equally accommodating of different conceptions of the good.[29] Another way of putting this is that the neutral state is one that extends to holders of differing conceptions of the good the treatment (i.e. benefits, subsidies, exemptions, hindrances) that will enable them to enjoy fair opportunity for self-determination.

However, attempts to conceive neutrality in terms of equality of opportunity (Raz and Clarke) or fair opportunity for self-determination (Patten) run into a stubborn objection, which for simplicity I will consider in relation to Patten's

[26] Raz, J., *The Morality of Freedom,* p. 115.
[27] Clarke, S., "Consequential Neutrality Revivified", pp. 116-7. Although Clarke himself identifies with neutrality of consequences, he can be said to advocate neutrality as equality of opportunity because the effects that he focuses on are those relevant to the maintenance of equality of opportunity. In general, neutrality as equality of opportunity can be viewed as one particular species of neutrality of consequences, where the baseline is equality of opportunity and thus where the relevant consequences are those that affect equality of opportunity.
[28] Patten, A., *Equal Recognition: The Moral Foundations of Minority Rights* (Princeton, NJ: Princeton University Press, 2014), chapter 4.
[29] Ibid., p. 115.

account. The objection is that neutrality as fair opportunity for self-determination is too weak in the sense that it judges as neutral arrangements that are intuitively non-neutral. Consider, in Cordelli's examples, the following instances of symbolic religious establishment: crucifixes are displayed in Italian courtrooms; most Western countries have statutory holidays around Christmas, as well as sponsored Christmas trees and nativity displays; a taxpayer-funded chaplain opens each session of the US Senate with a prayer.[30] Patten's neutrality of treatment fails to recognize the non-neutral character of these instances of symbolic religious establishment because, on his view, if policies do not vitiate fair opportunity for self-determination then they do not violate neutrality of treatment. Standard ways of undermining opportunity for self-determination would be to deny someone the *liberty* to be self-determining, or the *resources minimally necessary* to be self-determining, or the *capabilities needed* to be self-determining.[31] But Christmas trees and Senate chaplains do not seem to rise to these levels of vitiation of opportunity for self-determination, and so their non-neutral character would go undetected.

In a recent paper, Patten has defended his account against this objection. Patten concedes that the baseline value of fair opportunity for self-determination needs to be supplemented by some "other fairness-related values" such as "fair access to the state's symbolic resources".[32] Once we do this, we see that "even purely symbolic establishment detracts from a person's opportunities".[33]

In my view this reply takes us back to the other extreme. That is, whereas Patten's original account was too weak to detect non-neutral symbolism that fell short of undermining fair opportunity, this revised version is too strong and pushes us towards an hypersensitive and unattainable form of neutrality reminiscent of the version of consequential neutrality canvassed in Section 1. The point is not that symbols and expressive messages can *never* undermine fair opportunity for self-determination and other fairness-based values. If the state systematically put out a powerful symbolic message that led to some

[30] Cordelli, C., "Neutrality of What?", *Critical Review of International Social and Political Philosophy* 20 (2017), pp. 40-2.
[31] Ibid., p. 40.
[32] Patten, A., "Equal Citizenship, Neutrality, and Democracy: A Reply to Critics of *Equal Recognition*", *Critical Review of International Social and Political Philosophy* 20 (2017), p. 132.
[33] Ibid.

groups being stigmatized and marginalized, then that, of course, would violate fair opportunity for self-determination. The point is just that if, because of an expanded range of "fairness-related values" undergirding neutrality, the threshold of what counts as a barrier to opportunity became so low that even a state-funded nativity display would be sufficient to undermine opportunity for self-determination, then it would be unclear what kind of policies *wouldn't* undermine opportunity and thus what kind of policies *could ever* pass the test of neutrality. For instance, if a state-funded Christmas tree undermines a person's opportunity to pursue a non-Christian way of life, then it seems that a state-funded advertisement about joining the National Health Service would undermine a person's opportunity to pursue a career outside of medicine. Expanding the moral basis of neutrality in this way, in other words, seems to leave us with an implausibly fragile account of what an "opportunity" for self-determination is and thus with an excessively demanding account of neutrality.

5 A pluralist account of neutrality

I have tried to organize and evaluate the main theories of neutrality. In answer to the question of what is meant by "state neutrality", some, as I have explained, say that the state should be neutral in the sense of ensuring that its policies do not have the *effect or consequence* of making some conception of the good life more successful or popular or easily realizable than others (Goodin, Reeve, Wilkins). Others say that the state should be neutral in the sense that the *reasons and justifications* given in support of its policies must not be grounded in any conception of the good life but rather must be based on publicly accessible or intelligible considerations (Ackerman, Barry, Dworkin, Gaus, Larmore, Lister, Quong, Rawls, Vallier). Still others say neutrality requires that, when legislators and government officials make public policy, they do not *aim or intend* to promote some conceptions of the good and to disadvantage others (Brighouse, Kramer). And yet others say that neutrality requires ensuring that all conceptions of the good within society have some *fair or equal opportunity* to prosper (Clarke, Patten, Raz).

But which should we select amongst these rival formulations of the liberal principle of state neutrality? Naturally, each of these accounts captures *something* of the concept of neutrality. But, as I have pointed out, each one,

considered alone, has implausible implications. Indeed the objections to one account often make reference to the values and insights present in another. For instance, neutrality of justification and neutrality of intention were criticized for being insufficiently sensitive to consequences. What this suggests is that all the accounts have something to be said for them.

My own view, then, is that the most plausible account of neutrality will be pluralist in character. Neutrality, on this view, is a complex and multifaceted concept with consequential, justificatory, and intentional dimensions — a concept not amenable to a unidimensional analysis. Consequences, justifications and intentions are all relevant to neutrality; but no one of these three dimensions captures the concept of neutrality in its entirety.

In what follows, I present two different forms of pluralism about neutrality — an all-or-nothing pluralism with thresholds, and a matter-of-degree pluralism without thresholds — and argue for the latter.[34]

5.1 All-or-nothing pluralism with thresholds

Here is a natural first attempt to formulate a pluralist account of neutrality, an attempt that stays close to the spirit of the original proposals and involves simply conjoining the three previous analyses of neutrality:

> A state is neutral if and only if: (i) the reasons and justifications given in support of its laws, policies and institutions do not appeal to any particular controversial conception of the good life but rather are based on publicly accessible or intelligible considerations; *and* (ii) the state does not act with the aim or intention of promoting one or another conception of the good life; *and* (iii) the effects of state action are not such as to make it exceedingly difficult for some individuals to live out their reasonable conception of the good life.[35]

[34] I thank Jeff McMahan for helping me to distinguish between these two versions of pluralism about neutrality.

[35] Neutrality as equality of opportunity disappears because it gets folded into neutrality of consequences for reasons given in footnote 27.

This form of pluralism combines the insights of each of the consequential, justificatory and intentional accounts of neutrality. But it retains the all-or-nothing structure characteristic of the mainstream analyses of neutrality and this structure gives rise to two closely related problems.

First, the "exceeding difficulty" threshold in clause (iii) might still be too strong. Recall from Section 1 that critics rejected consequential neutrality on the grounds that *any* law, institution or policy is bound to have differential effects across the full range of conceptions of the good that citizens hold. This objection, in this form, no longer applies because the consequential clause has been weakened from requiring policies to have *absolutely equal effects* to requiring only that the (inevitably differential) effects of policies not render it *exceedingly difficult* for some individuals to pursue their reasonable conception of the good. But a revived version of this objection might say that any law, institution or policy is *still* bound to make it exceedingly difficult for *some* individuals to pursue their reasonable yet highly eccentric conception of the good. You can't please absolutely everyone. Given the sheer enormity and variety of conceptions of the good within society, that is, do not the "the facts of commonsense political sociology" tell against the practicability of *even this weakened* consequential requirement?

Perhaps the defender of the all-or-nothing account might reply by further shifting the threshold from "exceeding difficulty" to something even weaker (for example, "virtual impossibility"). Beyond having a whiff of the ad hoc, the problem with this reply is that such a threshold would start to move towards the other extreme of being too insensitive to consequences. This produces the following dilemma.[36] If, on the one hand, we set a very strong and stringent threshold (for example, state action must not *differentially affect* conceptions of the good), then neutrality will be impossible to achieve. Yet if, on the other hand, we set a very weak threshold (for example, state action must not *make it virtually impossible* for individuals to pursue their conceptions of the good), then our account becomes insufficiently sensitive to consequences and thus many non-neutral consequences that significantly advantage some conceptions of the good over others go unrecognized.

[36] I take this to be the same problem faced by Patten's account, noted in the final paragraph of Section 4.

The second problem is that this account treats neutrality as an all-or-nothing affair. In doing so — in treating neutrality and perfectionism as mutually exclusive — it follows one of the standard assumptions in the perfectionism-neutrality literature, an assumption revealed in the fact that neutrality is also referred to as "antiperfectionism". For instance, Quong defines a perfectionist as someone who answers "yes", and a neutralist as someone who answers "no", to the question: "Is it permissible for a liberal state to promote or discourage some activities, ideals, or ways of life on grounds relating to their inherent or intrinsic value, or on the basis of other metaphysical claims?"[37] One must, in other words, be *either* a perfectionist *or* a neutralist.

I find this implausibly dichotomous. It seems clear that states can be more or less neutral, and that pursuing certain policies can increase or decrease the neutrality of the state without wholly and utterly establishing or eliminating such neutrality. The point, to clarify, is not just about linguistic usage: it is not just that we can grammatically describe states as "more neutral" and thus that we need an account of neutrality that lets us do this. Rather, the point is that the structure of the concept of neutrality is most plausibly thought of as a matter of degree and that the binary, all-or-nothing account lumps together political arrangements that are worth distinguishing. This becomes particularly clear when we move away from comparing two stark opposites (for instance, a highly perfectionist state and a highly neutral state) and begin to compare a wider spectrum of cases (for instance, the only slightly perfectionist state and the highly perfectionist state).

One can think in terms of two extremes, each one of which is to all intents and purposes an impossibility. At one extreme, there is the maximally neutral state: where the state justifies its policies in terms of the promotion of health and social peace and other such public values, where the intentions of all state officials track only these sorts of public values, and where state action does not have any discernable differential effect, either advantageous or disadvantageous, across the full range of conceptions of the good and ways of life that exist among its citizens. At the other extreme is the maximally perfectionist state: where, for instance, the state justifies its policies in terms of

[37] Quong, J., *Liberalism Without Perfection*, pp. 15-21, esp. p. 19.

Biblical scripture and the promotion of Christian ideals, where the intentions of all state officials track only these sorts of Christian values, and where the effect of this state action is to massively advantage Christian citizens and to grievously disadvantage holders of all other conceptions of the good that exist within society. And, crucially, between these extremes it makes perfect sense to think that states can uphold the ideal of neutrality to a greater or lesser extent along its various dimensions. In short, neutrality comes in degrees.

5.2 Matter-of-degree pluralism without thresholds

A better kind of pluralist account of neutrality, and one that does not suffer from the problems faced by all-or-nothing pluralism with thresholds, is what I am calling matter-of-degree pluralism without thresholds. On this view:

> A state is neutral *to the extent that* (i) the reasons and justifications for its laws, policies and institutions do not appeal to any particular controversial conception of the good life but rather are based in publicly accessible or intelligible considerations; *and* (ii) the state does not act with the aim or intention of promoting one or another conception of the good life; *and* (iii) state action does not have the effect of advantaging some reasonable conceptions of the good life and of disadvantaging others.

In addition to the generally attractive pluralist sensitivity to justifications and intentions and consequences, the specific advantages of this account of neutrality, relative to the previous pluralist account, are clear: it does away with difficult-to-specify thresholds, and it avoids the tendency to treat neutrality as a binary all-or-nothing concept.

At this point, I need to explain how these three dimensions of neutrality relate to and interact with one another. How should the justificatory, intentional and consequential dimensions be balanced against each other? Are they all equally relevant to overall assessments of neutrality? Or do some dimensions have greater weight than others? What happens when the dimensions of neutrality conflict? That is, how should we think about the all-things-considered

neutrality of a state that is very neutral in one respect, say justifications, but very non-neutral in other respects, say consequences?

These are difficult questions and they represent a further line of inquiry that emerges from my discussion. Indeed, there may even be methodological reasons for thinking that there is no general answer to the question of how the individual dimensions contribute to overall neutrality in particular cases.[38] Nonetheless let me just suggest one way in which progress might be made here, namely by trying to connect these questions to an inquiry into the moral foundations of liberal neutrality.

Defenders of liberal neutrality are typically at pains to point out that neutrality is not a principle that flows from, or is justified by reference to, some comprehensive religious or ethical doctrine, for example a Millian doctrine of individuality and "experiments in living" — because to ground neutrality in a comprehensive doctrine would be to depart from the spirit of neutrality.[39] As Larmore puts it, liberals must give a "neutral justification for neutrality".[40] So, when pressed to spell out the moral foundations of neutrality, the idea of political liberals seems to be that the principle of neutrality can be justified by reference to very generic, minimalist, fundamental or uncontroversial notions acceptable to all, whatever their comprehensive worldview — notions such as respect for persons. For instance, Gaus grounds neutrality in fundamental reason-giving features of social morality and in the idea that one must give justificatory reasons when violating the presumption in favour of liberty.[41] Relatedly, Larmore argues that neutrality is grounded in respect: on his view, to justify coercion of another person by offering them reasons drawn from a comprehensive religious or ethical doctrine which that person might reasonably reject would be to treat them without due respect, to treat them

[38] See Frances Kamm's Principle of Contextual Interaction in her *Morality, Mortality, Vol. II* (Oxford: Oxford University Press, 1996), as well as Kagan, S., "The Additive Fallacy", *Ethics* 99 (1998), pp. 5-31.
[39] This, at least, is how neutrality is defended by political liberals. Comprehensive liberals, by contrast, have no compunction about grounding the principle of state neutrality in some comprehensive doctrine or conception of the good life.
[40] Larmore, C., *Patterns of Moral Complexity*, p. 53.
[41] Gaus, G., "Liberal Neutrality: A Compelling and Radical Principle", pp. 137-65.

"merely as means, as objects of coercion, and not also as ends, engaging with their distinctive capacity as persons".[42]

Inquiry into the moral foundations of neutrality is important because the conception of neutrality that one prefers will likely depend on the reasons one has for valuing neutrality in the first place. Some reasons for valuing neutrality (such as Gaus's ideas about fundamental norms of justification) naturally lend themselves to certain conceptions of neutrality (such as justificatory neutrality). And given that many political liberals agree with Gaus and Larmore that neutrality is grounded in fundamental norms of justification and reason-giving, it is hardly surprising that they too have focused entirely on the justificatory dimension of neutrality. But if one had a different view about the moral foundations of liberal neutrality, then that might explain the significance of the other dimensions of neutrality, such as the consequential dimension.

For instance, Lister grounds neutrality in notions of civic friendship and political community.[43] And these foundations might explain the role and significance of consequences and intentions. For it is as plausible to say that civic friendship cannot be maintained across great consequential and intentional divides as it is to say that civic friendship cannot be maintained across great justificatory divides. To motivate the claim that justificatory neutrality realizes civic friendship, Lister draws an analogy with marriage. In a marriage, says Lister, neither partner will make decisions on the basis of reasons that the other partner rejects because to do so would compromise their "shared aspiration to act together in concert".[44] And just as partners within a marriage would "agree to bracket the contested considerations…for the sake of acting in concert" when making family decisions, so too would members of a political community when making political decisions.[45] But this acting-in-concert rationale applies equally well to the consequential and intentional dimensions of neutrality. For suppose that, despite both partners agreeing to bracket contested considerations, the decisions ultimately made within a marriage systematically advantage the husband and the choices always align

[42] Larmore, C., "The Moral Basis of Political Liberalism", *The Journal of Philosophy* 96 (1999), p. 607.
[43] Lister, A., *Public Reason and Political Community* (London: Bloomsbury, 2013).
[44] Ibid., p. 108.
[45] Ibid.

with his needs, interests and preferences. Insofar as there is such a heavy tilt in favour of one partner, it is hard to see how this is a marriage in which the partners are "acting together in concert". Or suppose that a wife has non-neutral intentions: she cleverly manages to persuade her atheist husband to send their child to a religious school on the basis of neutral reasons such as the quality of the teaching, but with the non-neutral intention that their child becomes religious. Again, it is doubtful that this is a marriage in which the partners are "acting together in concert", despite the fact that the wife complies with the agreement to bracket contested considerations. So, short of simply stipulating that "acting together in concert" just is a matter of acting on the basis of shared reasons, there seems no reason to think that this rationale supports neutrality's justificatory dimension but not its consequential and intentional dimensions.

Others argue that social stability is the moral basis of neutrality. One way of understanding this idea is that justifying laws in terms that everyone can accept has stabilizing effects: it fosters trust and cooperation, and it averts feelings of resentment, frustration and anger.[46] But social stability arguably depends just as much on the *consequences* of state action and on the *intentions* of legislators as on justifications. For there would presumably also be destabilizing effects — including resentment, anger and the compromise of trust — if laws and policies made it extremely difficult for some groups to adhere to their deeply held convictions about the good life, or if it were discovered that legislators were designing laws with the intention of helping some segments of society and hindering others.

In short, then, we might come to a better understanding of the relations and interactions between, and the relative importance of, the different dimensions of neutrality by first coming to a better understanding of the moral foundations of neutrality and then thinking about how those foundational values are served by each of the three dimensions of neutrality.

There are different ways of envisaging the project that I am sketching here. One would be to see it as a critical project, an attempt to dethrone public reason as

[46] See Vallier, K., *Must Politics Be War? Restoring Our Trust in the Open Society* (Oxford: Oxford University Press, 2019).

the dominant strand of contemporary liberal theory by showing that public reason tells only part of the broader story worth telling and that, while it captures the justificatory and reason-giving aspect of neutrality, it neglects neutrality's intentional and consequential dimensions. But the reflections in the previous few paragraphs may also point to a sense in which mine is a reconciliationist project: those philosophers who identify themselves with one or another of the dimensions of neutrality might in fact at a deeper level be described as agreeing about pluralism but just disagreeing about relative weightings because of disagreements over the extent to which each of the three dimensions serves the moral foundations of neutrality.[47] Different theorists of neutrality, on this proposal, simply believe that their favoured dimension must be accorded much more weight (but surely not infinitely more weight) than the valid yet (to them) less significant dimensions of neutrality favoured by other theorists.

6 Objections and further lines of inquiry

I have tried to argue that neutrality is best understood as a complex, multifaceted concept with consequential, justificatory and intentional dimensions. This account departs from the existing accounts of neutrality both in its pluralism and its being a matter of degree. In this section I defend and clarify my account through responding to four objections; I then suggest some further lines of inquiry.

A first objection involves questioning the inclusion of clause (iii), which builds a consequential aspect into my account of neutrality. This clause stands in urgent need of defence because, as I indicated in Section 1, it is inevitable that state action will have differential impact across different conceptions of the good, and so the overwhelming majority of philosophers have considered consequential neutrality to be impossible or impracticable and have instead focused on either justificatory or intentional neutrality. Rawls, as we saw, puts it like this: "we must accept the facts of commonsense political sociology…Neutrality of effect or influence political liberalism abandons as

[47] There is an echo here of MacCallum's attempt to reconcile competing accounts of liberty within his "triadic" analysis. See MacCallum, G., "Negative and Positive Freedom", *The Philosophical Review* 76 (1967), pp. 312-34.

impracticable".[48] Sher uses language of entirely representative strength when he says that "*any* acceptable interpretation *must* construe neutrality as a property of the justifications of laws, policies, or institutions rather than their effects".[49] Arneson says that "it is immediately clear that *nobody* who wants to defend neutrality on the good would really want to defend neutrality of effect".[50] And for Kramer, "neutrality of effect is a disastrous nonstarter" and a "preposterous infeasibility".[51]

In my view, this vehement and near-unanimous rejection of consequences by theorists of neutrality is unwarranted. It is one thing to say that it is impossible to guarantee *absolutely equal* effects or that policies are bound to have *some* differential impact; but it is quite another thing to say that, therefore, *any* effects are compatible with neutrality and that consequences have no bearing on neutrality whatsoever. It does not follow, in other words, from a rejection of a strict outcome-egalitarian *version* of consequential neutrality that consequential neutrality must be rejected *generally*. It seems to me that, even if it would be "impracticable" (in Rawls's term) to try to control *all* the slings and arrows of fortune, neutralists should at least try to prevent or compensate for the most egregiously disadvantaging of the effects of state action. After all, should we not be concerned, at the bar of neutrality, if political structures have the cumulative effect — intended or not, neutrally justified or not — of making it extremely difficult for some citizens to live in accordance with their conception of the good? Can we not question the neutrality of a system that, systematically over time, has the spillover effect of killing off some fine and subtle culture, some cherished form of human excellence, some worthy pattern of thought and action?

Rawls, in a revealing admission, seems to be aware of the dangers of his neglect of the unintended consequences of liberal politics. He acknowledges, in a footnote, that "abandoning neutrality of effect or influence as impractical may lead to an excessively secular public life and background culture" and elsewhere says that to provide children with a liberal education may be "in effect, though

[48] Rawls, J., *Political Liberalism*, pp. 193-4.
[49] Sher, G., *Beyond Neutrality*, p. 22 (emphases added).
[50] Arneson, R., "Liberal Neutrality on the Good: An Autopsy", p. 193 (emphasis added).
[51] Kramer, *Liberalism with Excellence,* p. 13.

not in intention, to educate them to a comprehensive liberal conception".[52] "Doing the one," Rawls says, "may lead to the other, if only because once we know the one, we may of our own accord go on to the other".[53] Ultimately, concludes Rawls, these unintended spillovers "may have to be accepted, often with regret".[54] Here though I think it is right to ask: "But is regret, however sincere, enough?"[55] For can there really be no complaint, from the point of view of neutrality, if the liberal state pursues policies that contribute so directly, so powerfully, so foreseeably to the demise of certain ways of life? Even if pure neutrality of effect is impossible, would a state that aspires to neutrality not at least try — as a way, perhaps, of giving substance to its professions of regret and commiseration — to prevent, to minimize, or to compensate for these cumulatively powerful corrosive effects?

It is for this reason that I am inclined to say that any account of neutrality must contain some consequential aspect, as in my clause (iii). Of course, it is true that much state action will have the effect of "advantaging some conceptions of the good life and of disadvantaging others", as I put it there, and so no state will perfectly satisfy the condition in clause (iii). But, while this might prove fatal for the kind of simple all-or-nothing consequential accounts of neutrality we assessed in Section 1 (because on those views it would imply that no state is ever neutral), it need not be so for the pluralist, matter-of-degree account. After all, on the matter-of-degree view things are more nuanced, and the neutrality of the state varies according to how slight or grievous the differentially advantaging effects of state policy are.

At this point — and this is the second objection — political liberals might protest that, with the admitted exception of Barry,[56] they have long been attuned to the importance of counteracting especially burdensome consequences imposed on certain groups by neutrally justified laws and that they have frequently defended exemptions, subsidies and other disadvantage-countering measures on varied grounds including fair equality of opportunity,

[52] Rawls, J., *Political Liberalism*, p. 194 fn. 28, p. 199.
[53] Ibid., p. 199.
[54] Ibid., p. 200.
[55] Wilkins, B., "A Third Principle of Justice", p. 358.
[56] Barry, B., *Culture and Equality* (Cambridge: Polity Press, 2001).

rights of conscience, and the social bases of self-respect.[57] Indeed, some of the leading figures in multiculturalist thought — most notably Kymlicka and Patten — are liberal neutralists, and they have been very alert to the vulnerability and fragility of various cultural practices and institutions. So it might be thought that political liberals can already explain all that they need to explain. Isn't justificatory neutrality plus special accommodation (based on equality of opportunity, or conscience, or self-respect) enough? What is to be gained from subsuming these matters under the heading of neutrality?[58]

Let me make two points in reply. First, I do not think that the package comprising justificatory neutrality plus special accommodations is as extensionally adequate as pluralist neutrality. For there are a number of cases of non-neutral policy that only pluralist neutrality can recognize. Consider again symbolic religious establishment which is neutrally justified (perhaps the crucifixes in courtrooms encourage witnesses to tell the truth) and which, while imposing a small cost on some citizens, does not rise to the level of undermining their self-respect or their rights to conscience or their opportunity for self-determination. The non-neutral character of such symbolic establishment can be explained by pluralist neutrality, but goes undetected by justificatory neutrality plus special accommodation based on self-respect, conscience or fair opportunity for self-determination. Or consider a case in which devoutly Christian state officials choose to enact a set of decently egalitarian and liberal policies on the basis of theological considerations and with the intention that doing so will best express the law of God as articulated in the Bible. This case clearly offends against the intentional dimension of pluralist neutrality. By contrast, insofar as this state's policies can (alongside the theological justifications) be given a neutral justification, and insofar as the policies secure for all autonomy and self-respect and fair opportunity for self-determination, it could not be faulted in any respect by liberals who espouse justificatory neutrality plus special accommodation.

[57] See, e.g., Quong, J., "Cultural Exemptions, Expensive Tastes, and Equal Opportunities", *Journal of Applied Philosophy* 23 (2006), pp. 55-73; Laborde, C., *Liberalism's Religion* (Cambridge: Harvard University Press, 2017).

[58] I thank three reviewers for putting this objection to me, and in particular the reviewer whose statement of the objection I have followed closely here.

Perhaps political liberals will be tempted to try to accommodate the cases of theologically motivated egalitarian liberalism and of symbolic religious establishment by building intentions and consequences into justificatory neutrality, such that a policy is publicly justifiable only if is not enacted with the wrong aims or intentions, and only if that policy does not have the effect of advantaging some conceptions of the good life and disadvantaging others. Given a case that they cannot explain, that is, political liberals might be tempted to simply *expand* the available justificatory factors until neutrality of justification speaks against the case in question. However, while this approach is formally possible, it would represent a pyrrhic victory for proponents of justificatory neutrality because it would convert justificatory neutrality into pluralist neutrality in all but name.

Alternatively, perhaps political liberals will be tempted to bite the bullet here and to maintain that there really is nothing at all objectionable about such theologically motivated egalitarian liberalism provided that it is genuinely egalitarian-liberal in substance. This thought might stem from, or be reinforced by, a general suspicion of the view that intentions are relevant to political legitimacy. Consider, however, two otherwise identical just systems of laws. The only difference is that in the first case legislators are motivated by the content of their democratic mandate, whereas in the second case legislators pass the just laws because they have been promised benefits by a group of wealthy donors if they do so. The fact that the legislators' intentions track these promised benefits rather than the terms of their democratic mandate arguably deprives the laws of their legitimacy because they are not imposed on citizens in their name, for their sake, and on grounds of respect for them and their interests. This case suggests, then, that at least in certain instances untoward intentions and motives can cast doubt on political legitimacy.[59]

The second point, though, is that even if liberal neutralists are able to appeal to other principles within liberal theory besides neutrality in order to explain and support disadvantage-countering measures, this does not undermine, or indeed even touch, my central claim. What I claim is that neutrality has a consequential aspect and thus that our reasons for countering disadvantage are reasons *of neutrality*. This is just what follows if the questions posed in the

[59] I thank a reviewer for suggesting this case.

preceding paragraphs are answered in the affirmative; and this is what is denied by liberal neutralists who reduce neutrality to justificatory neutrality and to the giving of public justifications and thus whose concern with special accommodations generally has more to do with their being *liberal* than their being *neutralist*. I do not, however, claim that our reasons for countering disadvantage are *only* reasons of neutrality. There might also be reasons of fair opportunity, of self-respect, of autonomy, or of conscience for countering the disadvantaging or burdensome consequences of state action. So even if the objector's charge is correct, all this would show is that special accommodations and disadvantage-countering measures are morally overdetermined, not that there are no reasons of neutrality in play here.

The consequence-sensitivity of pluralist neutrality also gives rise to a third objection, namely that pluralist neutrality involves an objectionable indulging of expensive tastes.[60] After all, if states try, for the sake of neutrality, to prevent or compensate for the relative disadvantages faced by adherents of minority or unpopular conceptions of the good, then this would seem to amount to the accommodation of those minority groups' expensive tastes and thus to some people being forced to subsidize the preferences of others. One might therefore worry that pluralist neutrality is incompatible with the idea that individuals should take personal responsibility for their conception of the good life and should bear the costs and burdens associated with its pursuit.

This objection allows me to clarify the nature of my argument: all I have been arguing is that neutrality should be understood in terms of a general three-dimensional schema that involves consequences and intentions as well as justifications. And according to the consequential dimension of this schema, a state acts non-neutrally when its laws, policies and institutions have the effect of advantaging some conceptions of the good life and of disadvantaging others. Importantly, this leaves open what it means to advantage or disadvantage. I have not, that is, specified what the metric or "currency" of advantage is. What the problem of expensive tastes demonstrates, I think, is that we should steer clear of purely welfarist conceptions of advantage — and perhaps it encourages us to view advantage in terms of resources. But the consequential dimension of pluralist neutrality is not wedded to any particular currency, and beyond

[60] I thank two reviewers for pressing me on this point.

expressing my scepticism about pure welfarism I do not here take a stand on this question: my general schema is compatible with whatever currency best overcomes the problem of expensive tastes and other related problems such as the baseline problem mentioned in Section 1, just as it is compatible with different ways of understanding justification and intention. The other point worth stressing here, though, is that on my view neutrality is not an absolute constraint and thus there may indeed be cases in which it is overridden by considerations of personal responsibility. So the proponent of pluralist neutrality would not in any case be all-things-considered committed to the indulgence of all instances of expensive tastes.

A fourth and more programmatic objection involves accepting my account, as an analysis of the concept of state neutrality, but rejecting its broader implications for liberal theory. I have argued that neutrality is best construed in pluralist terms and, therefore, that public reason liberalism, *insofar as it aspires to neutrality*, tells only part of the full story. But — so this objection goes — I have not argued for the italicized clause, and indeed this clause can be rejected. For political liberals and public reason liberals may claim that they were never really that interested in neutrality per se: the aim, they may say, was never to ensure that laws and institutions were *neutral* but rather that they were *publicly justified*. Of course, it is true that early political liberals such as Dworkin and Larmore did pay attention to the concept of neutrality and did define their views in terms of neutrality; but, so this thought goes, contemporary liberalism no longer needs to be tied to the concept of neutrality and may even find it unhelpful or misleading (as indeed Rawls did[61]), preferring, instead, to talk about what is and is not publicly justified. In short, then, the objection is that if political liberals in fact care about public justification on its own terms as the best account of political morality and not insofar as it is the best interpretation of some commitment to neutrality, then it does not follow from accepting pluralist neutrality that they must now also take into account consequences and intentions.[62]

[61] See Rawls, J., *Political Liberalism*, p. 191.
[62] I thank Anthony Taylor for putting to me an objection along these lines.

I see two problems with this move.[63] The first is that many contemporary political liberals *do* describe their views as bound up with, or as interpretations of, neutrality and I see no reason not to take these theorists at their word. For instance, Quong states that the "main thesis" of his book is that "government should remain *neutral* on the issue of the good life", Gaus thinks that "liberal neutrality" is a "compelling and radical principle" that prohibits using state power to make others "more perfect in our own eyes", and Vallier says, "I interpret the ideal of liberal neutrality as the ideal of public justification".[64] Of course, it is possible that these political liberals are misspeaking, or speaking loosely, and that what is really morally operative in their view is public justification and not neutrality per se. But this leads to the second problem, which is simply that without a robust commitment to neutrality political liberals cannot say what they will want to say about a range of cases in which state action, despite being publicly justified, is intuitively non-neutral — cases such as Kramer's Convivia, the gradual secularization of public culture, the adoption of nation-building policies that inculcate a fairly thick common way of life, symbolic religious establishment, and theologically motivated liberal egalitarianism. Political liberals who focus exclusively on public justification do not seem to see anything untoward or troubling in these cases. So for political liberals to be able to draw the distinction, highly significant for liberalism, between a state that favours some religions and conceptions of the good over others and a state that refrains from such favouritism — where this favouritism is not fully captured by justificatory considerations — they must, I suggest, frame their view in terms of pluralist neutrality in general rather than justificatory neutrality (or public justification) in particular.

Finally, let me suggest some further lines of inquiry that emerge from the foregoing discussion. Some have already been mentioned: How are the different dimensions of neutrality to be weighed, ordered or balanced against each other? And does this depend, in turn, on the moral foundations of neutrality and the extent to which each dimension serves those foundations? In terms of what currency of advantage should the consequential dimension of pluralist neutrality be understood? Other questions will be about

[63] For a further discussion, see Section 5 of Chapter 4.
[64] Quong, J., *Liberalism Without Perfection*, p. 2 (emphasis in original); Gaus, G., "Liberal Neutrality: A Compelling and Radical Principle"; Vallier, K., *Liberal Politics and Public Faith*, p. 46.

implementation: How, as an epistemic and practical matter, should the effects and influences of state action on conceptions of the good life be ascertained? What kind of heuristics and rules of thumb would be suitable for these political purposes?

Another line of inquiry concerns the implications of this account of state neutrality on contemporary debates about liberalism, neutrality and perfectionism. What would be gained, and indeed what might be lost, by re-constructing the debate in this way? On my reading of the current state of the debate, there is now a near-unanimous consensus that neutrality is best understood in terms of justification and, therefore, that the neutral state is one that adheres to some principle of public justification. One of the implications of the pluralism of my account, then, is that this justificatory turn is one-sided and constraining, and ignores other important aspects of the liberal aspiration to neutrality. Indeed, pluralist neutrality might bring liberals back into closer touch with their tradition of thought. For instance, while the case can be made that J. S. Mill was concerned with public justification,[65] his concerns are arguably more aptly characterized in terms of consequences. Mill's liberalism grew out of a fear that the "despotism of custom" and excessive state interference were having the effect of channeling individuals into stultifying, conformist ways of life not of their own choosing; of quashing individuality, "eccentricity of conduct" and "experiments of living"; and of "maim[ing] by compression, like a Chinese lady's foot, every part of human nature which stands out prominently".[66] The concern seems to be about the *influences and effects* of state (and non-state) action just as much as its justifications. At the same time, viewing neutrality as a matter of degree, rather than as an all-or-nothing constraint, makes possible a richer and more nuanced debate between neutralists and perfectionists. For if a state is always neutral to some extent and in some respects, while simultaneously perfectionist to some other extent and in some other respects — if, in other words, states will always be part-neutral and part-perfectionist — then this shifts the focus of the debate away from the question of *whether or not* states should be perfectionist (or neutral) and

[65] See, e.g., Turner, P. "Social Morality in Mill", in P. N. Turner and G. Gaus (eds), *Public Reason in Political Philosophy: Classic Sources and Contemporary Commentaries* (New York, NY: Routledge, 2018), pp. 375-400.
[66] Mill, J. S., *On Liberty* (London: Penguin Books, 1985), p. 120, 126, 135, 136.

towards the question of *to what extent and in relation to what ends* states should be perfectionist (or neutral).

Deontic Perfectionism

It is a central tenet of much contemporary liberal political philosophy that the state should be, in some sense, neutral among the competing conceptions of the good life that exist within society.[1] A neutral state, it is often said, is one that does not impose or promote any particular ethical, religious or metaphysical doctrine and instead simply upholds a fair framework of laws and institutions within which each individual can freely pursue her own conception of the good life. By contrast, perfectionists hold that one function of the state is to encourage citizens to pursue virtuous, meaningful, or worthwhile ways of life and to discourage them from pursuing degrading or hollow ways of life. To do this, perfectionist states will naturally need to take a stand on a variety of ethical questions, in particular the question of what constitutes human flourishing and excellence.

Perfectionists, however, have not generally been clear about whether they take perfectionist interventions to be permissible or mandatory. Is the perfectionist thesis, that is, that the state *may* promote the good life or, more strongly, that it *must* do so? From what they say, most perfectionists seem quite content to defend the weaker permissibility claim. In this paper, I argue that perfectionists should set their sights higher. I thus defend what I call *deontic* perfectionism, according to which the promotion of art, culture, virtue and other aspects of the good life is a *duty* of the state. In Section 1, I introduce the distinction between deontic and non-deontic forms of perfectionism, arguing that despite its simplicity this distinction has been neglected and that the majority of perfectionists would appear to endorse non-deontic perfectionism. In Section 2, I make a case for deontic perfectionism by defending the idea of a right to a flourishing life. In Section 3, I show how "going deontic" can help perfectionists rebut a number of powerful objections that have recently been leveled against perfectionism. Finally, I defend deontic perfectionism against the charge that it violates the public justification principle (Section 4).

[1] See, e.g., Larmore, C., *Patterns of Moral Complexity* (Cambridge: Cambridge University Press, 1987); Rawls, J., *Political Liberalism: Expanded Edition* (New York: Columbia University Press, 2005); Quong, J., *Liberalism Without Perfection* (Oxford: Oxford University Press, 2011); Patten, A., "Liberal Neutrality: A Reinterpretation and Defense", *The Journal of Political Philosophy* 20 (2012), pp. 249-72; Schouten, G., *Liberalism, Neutrality, and the Gendered Division of Labor* (Oxford: Oxford University Press, 2019).

1 Deontic and non-deontic perfectionism

Perfectionists hold that that one function of the state is to promote human flourishing and the good life. But they have not always been clear about what *kind* of function this is. According to what I call *non-deontic perfectionism*, it is permissible but not mandatory for states to promote the good life, whereas according to what I call *deontic perfectionism* it is mandatory for states to promote the good life.

These two forms of perfectionism admit of internal variation. There are various ways of being a deontic perfectionist, for instance. One could hold (as I shall argue in Section 2.3) that individuals have a right to the conditions for a flourishing life and that this right generates correlative duties on the part of states to promote the flourishing of the citizenry. Or one might hold that perfectionist duties derive from other duties such as the state's duty to promote justice or self-respect.[2] For instance, perhaps having a warranted sense of self-respect requires a commitment to pursuing a sound conception of the good life, and so the state's perfectionist duties to promote sound ideals of the good life derive from its general duties to promote the self-respect of citizens. Nor is deontic perfectionism as a political theory wedded to deontology as a moral theory. For a consequentialist might take human flourishing to be part of her theory of the good and thus support deontic perfectionism — the view that states ought to promote human flourishing — on the grounds that states ought to do what promotes the most good. Similarly, there are a variety of potentially distinct ways of being a non-deontic perfectionist: one could hold that perfectionist policies are *permissible*, or that they are *legitimate*, or that they are *justified*, or that there is a *presumption* in favour of them, or that there are *reasons* to pursue them — where each of these notions on its own falls short of establishing a *duty* of perfectionism.

[2] As this point illustrates, the distinction between deontic and non-deontic perfectionism is related to yet distinct from the distinction, drawn by Quong, between "perfectionist justice" and "non-justice-based perfectionism" (see Quong, J., *Liberalism Without Perfection*, pp. 29-30). These distinctions diverge because not all duties are duties of justice. For instance, a deontic perfectionist might hold that states have a moral duty to promote impersonal goods (e.g. through preservation of sites of natural beauty or historical significance), but that this is not a duty of justice. One could therefore endorse deontic perfectionism without endorsing perfectionist justice.

Though most perfectionists have not explicitly distinguished between the claim that the state may promote the good and the claim that the state must promote the good, they have tended to state their views in non-deontic terms. In his book-length defence of perfectionism, Sher advances the view that "a government *may* legitimately promote the good", and in a recent paper he says, more explicitly, "although I don't think any government is under an *obligation of justice* to promote excellence among its citizens, I fully agree…that governments often have good *reason* to do just that".[3] Hurka argues that his perfectionism "approves" of "much non-coercive promotion of the good".[4] Chan states: "political perfectionism, which is the focus of discussion in this article, says that it is *permissible* for the state to design its political arrangement or policies with the aim of promoting what the state (or those citizens acting on the behalf of the state) thinks are worthwhile goods and ways of life".[5] In his concluding sentence he describes state perfectionism as "desirable, unavoidable and legitimate".[6] In *Liberal Perfectionism* Couto argues for the view that "the state is *permitted* and *justified* to protect and promote the opportunities for citizens to engage with intrinsically valuable goods on the basis of their intrinsic value".[7] And Metz defends what he calls "open perfectionism", which "recognizes that there are conceptions of the good that could be significantly furthered with the use of state resources and that the potential for a majority to live a certain way can, subject to certain constraints, *justify* such use".[8]

Of course, claiming that that perfectionist policies are approved of, permissible, desirable, or justified is strictly speaking compatible with the deontic perfectionist's more ambitious further claim that such policies morally required. But the point is that the kinds of claims made in the previous paragraph fall short of establishing perfectionist duties and are not supplemented and strengthened by further arguments that would establish any such duties, and thus these authors are most plausibly read as endorsing only

[3] Sher, G., *Beyond Neutrality: Perfectionism and Politics* (Cambridge: Cambridge University Press, 1997), p. 1 (emphasis added); Sher, G., "Confessions of a Quidnunc", *The American Journal of Jurisprudence* 63 (2018), pp. 60-1 (emphases added).
[4] Hurka, T., *Perfectionism* (Oxford: Oxford University Press, 1993), p. 159.
[5] Chan, J., "Legitimacy, Unanimity, and Perfectionism", *Philosophy & Public Affairs* 29 (2000), p. 35 (emphasis added).
[6] Ibid., p. 42.
[7] Couto, A., *Liberal Perfectionism* (Berlin: De Gruyter, 2014), p. 33 (emphasis added).
[8] Metz, T., "Respect for Persons and Perfectionist Politics", *Philosophy & Public Affairs* 30 (2001), p. 442 (emphasis added).

non-deontic perfectionism. One reason for this lowered sense of ambition is, I suspect, dialectical. That is to say, it is neutralists, and in particular political liberals, who have set the tone of the neutrality-perfectionism debate and who have defined that debate as one about whether it is permissible for the state to act in pursuit of perfectionist ideals. Quong, for instance, follows Rawls in formulating the neutrality-perfectionism debate as a debate about "what liberal states can permissibly or legitimately do", and in particular about whether "it is permissible for a liberal state to promote or discourage some activities, ideals, or ways of life on grounds relating to their inherent or intrinsic value, or on the basis of other metaphysical claims".[9] This formulation of the debate has, I suspect, put perfectionists on the back foot, and so most perfectionists have been content to show merely that perfectionist policies are legitimate — that they are not *de*-legitimizing — without going the extra step of arguing that perfectionist policies are in fact a necessary condition of state legitimacy or otherwise morally required.

In my view, perfectionists should set their sights higher. They should not be content to show merely that perfectionism is legitimate, or to show merely that there is no general exclusionary neutrality principle that forbids appeal to considerations about the good life and human flourishing when justifying political arrangements. Rather, they should, I think, defend the more ambitious deontic form of perfectionism, according to which states are not only permitted but also morally required to promote the good life. As Wall — one of the few perfectionists who argues for this stronger deontic claim — puts it, "political authorities *should* take an active role in creating and maintaining social conditions that best enable their subjects to lead valuable and worthwhile lives".[10]

2 The case for deontic perfectionism

In this section, I make a case for deontic perfectionism. In Sections 2.1 and 2.2, I consider and reject two recent arguments from Wall, who seeks to establish that perfectionist duties derive from the state's general duties to promote self-

[9] Quong, J., *Liberalism Without Perfection*, pp. 15-6.
[10] Wall, S., *Liberalism, Perfectionism and Restraint* (Cambridge: Cambridge University Press, 1998), p. 8 (emphasis added). The other perfectionist who defends the stronger deontic thesis is Raz: see his *The Morality of Freedom* (Oxford: Oxford University Press, 1986), esp. pp. 407-20.

respect and fairness. I then argue, in Section 2.3, that deontic perfectionism is grounded in the right of citizens to the conditions for a flourishing life.

2.1 From self-respect to deontic perfectionism

In a recent paper, Wall briefly argues that the state's perfectionist duties fall out of its general duties to secure and promote the conditions for the warranted self-respect of its citizens.[11] The most interesting and novel premise of Wall's argument is this: "to have a sense of self-respect, and to merit it, a person must be committed to pursuing a sound conception of the good and must care about his character".[12] Thus any state that has a duty to secure the conditions necessary for citizens' warranted self-respect will need — in a perfectionist vein — to ensure that citizens are committed to pursuing a sound conception of the good. If this tight conceptual connection between warranted self-respect and pursuit of a sound conception of the good does indeed exist, then this would generate a dialectically effective argument for deontic perfectionism because even Rawlsian neutralists generally accept that states have a duty to secure and promote the conditions necessary for citizens' possession of a warranted sense of self-respect.[13]

But when Wall seeks to motivate the idea that warranted self-respect requires commitment to a sound conception of the good, he gives, I believe, the wrong kind of case. Wall invites us to consider an individual, Charlie, who "critically evaluate[s] his conception of the good" and "comes to see that it is misguided".[14] As Walls plausibly says, Charlie must now renounce his commitment to his conception of the good. Charlie, that is, could not simultaneously enjoy warranted self-respect and judge as unsound the conception of the good life to which he is committed: "in being indifferent to

[11] Wall, S., "Enforcing Morality", *Criminal Law and Philosophy* 7 (2013), pp. 455-71, at pp. 460-5. Kramer's "aspirational perfectionism" also grounds perfectionism on the Rawlsian notion of self-respect. See Kramer, M., *Liberalism with Excellence* (Oxford: Oxford University Press, 2017).
[12] Ibid., p. 461.
[13] Of course, for those who are not already devotees of Rawlsianism, the idea that states have a duty to promote the necessary conditions for citizens to harbour a warranted sense of self-respect may seem far more contrived, and decidedly less intuitive, than the idea that states should help citizens to lead good, flourishing lives. So, unless more is said about the source of the duty to promote the conditions for warranted self-respect, the argument from self-respect to deontic perfectionism risks justifying the more plausible by reference to the less plausible. (I thank Jeff McMahan for this point.) But here I grant the existence of a duty to promote the conditions for warranted self-respect.
[14] Wall, S., "Enforcing Morality", p. 463.

whether or not he, in fact, lives well, Charlie would thereby show that he did not really respect himself".[15] Any self-respect in this case would, says Wall, be "deluded", not "warranted".[16] Thus, warranted self-respect "rests on the judgement or conviction that one's plan of life or conception of the good is worth carrying out".[17]

The problem with the Charlie case is that it does not support the claim that warranted self-respect requires commitment to a conception of the good life that is sound. What it supports is the claim that warranted self-respect requires commitment to a conception of the good life that one *believes* is sound. So, assuming that the state has a duty to promote warranted self-respect, the Charlie case could only support a kind of subjective perfectionism (the state has a duty to promote conditions in which citizens will be committed to conceptions of the good that they believe to be sound), not the objective perfectionism (the state has a duty to promote conditions in which citizens will be committed to conceptions of the good that are *genuinely* sound) that Wall intends.

To test whether warranted self-respect requires commitment to a sound conception of the good, we thus need to consider an individual who justifiably believes that his conception of the good is sound but who is simply mistaken about this. So consider Dan, an activist whose life is structured around campaigning for the reintroduction of the death penalty in England. He has been convinced, we can suppose, by Kramer's "purgative rationale" according to which certain crimes are so extravagantly and defilingly evil that the continuation of the evildoer's life would be a blot on the moral character of the community to which he belongs and would constitute an affront to humanity, and thus the death penalty is called for to purge the community of that taint.[18] Dan is admired amongst advocates of the death penalty, and he is often invited to give talks about the ethics of capital punishment. But suppose also, ex hypothesi, that the death penalty is always impermissible, and thus that Dan's pro-death-penalty stance is mistaken. So Dan is committed to an unsound conception of the good (despite being committed to a conception of the good the good that he *judges* to be sound).

---

[15] Ibid.
[16] Ibid.
[17] Ibid., p. 462.
[18] See Kramer, M., *The Ethics of Capital Punishment: A Philosophical Investigation of Evil and Its Consequences* (Oxford: Oxford University Press, 2011).

The question, then, is this: Must Dan — who is committed to a reasonable but ultimately unsound conception of the good — lack warranted self-respect? If he does respect himself, must he be "deluded" in doing so? It is difficult to see why. Dan certainly seems to fit the profile of a warrantedly self-respecting individual: he cares about his beliefs and thinks they matter enough to be worth subjecting to critical scrutiny, he confidently protests behavior that he reasonably and justifiably (if mistakenly) takes to be wrongful, he feels a sense of pride and accomplishment when his activism bears fruit, and so on. Of course, we might think that Dan's self-respect is less *valuable* than it would be if it were premised on a sound conception of the good. But that is no reason for thinking that his sense of self-respect is unwarranted or deluded. In general, then, it seems that the objective soundness of one's conception of the good is an inordinately strict requirement for having a warranted sense of self-respect.[19]

Let me briefly defend this argument against two objections. First, is it not possible to specify Dan's conception of the good as *being an effective campaigner for positive political change* rather than as *campaigning to get the death penalty reintroduced*? In this case, the unsoundness of his views about the death penalty would not suffice to render unsound his (more abstractly stated) conception of the good, and so this would no longer be a counterexample in which an individual has warranted self-respect despite having an unsound conception of the good.[20]

Tinkering with the level of abstraction at which we state conceptions of the good does not rescue Wall's argument, however. It merely moves the bump in the rug. Thus suppose ex hypothesi that Buddhists and other apolitical religions are right: we should reject the power struggles, the passions and the divisive activism of earthly politics in favour of a life of spiritual contemplation. Would the objective unsoundness of the life of politics mean that Dan (who,

[19] It is worth adding here that, even if Dan does not qualify as having warranted self-respect, it is unclear how steering him, in Wall's preferred perfectionist way, into having sound views about capital punishment could help him to enjoy warranted self-respect. Would Dan really have enhanced grounds for self-respect if the only reason he had the right view about capital punishment was that he had been "nudged" by the government? There is thus a gap between establishing that warranted self-respect requires commitment to a sound conception of the good, and establishing that states (who have a duty to secure warranted self-respect) should promote sound conceptions of the good because efforts to promote the good (or at least certain kinds of efforts to promote the good) might themselves compromise the warrants for self-respect. But, as with the concession in footnote 13, I will not press this point further.

[20] I thank David Miller for putting this objection to me.

recall, confidently protests behavior that he sees as wrongful, feels pride in his accomplishments, and so on) must be deluded to respect himself? Intuitively, it is unclear why he must.

Second, Wall might argue that the case of Dan only serves to demonstrate that one can be self-respecting whilst having an unsound conception of the good; it does not demonstrate that one can be *warrantedly* self-respecting whilst having an unsound conception of the good. Insofar as having an objectively sound conception of the good is *defined* into the notion of warrantedness, cases such as Dan's, which involve warrantedly self-respecting individuals who subscribe to unsound conceptions of the good, are ruled out conceptually.

However, if warrantedness rather than self-respect carries too much of the argumentative load here, the move from warranted self-respect to deontic perfectionism quickly becomes a tautology — a matter of conceptual fiat rather than an intuitively persuasive argument. Of course, warrantedness can play *some* role in Wall's argument: it is plausible that someone who is systematically deluded (e.g. someone in Nozick's "experience machine" who takes pride in illusory accomplishments), or who is guided by a deeply objectionable ideology (e.g. a racist who takes great pride and satisfaction in his racist behaviour), is not warrantedly self-respecting, however many of the hallmarks of self-respect he might possess. But if warrantedness does too much moral work beyond this, much of the intuitive force of the value of warranted self-respect is lost. In such a case, those who are not already convinced by the idea of perfectionist duties are likely to respond that, while they care about warranted self-respect in *some* sense (in particular, where warrantedness plays only a fairly minimal role), they simply do not care about *this* specific account of warranted self-respect — an account that feels very custom-built, even ad hoc, and that functions like a kind of intellectual Trojan Horse.

2.2 From fairness to deontic perfectionism

In another recent discussion, Wall has advanced a brief yet interesting argument that the state's duties to promote the good life derive from its general

duties of fairness.[21] Wall first argues that "fairness sometimes requires that, holding other things constant, we should favor those who are less well off over those who are better off".[22] Wall then argues that "moral environments that offer their inhabitants bad options, when compared to available alternatives, will tend to disadvantage those who are less well off".[23] As Wall also puts this second premise: "the failure to engage in [perfectionist politics] comes at the expense of those who fare least well".[24] So since fairness requires priority for the worst off, and since the worst off are particularly disadvantaged by liberal neutrality, fairness requires perfectionist policies and interventions.

Libertarians will want to reject Wall's first premise — namely, that fairness requires priority for the worst off. But my worry is with his second premise: that the worst off are particularly disadvantaged by neutrality (and so are particularly advantaged by perfectionism). To be clear: I believe that neutrality does disadvantage citizens by cutting them off from important goods. What I am questioning is Wall's claim that neutrality is *differentially* disadvantageous to the worst off, such that it would raise concerns about fairness.

So what reason does Wall give for thinking that a society that does not enact perfectionist policies would be particularly disadvantageous to the worst off? His main reason seems to be that those who are worse off are worse decision-makers and thus are more vulnerable to bad options. He says: "Naturally, and speaking generally, those who are bad decision-makers will tend to be less well off overall than those who are good decision-makers".[25] Thus a society that conducts various state-run lotteries, for instance, "will disadvantage those who are not good at making judgements of probability, or have trouble resisting the impulse to gamble", as compared to those who are able to see through such schemes.[26] A perfectionist politics, which removes this option, will thus benefit the worse off (who are worse decision-makers) more than the better off (who are better decision-makers). Wall also gives the example of heroin abuse. The existence of this option disadvantages the worse off and so the family of a

[21] Wall, S., "Moral Environmentalism", in C. Coons and M. Weber, *Paternalism: Theory and Practice* (Cambridge: Cambridge University Press, 2013), pp. 93-114, at pp. 98-100.
[22] Ibid., p. 99.
[23] Ibid.
[24] Ibid.
[25] Ibid.
[26] Ibid.

heroin addict might reasonably question how fair it is that "the moral environment in which he lived made it easier than it need be for people like him to ruin their lives".[27]

Wall may well be correct that the worse off are disproportionately disadvantaged by the existence of worthless options such as drugs and lotteries.[28] But, in general, perfectionism does not merely involve the *removal of worthless options*; a major part of the perfectionist programme is the *provision of worthwhile options*. By Wall's own logic, it would seem that the better off (insofar as they are better decision-makers) will be more likely to take advantage of these options, whereas the worse off (insofar as they are worse decision-makers) will be more likely to ignore them. Consider, for instance, Quong's list of standard perfectionist policies: "performance art, art galleries, public parks, works of literature, sights of cultural significance, education programmes for adults, and athletic events".[29] On the face of it, these kinds of options seem to cater to middle-class pastimes and are likely to be most advantageous to the better off.

So once a broad range of perfectionist policies is on the table — both policies that remove worthless options and policies that provide worthwhile options — it is not clear why we should think that any particular segment of society would stand to gain more or less from perfectionism than any other. Some perfectionist policies will likely be of most benefit to the worse off; others will likely be of most benefit to the better off. How do all these benefits wash out? It is difficult, if not impossible, to say from the armchair. What is needed here is some robust, empirically informed sociological model that explains why we should expect the worse off to be greater beneficiaries of perfectionist policies.

---

[27] Ibid., p. 100.

[28] Note, though, that the better off *also* make catastrophically bad decisions about how to lead a valuable life under liberal neutrality. Perhaps the better off do not waste their money on scratch cards and state-run lotteries; but on the other hand they may throw money around like confetti in luxury casinos. So, if anything, it may be that it is the *well off* who suffer most by the state's failure to engage in moral environmentalism, because the well off are more likely to lead lives of complacency, self-indulgence and frivolity. (Here Wall may say that the frivolous rich, while being *materially* well off, are not *actually* well off — well off, that is, on the right conception of well-being. Since the frivolous rich may in fact be amongst the worst off, this case may support, rather than refute, the claim that anti-perfectionism comes at the expense of the worst off. [I thank Steven Wall for putting this reply to me.] But, if well-being is construed in less material terms, then the first premise [which calls for priority for the worst off] becomes shakier. Do frivolous millionaires really have a fairness-based claim that government policies give their interests special priority?)

[29] Quong, J., *Liberalism Without Perfection*, p. 89.

In the absence of this, the argument from fairness to deontic perfectionism does not go through.

2.3 A right to a flourishing life

In my view, perfectionist duties follow not from duties of self-respect or fairness but rather from the right of all citizens to lead a flourishing life. This moral right correlates with the government's perfectionist duty to promote the good life.

A clarification is immediately in order: although the notion of a "right to a flourishing life" is a useful shorthand, what is really being guaranteed by this right is not a flourishing life *itself*. One important reason for this is that a right to lead a flourishing life would not be responsibility-sensitive. Such a view would be incompatible with the idea that individuals should take personal responsibility for their conception of the good life and should bear the costs and burdens associated with its pursuit — or, put differently, such a view would end up indulging some citizens' expensive tastes. Instead, then, what the right to a flourishing life guarantees is the *fair conditions* for living a flourishing life, understood to be the fair set of opportunities, resources, public culture and other conditions that supports and encourages the flourishing of citizens generally. This view is responsibility-sensitive because what citizens choose to do once these fair conditions have been provided is up to them.

At this point, however, deontic perfectionism might be thought to face a serious objection. For liberal neutralists can ask: In what sense does the neutral liberal state not already secure the fair conditions for human flourishing? After all, a Rawlsian well-ordered liberal society would eradicate the kind of poverty that prevents many individuals from devoting their time and energies towards the best things in life; it would encourage the development and exercise of the two moral powers, one of which is a capacity to rationally form, revise and pursue a conception of the good; it would provide space for individuals to interact and form associations aimed at various forms of excellence; it would combat non-material obstacles to flourishing such as propaganda and discriminatory speech. What, liberal neutralists might ask, is involved in providing the fair conditions for human flourishing above and beyond this?

Once each citizen has been given her fair share of social and economic goods, why hasn't she also been given the fair conditions necessary for her flourish? What does deontic perfectionism do that would not already be done by liberal neutrality?[30] Indeed, Quong rejects deontic perfectionism (though not under that name) and focuses almost entirely on non-deontic perfectionism precisely on the grounds that deontic perfectionism cannot "practically distinguish itself from non-perfectionist theories of distributive justice, such as Rawls' or Ronald Dworkin's theory."[31]

For deontic perfectionists, however, the prescriptions of Rawls, Dworkin and other liberal neutralists fall short of providing fair conditions for flourishing because perfectionist considerations strongly influence one's conception of social or distributive justice. The notion of "fair shares", "fair terms" and "fair conditions" are deeply (indeed, constitutively) shaped by specific claims about human flourishing. Perfectionist considerations, in other words, *define* what constitutes fair shares and fair conditions. To ask what more could be required over and above the establishment of the Rawlsian fair terms of social cooperation is thus to miss the point because, for the deontic perfectionist, what counts as fair terms and what counts as the rightful treatment of the citizenry by the state is precisely what is at stake here.[32]

To motivate and illustrate this move, let us consider a specific example of deontic perfectionism. Suppose that moral, intellectual and artistic excellence are components of the good life. Then, for deontic perfectionists of this kind, the natural and plausible thing to say is that social conditions cannot be fair unless they foster moral, intellectual and artistic excellence. For such deontic perfectionists, then, providing the fair conditions of flourishing would call for more than the liberal neutralist's efforts to secure association freedoms, to eradicate poverty and to combat discriminatory speech: it would also require

[30] I thank Max Afnan and Jonathan Quong for pressing me on this point.
[31] Quong, J., *Liberalism Without Perfection*, p. 122.
[32] I thank Jonathan Quong and Steven Wall for helping me to formulate this response. There is another available response to the indistinguishability objection, which is that even if deontic perfectionism is *extensionally* equivalent to the leading non-perfectionist theories of justice, this does not show that deontic perfectionism is redundant or uninteresting. After all, deontic perfectionism might give a more attractive underlying picture of political morality, one that justifies a range of policies in a more natural and plausible way. But since I believe there will be genuine divergences at the bottom line between deontic perfectionism and non-perfectionist theories of justice, I shall not press this point further. For further discussion of indistinguishability, see Section 4 of Chapter 3.

the cultivation and promotion of moral, intellectual and artistic excellence. These perfectionist ideals could be expected to heavily shape policy in a large range of areas including "public assistance, educational policy, the criminal and civil justice system, the prison system, city planning and land use, transportation policy, the tax code, support for cultural institutions, regulation of the entertainment industry, investment incentives, and the structure of institutions such as the military", to employ Sher's helpful list.[33] Even more concretely, and to give just one example: this kind of deontic perfectionism might plausibly seek to regulate the profusion of advertisements that direct energies away from various forms of excellence and towards the accumulation of material possessions. On the deontic perfectionist view, then, fair conditions would go considerably beyond those that would be established within a well-ordered liberal-neutral society.

Without these kinds of conditions — conditions that are shot through with perfectionist values — it is not, of course, *impossible* for citizens to realize moral, intellectual and artistic excellence. Under liberal neutrality, citizens clearly still have some chance of attaining moral, intellectual and artistic excellence. Nevertheless, as far as this form of deontic perfectionism is concerned, without these kinds of conditions citizens are being denied a *fair* chance to lead a flourishing life and so they have a legitimate complaint against the state. In this specific sense, and to this extent, deontic perfectionism entails that even a well-ordered, perfectly realized Rawlsian state fails to respect some of the rights of its citizens.

None of this, yet, is to say that deontic perfectionism is in the clear. Perhaps deontic perfectionism, insofar as it appeals to controversial claims about excellence, fails to take seriously the reasonable pluralism within a free society and so contravenes the public justification test — an objection I consider in Section 4. Or perhaps there are no perfectionist duties because there is no such thing as a right to the fair conditions for a flourishing life — an objection I consider in a moment. The point is just that the move from non-deontic to deontic perfectionism need not involve the abandonment of perfectionism: by building perfectionist content into one's conception of fairness and of fair terms, as deontic perfectionists will find it natural to do, it is possible to

[33] Sher, G., *Beyond Neutrality*, p. 246.

overcome Quong's charge that a deontic form of perfectionism aimed at the provision of fair conditions for flourishing would no longer be distinguishable from liberal neutrality.

A crucial claim here, of course, is that there exists a right to the fair conditions of human flourishing — or, as I will put it elliptically, a right to a flourishing life. But while many are willing to accept the idea of a right to a "minimally decent life", it is not generally thought that there is a right to a *flourishing* life. Griffin, for instance, urges that human rights should be understood as "protections of that somewhat austere state, a characteristically human life, not of a happy or perfected or flourishing human life".[34] Similarly, Nickel says: "human rights are not ideals of the good life for humans; they are rather concerned with ensuring the conditions, negative and positive, of a minimally good life".[35] And although I am positing a *moral* right to a flourishing life, and not a *human* right (where human rights are typically viewed as a more stringent subset of moral rights), Griffin, Nickel and others are likely to object on these sorts of grounds to the suggestion that there is even a moral right to a flourishing life. So what justifies the claim that there is a right to a flourishing life?

Defending any particular approach to the justification of rights lies beyond the scope of this paper. So I shall adopt an ecumenical strategy: I shall assume two prominent approaches and show that, on either of these approaches, it plausibly follows that there exists a right to a flourishing life.

First, then, let us consider whether a Rawlsian original position heuristic might yield a right to a flourishing life. Here we determine the system of rights and duties by considering what social terms would be chosen by rational and reasonable contractors situated behind a veil of ignorance. In particular, we want to know whether parties in the original position would select social terms that are deeply shaped by perfectionist considerations. In *A Theory of Justice*, Rawls discusses some proposals along these lines.[36] He first considers an extreme Nietzschean form of perfectionism, according to which rights and

[34] Griffin, J., *On Human Rights* (Oxford: Oxford University Press, 2008), p. 34.
[35] Nickel, J., *Making Sense of Human Rights* (Oxford: Blackwell, 2007), p. 138.
[36] See Rawls, J., *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971), pp. 325-332.

duties are defined in such a way as to "maximize the achievement of human excellence in art, science, and culture".[37] As Rawls rightly notes, this would be rejected by parties in the original position because (inter alia) it "might lead to a lesser religious or other liberty, if not to a loss of freedom altogether to advance many of one's spiritual ends".[38]

But this still leaves open the more moderate view of the kind I sketched above, according to which perfectionist considerations such as moral, intellectual and artistic excellence can play a role in shaping the terms of society *conditional* on the other lexically prior principles of justice being met. In addition, that is, to selecting social terms that guarantee equal basic rights and liberties, that require positions and offices to be allocated on the basis of fair equality of opportunity, and that give special distributive priority to the least advantaged members of society, why would parties in the original position not *also* think it reasonable and prudent to select social conditions that are conducive to their moral, intellectual and artistic excellence? Surely a party in the original position would no more want to gamble with his flourishing than he would want to "gamble with his liberties"?[39] Of course, a contractor might not want to see the state promotion of a highly specific conception of the good such as Christianity, as this would put in jeopardy her religious freedom if she turns out to be a non-Christian. But this hardly rules out a more moderate perfectionism of the kind I sketch above, one which seeks to promote the flourishing of the citizenry by establishing social conditions that are determined by and conducive to generic perfectionist ideals such as moral, intellectual and artistic excellence, all done in a way that is respectful of basic civil, political and economic rights. Rawls foregoes "any further advantages that might be won" by going beyond the two principles of justice because "the hardship if things turn out badly [may be] intolerable": but it is difficult to see who could reasonably find it intolerable to live under a system that, in addition to upholding Rawls's two principles of justice, seeks to foster social conditions that encourage the citizens to lead flourishing lives.[40]

[37] Ibid., p. 325.
[38] Ibid., p. 327.
[39] Ibid., p. 177.
[40] Ibid., p. 175.

Rawls is aware of this more moderate kind of view and admits that it is "not easy to argue against".[41] In an interesting anticipation of his later work, Rawls's main worry seems to be about the disagreement and indeterminacy surrounding perfectionist considerations. "Criteria of excellence", says Rawls, "are imprecise as political principles, and their application to public questions is bound to be unsettled and idiosyncratic".[42] Giving the example of debates over sexual morality, Rawls argues that since "uncertainties plague perfectionist criteria", since we often appeal to perfectionist considerations "in an ad hoc manner", and since "in these matters we are likely to be influenced by subtle aesthetic preferences and personal feelings of propriety", perfectionism cannot constitute part of "a feasible basis of social justice".[43]

The first thing we should note here, however, is that (as Rawls himself admits) it is impossible to make highly precise public-policy judgements and measurements about *neutral* values such as freedom and equality: our judgements "[need] not be very exact" and only need to be "accurate enough to guide the main decisions" concerning the provision of fair social terms.[44] So the question is: can we make *sufficiently* determinate assessments about the good life to guide certain important political decisions?

I address in Section 4 the charge that perfectionist considerations are too controversial or indeterminate to play a role in politics. But, in brief, my view is that, contrary to what political liberals often suggest, perfectionist ideals such as moral, intellectual and artistic excellence are not radically indeterminate. They function in roughly the same way as neutral ideals such as freedom and fairness insofar as there is a fairly stable set of core or paradigmatic cases of moral, intellectual and artistic excellence as well as a range of borderline cases in which it is unclear whether or not these ideals are exemplified. So while Rawls may well be correct about the specific example he gives — judgements about sexual morality are too idiosyncratic to be a feasible basis of social justice — I am less convinced that this is true of perfectionist judgements generally. I am less convinced, that is, that appeal to a core handful of perfectionist axioms that have broad appeal and that are pitched at a high level of generality and

[41] Ibid., p. 330.
[42] Ibid.
[43] Ibid., p. 331.
[44] Ibid., p. 327.

abstraction (such as moral, intellectual and artistic excellence) would be too "ad hoc", "idiosyncratic" or "uncertain" to feasibly be part of social justice. Surely it is possible to reliably judge both that these are valuable ideals and that, say, a windfall for museums, art galleries and theatre companies could be expected to advance these ideals amongst the citizenry? Surely, too, it is possible to reliably judge a life wholly consumed by drug addiction as being degraded or unworthy? Indeed, as Rawls himself earlier concedes, "clearly there are standards in the arts and sciences for appraising creative efforts, at least within particular styles and traditions of thought", and so "comparisons of intrinsic value can obviously be made".[45] "They are not necessarily so vague," Rawls continues, "that they must fail as a workable basis for assigning rights".[46] It thus remains unclear why the original position would not yield a system of rights and duties that is shot through with perfectionist considerations.

A second prominent approach to rights is the interest theory, according to which rights are grounded in important interests.[47] To show that we have a right to a flourishing life, then, I need to show that flourishing is an important interest. It is difficult to show this in a precise way because there are no strict criteria for determining whether or not something is an important human interest. Raz, for instance, says that "'X has a right' if and only if X can have rights, and, other things being equal, an aspect of X's well-being (his interest) is a sufficient reason for holding some other person(s) to be under a duty".[48] But why shouldn't a citizen's interest in living a flourishing life — in pursuing moral, intellectual and artistic excellence — be a sufficient reason for holding other citizens under a duty to contribute towards the cultivation and maintenance of the social and political conditions conducive to such flourishing? Similarly, Tasioulas argues that "a right exists if an individual's interest in the object of the putative right…has the requisite sort of importance to justify the imposition of duties on others variously to respect, protect or advance that interest by securing to that individual the object of his right".[49] But, again, why shouldn't each citizen's interest in flourishing be important

[45] Ibid., p. 328.
[46] Ibid.
[47] See, e.g., Raz, J., *The Morality of Freedom*; Kramer, M., "Rights Without Trimmings" in M. Kramer, N. Simmonds, H. Steiner, *A Debate Over Rights* (Oxford: Oxford University Press, 1998), pp. 7-111.
[48] Raz, J., *The Morality of Freedom*, p. 166.
[49] Tasioulas, J., "On the Foundations of Human Rights" in R. Cruft, M. Liao and M. Renzo (eds), *Philosophical Foundations of Human Rights* (Oxford: Oxford University Press, 2015), p. 50.

enough to justify the imposition of duties on other citizens and on the state to respect, protect or advance such flourishing? What could be more important than individual flourishing?

We state the idea behind these rhetorical questions more precisely by distinguishing between two ways of understanding the notion of an interest.[50] The first view holds that rights depend on *subjectively important interests* — that is, on preferences that particular individuals care about deeply. On this view, to show that there is a right to a flourishing life one would need to show that people in general have a deeply held desire to flourish. This strikes me as a plausible assumption, at least in the context of ideal theory. After all, few people — again, in the ideal theory context — aspire to an *adequate* life: to only minimal education, basic literacy, and a morally mediocre character. Most set their sights higher. They have a deeply held desire to excel and flourish. Not all people desire to write the Great American Novel, of course. But the desire for flourishing need not be understood in this highly elite sense. Individuals also exhibit a desire for excellence and for flourishing when they desire to play a musical instrument well, to run a half marathon, to struggle admirably in the face of adversity, to be a kind and loyal spouse, to be a loving parent, to have a morally commendable character, to succeed professionally, to gain an understanding of the science of their day, and so on. In all these ways, it generally matters greatly to people that they live not just "adequately" or "minimally decently", but excellently, worthily, and flourishingly.

Of course, there are well-known problems with construing interests in terms of the actual or idealized desires of individuals. It does not, for instance, seem correct to say that an individual who devotes her life to some pointless activity such as counting blades of grass is acting in her interests, however ardently and deeply held this desire is.[51] Many have thus been tempted to understand interests in more objective terms. On this second view, there are certain goods, values and pursuits — such as freedom, rationality or physical health — that are in our interests, irrespective of our beliefs and desires. Thus, even though it might be subjectively important for an eccentric to count blades of grass all day every day, or for a drug addict to score his next fix, these pursuits might not be

[50] See Swanton, C., "The Concept of Interests", *Political Theory* 8 (1980), pp. 83-101.
[51] See Rawls, J., *A Theory of Justice*, p. 432.

in the individuals' interests objectively understood. On this account of interests, to show that there is a right to a flourishing life one would need to show that flourishing is in citizens' objective interests. This strikes me as even more plausible than the claim that citizens have subjectively important interests in flourishing. Indeed, the argument here is quasi-analytic. For when one thinks about the kinds of goods, activities and ideals that standardly feature on objectivists' lists of the attitude-independent objective goods — friendship, community, knowledge, wisdom, rationality, virtue, autonomy, appreciation of true beauty, and so on — these are precisely the kind of goods that perfectionists take to be constitutive components of human flourishing and that perfectionists seek to promote through politics. Whether interests are understood subjectively or objectively, then, it seems that individuals have an important interest in living a flourishing life.

3 The advantages of going deontic

The stronger deontic form of perfectionism has the advantage of giving perfectionists the resources to block a number of powerful recent objections. I shall illustrate this by examining three recent objections which allege respectively that perfectionism is illegitimate, that it is paternalistic, and that it manifests a meddlesome "quidnunc" mentality.[52] All of these objections, however, presuppose there are no perfectionist duties — a fair assumption in relation to mainstream non-deontic forms of perfectionism, but one which flatly begs the question against deontic perfectionists. Insofar as these objections trade on the absence of duties to promote the good life, then, deontic perfectionism is better placed to rebut them than are non-deontic forms of perfectionism which recognize only reasons, justifications and permissions to promote the good life.

3.1 Legitimacy

In his *Liberalism without Perfection,* Quong argues that perfectionist states lack legitimacy.[53] To make this argument, he notes that the legitimacy of

[52] Quong's autonomy-based objection to perfectionism also presupposes that there are no perfectionist duties and as such would have no force against deontic perfectionism: see Quong, J., *Liberalism Without Perfection*, pp. 45-72; Kramer, M., *Liberalism with Excellence*, pp. 46-63.
[53] Quong, J., *Liberalism Without Perfection*, pp. 108-36.

perfectionist states is generally defended by reference to "practical reason" theories of legitimacy. According to these theories — of which Raz's "service conception" is the foremost example — states are legitimate when they enable us to best comply with the reasons that apply to us. A surgeon who comes across a serious accident, for instance, has legitimate authority to issue commands to a bystander with no medical expertise because that bystander will best comply with the reasons that apply to her by following the surgeon's commands rather than by deciding for herself how to assist the victim.[54] On practical reason views, then, perfectionist states are legitimate because they enable citizens to flourish and thus to comply with the full range of practical reasons that apply to them.

But, says Quong, these practical reason models of legitimacy all suffer from a common problem: namely, that the mere fact that living by X's rules will help you to live well and to comply with the dictates of reason does not suffice to show that X has legitimate authority over you. We do not, for instance, think that just because a certain Peruvian tour company is superb in every way — so good, in fact, that what you have most reason to do is to become its customer — that it thereby gains legitimate authority over you and may force you to sign up to its tours when you visit Peru.[55] The "fundamental problem" with the practical reason accounts of legitimacy that have tended to underwrite perfectionist politics, then, is that they fail to explain "why the brute fact that *I* have reason to do something should affect what rights *you* have with regard to me".[56]

In light of this, Quong develops his own preferred "natural-duty" account of legitimacy, according to which legitimate authority depends on *rights and duties, not reasons*.[57] As Quong puts this view, "legitimate authority does not track what we have reason to do, it tracks what rights we have, and what duties we may be under".[58] This view explains the Peruvian tour company case: the tour company has no authority over me because they have no *right* to my

[54] Ibid., pp. 112-3.
[55] Ibid., pp. 108-9.
[56] Ibid., p. 115.
[57] For other defences of the natural-duty approach, see, e.g., Waldron, J., "Special Ties and Natural Duties", *Philosophy & Public Affairs* 22 (1993), pp. 22-7; Wellman, C., "Liberalism, Samaritarianism, and Political Legitimacy", *Philosophy & Public Affairs* 25 (1996), pp. 211-37.
[58] Quong, J., *Liberalism Without Perfection*, p. 116.

custom and because I have no corresponding *duty* to use their services (even though I do have *reason* to do so). And the natural-duty account also offers an alternative explanation of the surgeon case: the surgeon has legitimate authority over me because I am under a positive duty to aid and can best fulfill this duty by accepting the surgeon's commands as authoritative (and *not*, as Raz would say, because by following the surgeon's commands I best comply with the reasons that apply to me). Of course, this is not to say that talk of rights and duties has no connection to talk of reasons. Indeed, what we have most reason to do and what we are have a duty to do will often coincide. But, as the Peruvian tour case illustrates, these can come apart because rights and duties are determined by a set of reasons that is narrower than the full set of practical reasons, which includes friendship, humour, wisdom, and many other values.[59]

In short, then, legitimate authority depends on the ability of states to "enable each of their citizens to better fulfill the duties they are under", and not on their ability to help us lead the kind of flourishing lives that we have most reason to lead.[60] This is why, for Quong, neutral states, whose policies aim at ensuring citizens fulfill their duties of justice to others, are legitimate, whereas perfectionist states, whose policies aim at ensuring citizens live flourishing lives in accordance with the full set of practical reasons that apply to them, are not legitimate.

However, this objection assumes that there are no perfectionist duties. Quong's argument thus begs the question against deontic perfectionists. After all, if, as I have argued, each citizen has a right to flourish, and if the state and other citizens therefore have correlative duties to promote such flourishing, then the legitimacy of the perfectionist state could be grounded in these perfectionist duties. In such a case, it could be said not just of neutral states but also of perfectionist states that they possess legitimacy in virtue of planning and coordinating individual actions in ways that "enable each of their citizens to better fulfill the [perfectionist] duties they are under". Quong's preferred natural-duty account of legitimacy would, that is, apply equally well to perfectionism as to neutralist liberalism. It is important to note that Quong is

[59] Ibid., p. 119.
[60] Ibid., p. 131.

aware of this possibility and does consider a perfectionist response along these lines — though he is not ultimately convinced for reasons I discuss above.[61] But, for our purposes here, the crucial point is that the cogency of the legitimacy objection to perfectionism entirely hinges on, and reduces to, the claim that there are no perfectionist duties.

3.2 Paternalism

Quong also contends that perfectionism is objectionably paternalistic. To show this, he first argues for a "judgemental" account of paternalism according to which A acts paternalistically when she attempts to improve the welfare, good, interests or values of B, and when her act is motivated by a negative judgement about B's ability to make the decision that would effectively advance B's welfare, good, interests or values.[62] This account of paternalism, says Quong, "captures our sense that to treat someone paternalistically is to treat that person like a child in the specific sense of acting in that person's best interests because you believe, in this situation, the person lacks the ability to do so himself or herself".[63] And this account explains the prima facie wrongness of paternalism. For paternalism conflicts with the conception of moral status central to liberal political philosophy, according to which citizens are free and equal in virtue of their possession of two moral powers, namely the capacity for a sense of justice and the capacity to form, to revise and rationally to pursue a conception of one's good. Thus to treat someone paternalistically is prima facie wrong because it involves "treat[ing] that person as if he or she lacks the second moral power".[64]

Quong then seeks to demonstrate the paternalism of perfectionism by arguing that perfectionists have no satisfactory answer to a "simple but important" question: "Why is state action necessary at all to achieve perfectionist objectives?"[65] "Why not simply give each citizen their fair share of resources and let them make their own decisions [about how to lead a flourishing life]?"[66]

---

[61] Ibid., pp. 120-6. See text at footnote 31.
[62] See ibid., p. 80.
[63] Ibid., p. 81.
[64] Ibid., p. 101.
[65] Ibid., p. 74.
[66] Quong, J., "Liberalism Without Perfection: A Précis", *Philosophy and Public Issues* 2 (2012), p. 2.

Quong considers various answers given by perfectionists to this question and finds nearly all to involve a negative, infantilizing judgement about citizens' capacity to run their own lives, thus revealing "perfectionism's true paternalist colours".[67] For instance, some perfectionists have argued that the realization of ideals of the good life requires political action because, when left to their own devices, people act *irrationally* and fall prey to various cognitive biases that lead them to squander their resources. Other perfectionists, Quong notes, claim that citizens suffer from *weakness of will* and thus public action is required in order to steer them away from unworthy temptations and addictions.

In responding to the charge that there is no non-paternalistic justification for perfectionist public action, it is, I think, helpful to consider how neutralist political liberals might themselves respond to this same objection. Why, that is, is state action necessary at all to achieve *ordinary, non-perfectionist, justice-based* objectives? Is it not similarly paternalistic for political liberals to use state power to pursue neutral ends such as education, national defence and social justice? Do political liberals have non-paternalistic grounds for believing that these neutral ends would not be achieved in the absence of state action? In other words: why think that perfectionist state action is paternalistic and non-perfectionist state action non-paternalistic?

Quong says little about this question, but in a footnote he does suggest that state action in areas such as social justice and national defence can be non-paternalistically justified on the grounds that it is necessary for providing assurance to citizens that others are complying with their justice-based duties:

> The state's coercive power may be necessary to provide the requisite assurance to each citizen that others will do their fair share. It seems plausible to suppose that even if we are all committed to acting justly, we cannot know that everyone else is similarly committed. Assuming that it is not irrational or unreasonable to refuse to do one's part in a cooperative scheme without the requisite assurance that others will do likewise, there thus need be no negative judgement implied when the

[67] Quong, J., *Liberalism Without Perfection,* p. 86.

> state provides the requisite assurance by coercively enforcing certain duties of justice.[68]

The problem with this response, though, is that if we owe perfectionist duties to others, as deontic perfectionists maintain, then perfectionists can likewise argue that public action is necessary for providing each citizen with the assurance that others are complying with their perfectionist duties. Just as we need assurance that others are discharging their duty to pay taxes towards neutral ends such as education, we also need assurance that others are discharging their duty to pay taxes towards perfectionist ends such as art galleries. Without such assurance, many perfectionist duties would plausibly be left unfulfilled. Perfectionists, too, that is, can argue with great plausibility that the state's coercive power may be necessary to provide the requisite assurance to each citizen that others will do their fair share towards the promotion of the conditions that generally conduce to moral, intellectual and artistic excellence. So, assuming that refusing to do one's part in a cooperative scheme aimed at promoting human flourishing without the requisite assurance that others will do likewise is no more irrational than refusing to do one's part in a cooperative scheme aimed at promoting justice without such assurance, there need be no negative judgement implied when the perfectionist state provides this assurance by coercively enforcing certain perfectionist duties.

It is worth noting here that mutual assurance of duty compliance is not the only, or perhaps even the most important, non-paternalistic justification available to perfectionists.[69] Indeed, the idea that the state functions as a mutual assurance mechanism is quite a narrow conception of the role of the state. There are, in addition, other non-paternalistic justifications for

---

[68] Ibid., p. 103.

[69] Of course, perfectionists could also simply say (i) that there is *nothing* wrong with paternalism, or (ii) that the prima facie wrongness of paternalism is so *slight* that it would be routinely outweighed by the great benefits of perfectionist policies. For Quong, paternalism is prima facie wrong because it denigrates or disrespects a citizens' moral status by making "negative assumptions about citizens' abilities to know and rationally pursue their own good" (*Liberalism Without Perfection*, p. 86). But what if these assumptions are true, as psychological research increasingly suggests? Is it so denigrating or disrespectful to a person to proceed on the assumption — surely true of *all* human beings — that she will make rational mistakes about what is in her interests? To pose this question is not (as Quong suggests, ibid., p. 107) to call into question the "liberal conception of ourselves as free and equal". It is simply to call into question a particular *interpretation* of free and equal liberal citizenship — an interpretation that equates being treated as free and equal with being treated as if one were a superhumanly reliable practical reasoner.

perfectionism, such as arguments that have to do with externalities, coordination problems, and public goods.[70] My point here is simply that once we accept the deontic perfectionist's claim that there are perfectionist duties, we can answer the "why is state action necessary?" question in the same way that neutralists do, namely by appealing to the role of the state in providing each citizen with assurance that others have complied with their duties.

3.3 The quidnunc mentality

In his recent book *Liberalism with Excellence*, Kramer develops an intriguing and original objection to perfectionism in its mainstream "edificatory" form — namely, that edificatory perfectionist regimes are guilty of an objectionable "quidnunc" mentality, one akin in important respects to that of a village busybody who distributes "boxes of delicious homemade fudge" to "inspire her fellow inhabitants to lead lives of propriety" and who in general meddles in matters that are "none of [her] business".[71] Thus whereas many critics of edificatory perfectionism adopt a citizen-focused perspective — contending, for instance, that perfectionism treats citizens infantilizingly or manipulatively — Kramer instead adopts "a change of perspective" and focuses instead on the character and outlook of an edificatory-perfectionist system of governance and its state officials.[72]

For Kramer, the quidnunc mentality is objectionable in two main ways, namely by being "a form of self-aggrandizement and a form of self-abasement".[73] It is meddlesomely *self-aggrandizing* because perfectionist policies "are aimed neither at preventing the infliction of harm by some people on others nor at fulfilling any of the further responsibilities" of government such as promoting public order.[74] And it is demeaningly *self-abasing* because "it presumes that the success of a system of governance is partly dependent on the willingness of citizens to alter their private pursuits in response to inducements by the

[70] See Quong, J., *Liberalism Without Perfection*, pp. 88-91; Kramer, M., *Liberalism with Excellence*, pp. 63-91.
[71] Kramer, M., *Liberalism with Excellence* (Oxford: Oxford University Press, 2017), pp. 251-96, at p. 282, 287. I discuss Kramer's objection in more depth in Chapter 3.
[72] Kramer, M., *Liberalism with Excellence*, p. 254.
[73] Ibid., p. 277.
[74] Ibid.

system's officials".[75] As he elsewhere puts it, the edificatory-perfectionist state is self-abasing because it makes its success "dependent on choices *that are none of its business*" — choices that are private, choices about which it should not care, choices that should be beneath its notice.[76] An edificatory perfectionist state thus violates a "deontological ethic of self-restraint" — a wrong that "tarnishes its moral integrity", that "far-reachingly informs and taints" its relationship with the citizenry, and that "deprives [its edificatory-perfectionist laws] of moral legitimacy".[77]

This argument has no force, however, against the deontic perfectionist, who holds that governments have duties of edification towards their citizens. For deontic perfectionists, the choices that citizens make about how to live — for instance, the choice to waste away one's life in front of the TV — are not "private pursuits" but are instead a legitimate concern of the state. The state might, for instance, have duties to steer people away from a life wasted away in front of the TV and towards more worthwhile and meaningful avenues. In such a case, an edificatory-perfectionist state is no more self-abasing or self-aggrandizing than a neutral state that pursues imperatives such as justice and security, because their edificatory-perfectionist policies are in fact "necessary for the fulfillment of the elementary moral responsibilities that are incumbent on any system of government".[78] Another way of putting this point is as follows: when Kramer says that edificatory perfectionist states stick their nose where it does not belong, he already assumes that states have no duties to promote the edification of the citizenry; but he cannot help himself to this assumption because whether the edificatory impulse falls within or outside the proper bounds of government is precisely what is at stake here. As before, then, the whole cogency of the quidnunc mentality objection rests on the presupposition that perfectionist duties do not exist — a presupposition that is explicitly rejected by deontic perfectionism.

4 Does deontic perfectionism violate public justification?

[75] Ibid.
[76] Ibid., p. 289 (emphasis added).
[77] Ibid., p. 253, 275, 277, 278.
[78] Ibid., p. 285.

In this section and the next, I seek to defend deontic perfectionism against two objections. The first objection is that deontic perfectionism runs afoul of the public justification principle — the requirement, that is, that political decisions be justifiable to citizens in terms that they could reasonably accept.[79] As political liberals and public reason liberals frequently remark, perfectionist ideals of the good life and of human flourishing are controversial. So wouldn't deontic perfectionism involve the promotion of ends that some citizens reasonably reject? Couldn't the deontic-perfectionist claim that states have a duty to subsidize art galleries, for instance, be reasonably rejected by those tax-paying citizens who prefer to stay at home and watch TV? Taking the fact of reasonable pluralism seriously, would efforts by the state to discharge its perfectionist duties survive the public justification test? "The deeply objectionable characteristic" of perfectionist politics, says Gaus in a forceful rendition of this worry, is that it makes the terms of our common life "dictated by the conscience of some, demanding that others simply obey".[80] He continues:

> Social interaction is to be on their terms, directed by their conscience and their convictions about moral truth. Those who do not share this conscience must submit to the truth. It is precisely this imperious claim of private conscience that the social contract-liberal tradition has rejected.[81]

Broadly speaking, there are two ways that deontic perfectionists can reply to this objection. The first and more hardline response is to simply reject the public justification principle. The public justification principle, after all, is not self-evident. Many remain unconvinced that public justification is a genuine constraint on government action. At most, say the critics, acceptability-to-all is a *pro tanto* value: it would be nice, of course, if all reasonable citizens could accept all laws, but this kind of acceptability is just another item on the long and dense list of political desiderata, and not a particularly weighty item at

---

[79] This principle is of course most closely associated with Rawls's *Political Liberalism*. For a more recent account of public reason liberalism, see Gaus, G., *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World* (Cambridge: Cambridge University Press, 2010).
[80] Gaus, G., *The Order of Public Reason*, p. 230.
[81] Ibid., pp. 230-1.

that.[82] So deontic perfectionists can just throw the ball back into the objector's court by pointing out that public reason liberals need to provide a justification (or grounding, or moral foundation) for public justification. Public reason liberals need, that is, to answer the question: Why reason publicly? Why care about justification to all reasonable citizens? Why not simply implement the political rules and institutions that are *correct* — or at least that are most likely to be correct — regardless of whether those rules and institutions could be reasonably rejected by some citizens?

Of course, defenders of the public justification principle have sought to provide answers to these important questions. Larmore, for instance, argues that public justification is required by respect for persons: "if we try to bring about conformity to a rule of conduct solely by the threat of force [and not by offering reasons that others recognize as valid], we shall be treating persons merely as a means, as objects of coercion, and not also as ends, engaging directly their distinctive capacity as persons".[83] But each of these attempts has been subjected to powerful critiques. For instance, while Larmore is surely correct that *some* failures to deliberate in terms of public reasons and to make publicly justified decisions would indeed be disrespectful (e.g. dictatorship or theocracy), it is unclear why *all* alternative models of democratic deliberation and decision-making short of public reason liberalism are incompatible with respect for persons. Even if a dictatorship treats persons as mere objects of coercion, why does a deliberative democratic approach which requires citizens to favour only policies for which they have a high degree of rational justification, to deliberate openly and honestly, and to refrain from supporting laws that violate the dignity of others — but which does not require citizens to withhold support for policies that are not publicly justified — treat citizens disrespectfully and as mere objects of coercion?[84] Needless to say, public reason liberals cannot answer this question by saying that deliberative democracy is

[82] See, e.g., Enoch, D., "Against Public Reason", in D. Sobel, P. Vallentyne and S. Wall (eds), *Oxford Studies in Political Philosophy*, vol. 1 (Oxford: Oxford University Press, 2015), pp. 112-44, esp. at pp. 138-40.
[83] Larmore, C., "The Moral Basis of Political Liberalism", *The Journal of Philosophy* 96 (1999), p. 607. For other renditions of the argument from respect to public justification, see Boettcher, J., "Respect, Recognition, and Public Reason", *Social Theory and Practice* 33 (2007), pp. 223-49; Nussbaum, M., "Perfectionist Liberalism and Political Liberalism", *Philosophy & Public Affairs* 39 (2011), pp. 3-45.
[84] See Eberle, C., *Religious Conviction in Liberal Politics* (Cambridge: Cambridge University Press, 2002), pp. 84-151.

disrespectful because it permits states to advance laws that some citizens reasonably reject — for that would beg the question against deliberative democrats, for whom the norms of deliberative democracy are perfectly adequate to expresses respect for persons.[85]

The second and more conciliatory response is to accept that public justification is a genuine constraint on governmental action but to argue that a good amount of deontic perfectionism would survive the public justification test. This is the response of so-called "moderate perfectionists", and is the response that I myself favour and have sought to defend in other work.[86] Deontic perfectionists, that is, can follow Rawls in taking a "political turn" and in presenting their view in a way that is "political not metaphysical". A "political" form of deontic perfectionism is one that combines the contractualist commitment to public justification favoured by political liberals with the deontic-perfectionist claim that the state must promote the good life. By taking a political turn of this kind, deontic perfectionists can incorporate many of the central insights of political liberalism — that political philosophers must show greater sensitivity to reasonable pluralism, and that the terms of our common political life should as far as possible be justifiable to all in order to realize a valuable form of civic friendship amongst the citizenry — whilst also avoiding

[85] Moreover, even if it could be shown that the failure to publicly justify is disrespectful *in some sense*, this does not look like an especially grievous kind of disrespect. As Lott nicely puts it, "respect comes in shades": it is disrespectful to cut in front of someone in a queue; it is even more disrespectful to spit on someone who supports a different football team; and it is yet more disrespectful to maim and torture someone on the basis of race or religion. The kind of disrespect involved in the failure to reason publicly looks fairly slight. Public reason, that is, looks to be at the periphery of our concept of respect rather than a paradigmatic instance of disrespect. (See Lott, M., "Restraint on Reasons and Reasons for Restraint: A Problem for Rawls' Ideal of Public Reason", *Pacific Philosophical Quarterly* 87 (2006), pp. 87-9.) This worry about weightiness becomes particularly acute once we observe that there are other important senses or dimensions of respect that are in tension with the dimension of respect picked out by public reason liberals. For Galston, for instance, "we show others respect when we offer them, as explanation, what we take to be our true and best reasons for acting as we do". (See Galston, W., *Liberal Purposes: Goods, Virtues, and Diversity in the Liberal State* (Cambridge: Cambridge University Press, 1991), p. 109.) So, in order to ground the kind of robust public justification principle that they are after — a principle of "very great and normally overriding weight", in Rawls's words (*Political Liberalism*, p. 241) — public reason liberals need to explain why we should care so much about the specific kind of respect (or disrespect) expressed in giving (or failing to give) public reasons. Why does this highly specific and seemingly peripheral kind of respect generally take priority over (a) other senses or dimensions of respect and (b) other values in general?

[86] See Caney, S., "Liberal Legitimacy, Reasonable Disagreement, and Justice", *Critical Review of International Social and Political Philosophy* 1 (1998), pp. 19-36; Chan, J., "Legitimacy, Unanimity, and Perfectionism", *Philosophy & Public Affairs* 29 (2000), pp. 5-42. I defend this view in Chapter 4.

its rather less appealing requirement that we "cordon off political morality from our best understanding of human flourishing".[87] On this approach, then, certain perfectionist commitments, such as moral, intellectual and artistic excellence, are built into the notion of reasonableness, in much the same way that many political liberals make reasonableness conditional on acceptance of certain liberal commitments such as freedom, fairness and equality. Deontic-perfectionist laws and policies such as state subsidies for art galleries would thus be expected pass the test of public justification because a citizen who rejects the value of artistic excellence would count as unreasonable and so her veto would have no more normative force than that of a citizen who rejects the value of equality.

The vital question that arises here is: Is this the right justificatory constituency? What justifies the claim that reasonableness is conditional not only on acceptance of liberal values but on acceptance of perfectionist values too? In particular, should those who reject the value of artistic excellence really be excluded as "unreasonable"? Without a clear answer to these questions, the strategy of baking perfectionist content into the constituency of reasonable citizens looks unmotivated and gerrymandered — and thus unconvincing in comparison to Rawls's and Quong's more deeply theorized neutralist accounts of reasonableness.

It is not possible to address this concern fully here. But here is the sketch of a reply. For Rawls, the bounds of reasonableness are set by "a fundamental organizing idea within which all ideas and principles can be systematically connected and related. This organizing idea is of society as a fair system of social cooperation between free and equal persons".[88] I agree with this way of supplying content to reasonableness. But I am doubtful that the political liberal conception of society — a fair system of social cooperation between free and equal persons — is the only, let alone the best, such conception available. In particular, in other work I argue that a plausible and arguably superior alternative conception of society is that of society as a fair scheme of social cooperation and a striving for human flourishing between free and equal persons. The exclusion of the art-hater from the constituency of reasonable

[87] Wall, S., *Liberalism, Perfectionism and Restraint* (Cambridge: Cambridge University Press, 1998), p. 123.
[88] Rawls, J., *Political Liberalism*, p. 9.

citizens, then, is not an ad hoc fix to ensure that perfectionist policy achieves public justification but rather flows from deeper ideas about the nature of society which set the bounds of reasonableness.[89] In other words, the exclusion of the art-hater may be stipulative, but it is not *arbitrarily* stipulative because there are deeper reasons for not giving her a veto.

Of course, there are limits to how far this strategy can go whilst remaining faithful to the basic public reason motivations.[90] It would clearly make a mockery of the public justification apparatus if we held a view such as "Christianity: Political not Metaphysical", according to which reasonableness is conditional on acceptance of Christian doctrine, where this conception of the constituency of reasonable citizens is in turn justified by appeal to a Christian conception of society. A view that defines "reasonable disagreement" in a way that bears no relation at all to the folk concept is not aptly characterized as a "public reason" view. But I see no reason to think that these limits of fidelity are necessarily coextensive with liberal neutrality. A moderately perfectionist view of the kind that I propose — according to which members of the justificatory constituency endorse both liberal axioms (such as freedom, equality and fairness) and perfectionist axioms (such as moral, intellectual and artistic excellence) — would still leave room for the fact of reasonable disagreement to play a role (such as with respect to religious matters), albeit a more circumscribed role than it plays within political liberalism, and so such a view could plausibly be regarded as true to the spirit of the public reason project and faithful to its underlying motivations. This kind of perfectionism, I hope to have shown, can thus occupy a sweet spot by being thin enough that it satisfies a plausible specification of the public justification condition, yet thick enough that its conception of fair terms is distinguishable from the views of leading liberal neutralists such as Rawls, Dworkin and Quong.

[89] Naturally, this move raises further questions and worries. One such objection is: Doesn't this make the idea of public justification redundant? Isn't all the moral work being done by the perfectionist conception of society that sets the bounds of reasonableness? I address this worry in Section 4 of Chapter 4.

[90] I thank Tom Sinclair for pressing me to be clearer about this.

Does Edificatory Perfectionism Express a Quidnunc Mentality?

It is a central tenet of much contemporary liberal political philosophy that the state should be, in some sense, neutral among the competing reasonable conceptions of the good life that exist within society.[1] A neutral state, it is often said, is one that does not impose or promote any particular ethical, religious or metaphysical doctrine and instead simply upholds a fair system of laws and institutions within which each individual can freely pursue her own conception of the good life. By contrast, perfectionists hold that an important function of the state is to encourage citizens to pursue virtuous, meaningful, or worthwhile ways of life and to discourage them from pursuing degrading or hollow ways of life. For perfectionists, then, liberal neutrality needlessly "cordon[s] off political morality from our best understanding of human flourishing".[2]

Over the years, a wide variety of objections and criticisms have been leveled against perfectionist politics. Some have argued that states should remain neutral because citizens reasonably disagree about the question of what makes for a good life and so perfectionism would inevitably involve the promotion of ends that some citizens could reasonably reject.[3] Others have alleged that perfectionist public policy is paternalistic insofar as it is generally premised on the idea that citizens are weak-willed and cognitively biased and so would, if left to their own devices, squander their resources on worthless pursuits and allurements.[4] Still others have objected that perfectionist political action violates the Millian principle that each person is the best judge of his or her own interests,[5] that it is manipulative and invasive of autonomy,[6] that it is

[1] See, e.g., Larmore, C., *Patterns of Moral Complexity* (Cambridge: Cambridge University Press, 1987); Rawls, J., *Political Liberalism: Expanded Edition* (New York: Columbia University Press, 2005); Quong, J., *Liberalism Without Perfection* (Oxford: Oxford University Press, 2011); Patten, A., "Liberal Neutrality: A Reinterpretation and Defense", *The Journal of Political Philosophy* 20 (2012), pp. 249-72; Schouten, G., *Liberalism, Neutrality, and the Gendered Division of Labor* (Oxford: Oxford University Press, 2019).

[2] Wall, S., *Liberalism, Perfectionism and Restraint* (Cambridge: Cambridge University Press, 1998), p. 123.

[3] Rawls, J., *Political Liberalism*.

[4] Quong, J., *Liberalism Without Perfection*, pp. 73-107.

[5] See discussion in Caney, S., "Consequentialist Defences of Liberal Neutrality", *The Philosophical Quarterly* 41 (1991), pp. 463-5.

[6] Quong, J., *Liberalism Without Perfection,* pp. 45-72.

elitist,[7] that it is intolerant of pluralism and diversity,[8] that it contravenes the "endorsement constraint" according to which political measures cannot improve a citizen's life unless he or she positively affirms them "from the inside",[9] and that it jeopardizes stable social cooperation and peace and even risks the abuse of state power.[10]

Recently, Matthew Kramer has developed an intriguing and highly original objection to perfectionism in its mainstream "edificatory" form — namely, that perfectionist regimes that seek the elevation and edification of the citizenry are guilty of a "quidnunc" mentality, one akin in important respects to that of a curtain-twitching village busybody who distributes "boxes of delicious homemade fudge" to "inspire her fellow inhabitants to lead lives of propriety" and who in general meddles in matters that are "none of [her] business".[11]

In this paper, I first examine, in Section 1, what Kramer means by the quidnunc mentality and why he considers such a mentality objectionable. I then argue, in Section 2, that the quidnunc mentality objection fails because it begs the question against edificatory perfectionism. I consider and reject, in Sections 3 and 4, two potential responses available to Kramer. I conclude, in Section 5, by commenting on a sense in which I think Kramer's quidnunc mentality objection, despite not succeeding in undermining edificatory perfectionism, nonetheless contains an underlying insight that opens up a fascinating and novel set of questions.

1 What is the quidnunc mentality and why is it wrong?

Though this paper will focus on the quidnunc mentality in particular, it is worth briefly situating that critique within Kramer's broader view. In *Liberalism with Excellence*, Kramer advocates a view that he classifies as perfectionist, and he defends perfectionism against a number of objections. However, Kramer's defence of perfectionism is decidedly "arm's-length" because he has little

[7] See discussion in Wall, S., *Liberalism, Perfectionism and Restraint*, pp. 15-7.
[8] Ibid., pp. 17-9.
[9] Dworkin, R., "Foundations of Liberal Equality" in S. Darwall (ed.), *Equal Freedom: Selected Tanner Lectures on Human Values* (Ann Arbor: University of Michigan Press, 1995), pp. 264-73.
[10] See discussion in Sher, G., *Beyond Neutrality* (Cambridge: Cambridge University Press, 1997), pp. 106-39.
[11] Kramer, M., *Liberalism with Excellence* (Oxford: Oxford University Press, 2017), pp. 251-96, at p. 282, 287.

sympathy for virtually all contemporary perfectionist theories such as those of Joseph Raz and Steven Wall.[12] Distancing himself from these existing perfectionist theories — which he groups together under the label "edificatory perfectionism" and which he vigorously assails on a number of fronts, including most significantly their quidnunc mentality — Kramer argues instead for a view he refers to as "aspirational perfectionism". Though edificatory and aspirational perfectionism largely converge on the bottom line insofar as they both permit the promotion of excellence in a range of areas of human endeavor, they differ in their underlying justificatory structure. For whereas edificatory perfectionists seek "to induce citizens to lead lives that are more wholesome or cultivated or worthily autonomous", an aspirational perfectionist eschews this "edificatory impulse" and instead promotes the attainment of excellence within society because this is "a necessary condition for the warrantedness of a robust sense of self-respect on the part of each individual who belongs to that society".[13] In what follows, I focus exclusively on edificatory perfectionism and so I will sometimes drop the qualification "edificatory" for the sake of readability.

One of the distinctive features of Kramer's quidnunc mentality objection to edificatory perfectionism is that it involves "a change of perspective".[14] For whereas many critics of edificatory perfectionism adopt a citizen-focused perspective — contending, for instance, that perfectionism treats citizens infantilizingly or manipulatively — Kramer focuses instead on the character and outlook of an edificatory-perfectionist system of governance and its state officials. In the course of his analysis, he uses a variety of richly suggestive terms to evoke aspects of this outlook: officious, overweening, prickly, heavy-handed, unaccommodating, overbearing, meddlesome, busybody, censorious, nosy, and, of course, its delightful official label "quidnunc".

---

[12] Ibid., p. 45.

[13] Ibid., p. 36, 279. A taxonomical quibble: in my view, Kramer's view is more aptly classified as a species of liberal neutrality than as a species of perfectionism. The fact that Kramer believes that many standard perfectionist policies such as state-supported arts are necessary for citizens' warranted self-respect does not suffice to make his view perfectionist. After all, many neutralists readily concede that there can be arguments from neutral values to perfectionist policies (see, e.g., Dworkin, R., "Can a Liberal State Support Art?" in *A Matter of Principle* (Oxford: Clarendon Press, 1985), pp. 221-33). Nor does Kramer's rejection of public reason suffice to make his view perfectionist (see *Liberalism with Excellence*, pp. 341-2); after all, comprehensive liberals uphold neutrality whilst rejecting public reason (see, e.g., Pallikkathayil, J. "Neither Perfectionism nor Political Liberalism," *Philosophy & Public Affairs* 44 (2016), pp. 171–96).

[14] Ibid., p. 254.

This shift of perspective, says Kramer, is important for understanding what makes edificatory perfectionism "always and everywhere morally wrong" and parallels his earlier work on the ethics of torture, in which he argues that any absolutist account of the wrongness of torture must refer not only to the interests, rights, and dignity of the victim but also to the perpetrators of torture, who, by stooping to this atrocious practice, tarnish their moral integrity as well as the moral integrity of the system of governance on whose behalf they act.[15] The perpetrator-focused perspective is necessary for explaining the absolute wrongness of torture because if we consider matters solely from the perspective of the victim it is difficult to account for the wrongness of certain instances of torture, such as the torture of a victim who is physiologically incapable of experiencing pain or the torture of monstrous individuals guilty of enormous crimes who (at least on certain retributivist accounts of punishment) might deserve to suffer. Similarly, there can be cases in which edificatory perfectionism is not disrespectful or otherwise problematic from the citizen perspective, and so the government-focused perspective is necessary for capturing what makes the "drive for edification" always and everywhere wrong.[16]

But what, more precisely, is the quidnunc mentality and what makes it wrongful (as opposed to merely less than ideal)? Here it is helpful to turn to the main analogy that Kramer uses to illustrate the quidnunc mentality, namely that of the village busybody. On Kramer's view, officials who uphold a perfectionist system of governance interact with their fellow citizens "in broadly the same fashion in which a busybody of some village interacts with her fellow villagers".[17] "She keeps track of how her fellow residents behave," Kramer continues, "and she hectors or ostracizes anyone whom she believes to be falling significantly short" — even though the falling short of her fellow residents "affect[s] her only through her own sentiments of offendedness and gratification".[18]

---

[15] Ibid., pp. 254-64, at p. 255. See also Kramer, M., *Torture and Moral Integrity: A Philosophical Enquiry* (Oxford: Oxford University Press, 2014).
[16] Kramer, M., *Liberalism with Excellence*, p. 264.
[17] Ibid., p. 276.
[18] Ibid.

Kramer argues that the attitude of the village busybody is objectionable in two main ways, namely by being "a form of self-aggrandizement and a form of self-abasement":

> At the level of an individual, the quidnunc mentality is quite evidently a failure of self-restraint. It is both a form of self-aggrandizement and a form of self-abasement. It is a meddlesome form of *self-aggrandizement* because it takes for granted that one's fellows can rightly be prevailed upon — through coercion or through vehement exhortation — to modify patterns of behavior that are not legitimately of concern to anybody except the individuals themselves who engage in those patterns of behavior. Yet the quidnunc mentality also consists in *self-abasement*, since it makes one's own sense of satisfaction partly dependent on the responsiveness of one's fellows to one's officious badgering about their private pursuits. Instead of abiding by a live-and-let-live ethos, a busybody partly ties the success of her own life to her effectiveness in not letting other people lead their lives as they see fit. That is, the success of her own life partly hinges on the willingness of other people to desist from conduct that offends her (even though the conduct does not otherwise harm her or anybody else). In precisely that respect, one's possession of a quidnunc mentality is a manifestation of weakness.[19]

Moving back to politics, Kramer contends that similar points apply to the mentality that characterizes edificatory-perfectionist governments:

> At the level of a system of governance, a quidnunc mentality that underlies various laws or policies is even more egregiously a failure of self-restraint. Again, it is both a form of self-aggrandizement and a form of self-abasement. It is a meddlesome form of *self-aggrandizement* because any laws or policies that emanate from it are aimed neither at preventing the infliction of harm by some people on others nor at fulfilling any of the further responsibilities [of government]. Rather, those laws or policies are aimed at prevailing upon people — coercively or enticingly — to abandon certain patterns of behavior that are not harmful to other people in any significant ways. Furthermore, the

[19] Ibid., p. 277, emphasis added.

> overweeningness of the quidnunc mentality in that regard is integrally linked to the sense in which such a mentality is also a self-abasing manifestation of weakness. It is an outlook of *self-abasingness* because it presumes that the success of a system of governance is partly dependent on the willingness of citizens to alter their private pursuits in response to inducements by the system's officials. Instead of remaining confined to the dispensation of information and the maintenance of fair conditions for informed decisions by citizens about the aforementioned pursuits, a system of governance with laws or policies that stem from a quidnunc mentality will have tied the success of its endeavors partly to its effectiveness in not letting citizens arrive at disfavored decisions about the ways in which they should lead their lives…A scheme of governance should not be dependent for its success on its prodding of citizens about their preferred lifestyles when no requirements of justice are at stake; a scheme of governance should not resemble a busybody whose equanimity hinges on the effectiveness of the similarly officious prodding in which she engages.[20]

In acting on the basis of a mentality that renders it "both grandiosely presumptuous and demeaningly vulnerable", the edificatory perfectionist state violates a "deontological ethic of self-restraint" — a wrong that "tarnishes its moral integrity", that "far-reachingly informs and taints" its relationship with its citizenry, and that "deprives [its edificatory-perfectionist laws] of moral legitimacy".[21]

Before proceeding, one important clarification is in order. In attributing to perfectionist governments a quidnunc mentality, Kramer stresses that he is not seeking to advance an empirical hypothesis about the actual mental states and self-understandings of the officials and functionaries of a perfectionist government.[22] It is not an exercise in social science or in mind-reading: the "imputation of a quidnunc mentality to edificatory perfectionism is not an exercise in psychological speculation from one's armchair" but is instead "the distillation of an outlook from which the drive for edification makes most

[20] Ibid., pp. 277-8, emphasis added.
[21] Ibid., p. 253, 275, 277, 278.
[22] Ibid., p. 254, 269-70.

sense".[23] That is, the quidnunc mentality is meant to refer to a normative property that, "at a high level of generality", would "most credibly account for [the] decisions and behavior" of perfectionist state officials, whether or not they are conscious of it.[24] In this sense, Kramer's argument is in line with the jurisprudential literature on legal interpretation and legislative intent, according to which the attribution of intention to a law or to a government involves not peering into the souls of legislators but rather trying to make sense of a law and the objectives that most plausibly account for its having been adopted and sustained.[25]

2 How the quidnunc mentality objection begs the question

There are at least five ways in which one might criticize the quidnunc mentality objection. First, one can question whether being self-abasing and self-aggrandizing in the Kramerian sense (that is, making one's sense of success dependent on the responsiveness of others to one's efforts to alter their private pursuits, and going beyond what is morally required of one, respectively) accurately captures what it means to be self-abasing and self-aggrandizing in the ordinary morally significant senses of those terms. For instance, does an individual who feels more successful when he manages to persuade his neighbour or his colleague to adopt a healthier, less calorific diet really thereby *abase* himself? If not, Kramer might be getting undue moral and rhetorical mileage out of evocative labels that do not fit the behavior to which they are applied. Second, even if village busybodies are self-abasing and self-aggrandizing in the ordinary morally significant sense, one might wonder how the attitudes of self-abasement and self-aggrandizement are relevant to permissibility as opposed to character evaluation.[26] Why not say that village busybodies act permissibly yet possess flawed characters — in the same way that if someone donates money to a famine-relief charity on the basis of a Mrs Jellyby mentality, or with an attitude of sanctimonious self-congratulation, then, while this shows that she has a flawed character, it does not make her

[23] Ibid., p. 254.
[24] Ibid., p. 270.
[25] See ibid., pp. 75-87.
[26] For the view that intentions are irrelevant to permissibility (though relevant to blameworthiness and character assessment), see, e.g., Thomson, J. J., "Self-Defense", *Philosophy & Public Affairs* 20 (1991), pp. 283-310; Scanlon, T., *Moral Dimensions: Permissibility, Meaning, Blame* (Cambridge, MA: Harvard University Press, 2008).

donation impermissible? Third, even if self-abasement and self-aggrandizement are relevant to permissibility, it is unclear why these vices are *so* great — and why manifesting them is *so* wrong — that they could form the basis of a deontological prohibition of the particularly stringent kind Kramer seeks. Even if it is *ordinarily* impermissible to be self-abasing and self-aggrandizing, that is, these traits are presumably all-things-considered permissible when something more important is at stake. Yet mightn't the importance of village busybodies' objectives — namely, the improvement of other people's lives — exceed in stringency considerations of self-abasement and self-aggrandizement, thereby generally rendering their actions all-things-considered permissible? A fourth line of potential criticism concerns the extrapolation Kramer makes from individual vices to state vices. It is, after all, debatable whether the properties of self-abasement and self-aggrandizement can meaningfully be imputed to states. Fifth, even if states can as a conceptual matter be bearers of attitudes and mentalities, it is unclear that self-abasement and self-aggrandizement have the same moral status when understood as supervenient properties of a collective "self" as they do when understood as character flaws of a village busybody. Given that state attitudes are not at all like individual attitudes, is it really plausible to suppose that they are relevant to permissibility in the way that individual attitudes and intentions are sometimes thought to be?[27]

Though I have some sympathy for each of these lines of argument, in this paper I grant (i) that Kramer accurately captures the morally significant essence of self-abasement and self-aggrandizement, (ii) that these vices are relevant to permissibility, (iii) that they are of sufficient moral significance to ground a particularly stringent prohibition, (iv) that they can be applied meaningfully at the state level, and (v) that they still retain their stringent-prohibition-grounding moral significance when applied to the state. Still, I argue that the quidnunc-mentality objection fails because it begs the question against edificatory perfectionism.

Edificatory perfectionism, recall, is supposedly *self-abasing* in the sense that "it presumes that the success of a system of governance is partly dependent on the willingness of citizens to alter their private pursuits in response to inducements

[27] I thank Jeff McMahan for helping me to distinguish these five lines of criticism.

by the system's officials".[28] As he elsewhere puts it, the edificatory-perfectionist state makes its success "dependent on choices that are none of its business" — choices that are private, choices about which it should not care, choices that should be beneath its notice.[29] Similarly, edificatory perfectionism is supposedly *self-aggrandizing* in the sense of being "aimed neither at preventing the infliction of harm by some people on others nor at fulfilling any of the further responsibilities" of government such as maintaining public order or the provision of public education.[30]

The problem with this argument, however, is that edificatory perfectionists can argue, and have argued, that governments have duties of edification towards their citizens. Call this view *deontic* edificatory perfectionism. For deontic edificatory perfectionists, the choices that citizens make about how to live — for instance, the choice to waste away one's life in front of the TV — are not "private pursuits" but are instead a legitimate concern of the state. The state might, for instance, have duties to steer people away from a life wasted away in front of the TV and towards more worthwhile and meaningful avenues. The charge of self-abasement and self-aggrandizement thus begs the question against deontic edificatory perfectionists for whom the edification of the citizenry is the proper "business" of the state and one of its "further responsibilities".[31] Another way of putting this problem is as follows: when Kramer says that edificatory perfectionist states stick their nose where it does not belong, he already assumes that states have no duties to promote the edification of the citizenry; but he cannot help himself to this assumption because whether the edificatory impulse falls within or outside the proper bounds of government is precisely what is at stake here.

Raz and Wall — whom Kramer takes to be the foremost exemplars of edificatory perfectionism — are both deontic edificatory perfectionists. Both, that is, defend edificatory-perfectionist duties. In *The Morality of Freedom*, Raz argues that "governments are subject to *autonomy-based duties* to provide the

[28] Kramer, M., *Liberalism with Excellence*, p. 277.
[29] Ibid., p. 289.
[30] Ibid., p. 277.
[31] This point is also made in Billingham, P. and Taylor, A., "Liberal Perfectionism, Moral Integrity, and Self-Respect", *The American Journal of Jurisprudence* 63 (2018), p. 69.

conditions of autonomy for people who lack them".[32] "The autonomy principle," he further explains, "permits and even *requires* governments to create morally valuable opportunities, and to eliminate repugnant ones".[33] Similarly, in *Liberalism, Perfectionism and Restraint*, Wall argues that "political authorities *should* take an active role in creating and maintaining social conditions that best enable their subjects to lead valuable and worthwhile lives".[34] In more recent work, Wall argues that edificatory-perfectionist duties are derivable from duties to promote welfare, from duties of fairness, and from duties to secure citizens' merited self-respect.[35]

If Raz's and Wall's substantive arguments are sound — if, in other words, it is true that one of the major responsibilities of government is to induce citizens to lead wholesome, cultivated, virtuous, or autonomous lives — then edificatory perfectionists do not stick their nose in where it does not belong and do not meddle in matters that are none of their business. In such a case, an edificatory-perfectionist state is no more self-abasing or self-aggrandizing than a neutral state that pursues imperatives such as justice and security because their edificatory-perfectionist policies are in fact "necessary for the fulfillment of the elementary moral responsibilities that are incumbent on any system of government".[36] So, if we must employ provincial analogies, deontic edificatory perfectionists are likely to view governmental efforts to encourage the worthwhile and to discourage the worthless as more analogous to the activities of the local magistrate than the local busybody.[37]

---

[32] Raz, J., *The Morality of Freedom* (Oxford: Oxford University Press, 1986), pp. 415 (emphasis added).
[33] Ibid., p. 417 (emphasis added).
[34] Wall, S., *Liberalism, Perfectionism and Restraint*, p. 8 (emphasis added).
[35] For the arguments from welfare and fairness, see Wall, S., "Moral Environmentalism", in C. Coons and M. Weber, *Paternalism: Theory and Practice* (Cambridge: Cambridge University Press, 2013), pp. 93-114; for the argument from merited self-resect, see Wall, S., "Enforcing Morality", *Criminal Law and Philosophy* 7 (2013), pp. 455-71. See also my arguments for edificatory-perfectionist duties in Section 2.3 of Chapter 2.
[36] Kramer, M., *Liberalism with Excellence*, p. 285.
[37] Another reason why the local magistrate is, as far as it goes, a better analogy of the edificatory-perfectionist state than the local busybody (in addition, that is, to the fact that local magistrates generally act in pursuit of objectives that are morally required, and hence are not self-aggrandizing or self-abasing in Kramer's sense) is that busybodies are generally *self-appointed*, whereas magistrates and edificatory-perfectionist governments are generally not. That is to say, in addition to self-aggrandizement and self-abasement, a further feature (though not one that Kramer focuses on) that makes the village busybody an unappealing character is that she *takes it upon herself* to promote the edification of others. But an edificatory-perfectionist system of governance receives a democratic mandate to pursue edificatory objectives such as opera and art galleries, and so it is not self-appointed. So insofar as the

Of course, Kramer would no doubt deny that there are any such edificatory-perfectionist duties. But, as outlined above, edificatory perfectionists such as Raz and Wall have provided extensive argumentation in support of edificatory-perfectionist duties. Thus the non-existence of edificatory-perfectionist duties cannot be simply assumed here. The moral of the story is that, in order to successfully wield the quidnunc mentality objection, one cannot remain at this level of abstraction. It is one's prior view of the state's duties and its legitimate domain of activity (and, correspondingly, one's prior view of the public/private distinction) that does all the real moral work in determining whether a quidnunc mentality is present. So Kramer needs to engage the substantive question of whether states really have moral duties to promote the flourishing and edification of the citizenry — for *that* is where the action really is. To do this, one needs either to provide some specific reason to reject the arguments of Raz and Wall, or to provide some general reason to think that duties of edification do not exist, or ideally both. In the absence of such substantive argument, the quidnunc mentality objection assumes precisely what needs to be shown.[38]

3 A substantive argument against edificatory-perfectionist duties?

While Kramer does not engage in specific rebuttals of Raz's and Wall's arguments in favour of edificatory-perfectionist duties, he has indicated that some general reason to doubt the existence of edificatory-perfectionist duties is given in his discussion of the passages from Shakespeare:[39]

> I there maintain that, outside the sway of the harm principle (the genuine harm principle [rather than Raz's perfectionist reformulation of

---

magistrate analogy does not trade on this morally relevant feature, it is preferable to the busybody analogy. (I thank Rowan Cruft for discussion of this point.)

[38] To be clear: this is not yet to say that edificatory-perfectionist policies are in the clear. Perhaps other arguments — for instance, political liberals' concern that perfectionist judgements are subject to reasonable disagreement and thus that perfectionist policies violate the public justification requirement — tell against them. (For an attempt to reconcile public justification and edificatory perfectionism, though, see Chapter 4.) The point is just that such policies cannot be impugned on the specific grounds that they manifest the mentality of a village busybody, with its unattractive combination of self-aggrandizement and self-abasement.

[39] See Kramer, M., *Liberalism with Excellence*, pp. 270-8.

> it — CT]) and the preservation of a societal infrastructure, the measure of the ethical strength of a system of governance is given by what it can safely ignore and does ignore. An ethical eagle can and does safely ignore the chirping of the little birds. Since the ethical strength of a system of governance is a key constituent of the political and social and economic conditions under which everyone can be warranted in harboring an ample sense of self-respect, edificatory-perfectionist duties would run athwart a system's responsibility for bringing about those conditions.[40]

This argument — that the conditions of citizens' warranted self-respect depend on the ethical strength of government, and that the ethical strength of government depends in turn on what it safely ignores — is developed too briefly in the relevant pages of *Liberalism with Excellence* to allow for a proper evaluation of its merits, and my understanding is that it will be explored in much greater depth in Kramer's book-in-progress *A Stoical Theory of Justice*. Nonetheless, let me comment specifically on the suggestion that this argument would rule out edificatory-perfectionist duties.

For this argument to rule out edificatory-perfectionist duties, it would have to be the case that those things that trouble edificatory perfectionists — such as individuals wasting away their lives in front of the TV — can be safely ignored by the state. But edificatory perfectionists will want to know why it is that citizens' failures to realize edification and flourishing can be safely ignored by the state, whereas citizens' failures to realize other values such as equality and justice *cannot* be safely ignored by the state. Why, to use the Shakespearean idiom, does the former failure but not the latter failure count as the trifling and ignorable "chirping of little birds"? A lot here will hinge, I think, one what is meant by "safely ignore". In general, one is not just safe but safe *from* something. So a natural question here is: *In what sense, and from what threat,* could states remain safe whilst ignoring the failure of some citizens to live edified and flourishing lives?[41]

---

[40] In written comments.

[41] Perhaps safety is to be cashed out with reference to the conditions for warranted self-resect: thus states can ignore some citizens' failures to live edified lives because the edification of the citizenry falls outside of the state's duty to promote the conditions of citizens' warranted self-resect. But this looks circular: the conditions of warranted self-respect are defined in terms of what can be safely ignored, and what can be safely ignored is defined in terms of the conditions

If, on the one hand, Kramer holds that citizens' failures to live edified and flourishing lives could be safely ignored in the sense that they could be ignored *without thereby threatening the continued existence of society*, then he is no doubt correct. Edificatory perfectionist should readily agree that their concerns can be safely ignored in this sense. But this interpretation would generate an overly demanding criterion of state action: it would rule out efforts to promote important neutral values such as social justice because (as the political experience of many less liberal countries today demonstrates) states can ignore social injustice without thereby threatening the collapse of society. Yet if, on the other hand, and in an attempt to rule in social justice, Kramer holds that citizens' failures to live edified and flourishing lives could be safely ignored in the sense that they could be ignored *without thereby threatening important values*, then edificatory perfectionists are unlikely to agree that their concerns can be safely ignored because they see flourishing as an important value. In short, it is unclear how the distinction between what can and what cannot be safely ignored tracks the distinction between what falls inside and what falls outside the edificatory impulse.[42]

Perhaps, as a third possibility, the idea is that citizens' failures to live edified and flourishing lives could be safely ignored in the sense that they could be ignored *without thereby threatening harm to others*. Something along these lines does seem to be playing a role in Kramer's argument; although he does not provide a specification or defence of Mill's harm principle, Kramer does make a number of passing references to it.

But if Kramer hitches the quidnunc mentality objection to the wagon of Mill's harm principle, he would face at least two challenges — one conceptual, the other normative. First, he would need to engage the vexed question of how to

---

of warranted self-respect. So we are back to the question at the outset of this section: why does the duty to promote the conditions for citizens' warranted self-respect not include the duty to promote the edification of the citizenry? (And the answer cannot be "Because the conditions for citizens' warranted self-respect depends on the ethical strength of government, which in turn depends on what it safely ignores, which in turn rules out edificatory-perfectionist duties", for that only raises the question, posed in the main text, of why citizens' failures to live edified lives can be safely ignored.)

[42] See Chan's related discussion of Nagel's antiperfectionist argument from practical necessity: Chan, J., "Legitimacy, Unanimity, and Perfectionism", *Philosophy & Public Affairs* 29 (2000), p. 24ff.

conceptualize harm, explaining, among other things, what is stopping edificatory perfectionists from arguing that states that fail to take measures to prevent their citizens from sinking into worthless and degraded ways of life count as harming the citizenry. Raz, for instance, argues that the harm principle is merely "a formal principle lacking specific concrete content and leading to no policy conclusions" until it has been "embedded" within a specific moral theory. [43] He thus contends that (autonomy-based) edificatory perfectionism is perfectly compatible with the harm principle once that principle is theorized "from the point of view of a morality which values autonomy".[44] As Raz puts this: "Sometimes failing to improve the situation of another is harming him…[A state which] fails to discharge [its] autonomy-based obligations [or its edificatory-perfectionist obligations more generally] towards others is harming them, even if those obligations are designed to promote the others' autonomy [or edification] rather than to prevent its deterioration."[45]

And second, even if Kramer could answer this conceptual question and could show that edificatory-perfectionist concerns could be safely ignored without threatening harm to others, he would face the normative question of why it is *this* conception of what can be safely ignored — as opposed to what can ignored *without thereby threatening harm* (*including harm to oneself*) or *without thereby threatening important values* — that has moral significance and that should be taken to constrain the behaviour of states. After all, critics of the harm principle, among whom some edificatory-perfectionists belong, often point to actions and omissions that do not harm others (in the intuitive sense of not causing others to be worse off than they would otherwise have been) yet that should plausibly still be regulated by political authorities. Familiar examples include: the exploitation of someone's misfortune through a mutually beneficial transaction, the failure to save a drowning child in a case of easy rescue, voluntary self-enslavement, the destruction of a species whose extinction will have no adverse impact on others, and issues arising from the non-identity problem. The harm principle (or at least a formulation of the harm principle that rules out edificatory perfectionist policies) would also seem

[43] Raz, J., *The Morality of Freedom*, p. 414.
[44] Ibid.
[45] Ibid., pp. 416-7. See also Feit, N., "Harming by Failing to Benefit", *Ethical Theory and Moral Practice* (forthcoming).

to rule out many paternalistic interventions — though the legitimacy of paternalistic measures is admittedly more contested than the legitimacy of the other measures mentioned alluded to above.[46] Indeed, it is arguable that some of the *most* serious wrongs — those wrongs that it is most important to discourage and even criminalize — are either non-harmful or else not seriously wrong in virtue of their harmfulness.[47] So while harm to others might be *sufficient* for state intervention, why is it also a *necessary* condition for state intervention? Indeed, does Kramer not suggest that harm to others is sufficient but not necessary for the justification of state action when he states that edificatory-perfectionist policies are "aimed neither at preventing the infliction of harm by some people on others *nor at fulfilling any of the further responsibilities* [of government]"?[48] In brief, then, Kramer would need to tell a deeper moral story about why we should accept Mill's harm principle (as does Mill himself, who argues for the harm principle variously on the basis of the best-judge principle, individuality and experimentation, utility, and so on) as opposed to the many alternative proposed principles of political morality (and thus the many alternative ways of cashing out the what-can-be-safely-ignored principle).

Let me make a final, somewhat speculative point about the critical leverage of the harm principle in the context of a dialectic in which we are trying to establish whether or not edificatory-perfectionist duties exist. The speculation is that the harm principle cannot tell us what duties exist because, in order to determine whether someone has "harmed" another, we already have to know what rights and duties there are. So looking to the harm principle to settle the substantive debate over edificatory-perfectionist duties gets things backwards, justificatorily speaking. Consider how we say that a parent who fails to provide for his own children harms them, whereas a parent who fails to provide for some unknown third-party's children does not harm them. Plausibly, the reason for this is that children have *rights* against their own parents, and parents have *duties* towards their own children. Determinations of harm, then, depend on prior judgements about what we owe to others, about how we are

[46] For a powerful recent defence, though, see Hanna, J., *In Our Best Interest: A Defense of Paternalism* (Oxford: Oxford University Press, 2018).
[47] See, e.g., Tadros, V., *Wrongs and Crimes* (Oxford: Oxford University Press, 2016), esp. p. 95, 102, 201-3, 216-22, 292-7.
[48] Kramer, M., *Liberalism With Excellence*, p. 277, emphasis added.

entitled to be treated, and about the allocation of rights and duties.[49] Indeed, some have taken this point so far as to suggest that we should "let harm go the way of phlogiston": since harm is notoriously difficult to conceptualize, and since "nobody really needs to talk about harm", serious moral and political theorizing should instead appeal to more well-behaved and fundamental moral concepts such as rights and duties.[50] So it strikes me as unlikely that the harm principle will help us determine whether or not edificatory-perfectionist duties exist: rather, the harm principle seems to come in later in the day, as it were, only once the deep and difficult questions about our rights and duties have been figured out. This does not mean that the harm principle has no role to play in politics. Perhaps it is a good heuristic or rule of thumb. But it is hard to see how the harm principle could serve as a fundamental principle of political morality.

4 Abandoning perfectionism?

At one point, Kramer considers the possibility of a deontic edificatory perfectionism (though not under that name), as outlined in Section 2.[51] He says that "autonomy-centered perfectionists", for whom "the nurturing of each individual's autonomy" is a "matter of justice", may argue that the quidnunc mentality objection simply begs the question by assuming that the promotion of autonomy is not a proper responsibility of government.[52] "A riposte of [this] kind", Kramer continues, "could most promisingly be advanced in a Rawlsian guise" whereby autonomy is viewed as a "primary natural good in Rawls's sense".[53]

But the main problem with presenting edificatory perfectionism as a doctrine of justice, argues Kramer, is that it would no longer be distinctively perfectionist:

[49] See also Wall, S., "Moral Environmentalism", p. 107.
[50] Bradley, B., "Doing Away with Harm", *Philosophy & Phenomenological Research* 85 (2012), p. 391, 411. See also Holtug, N., "The Harm Principle", *Ethical Theory and Moral Practice* 5 (2002), pp. 357-89.
[51] Kramer, M., *Liberalism with Excellence*, pp. 289-96.
[52] Ibid., pp. 289-90.
[53] Ibid., p. 289.

> When the attainment of autonomy is taken to be a primary natural good…and when opportunities to attain autonomy are taken to be distribuenda covered by the principles of justice that should apportion primary goods in any liberal-democratic society, the proponents of an autonomy-centered account of justice are not disagreeing with Rawlsians over the doctrine of liberal neutrality. Instead, at the very most, they are disagreeing with Rawlsians over the index of primary natural and social goods…Hence [the view] would not itself be perfectionist. It would be concerned with modifying or clarifying Rawls's thin theory of the good rather than with affirming that a system of governance can legitimately aim to promote some thicker conception(s) of the good.[54]

Indeed, Kramer argues that, given that Rawls himself recognizes the value of autonomy, this approach may end up being not just similar but *identical* to Rawlsianism: "instead of turning autonomy-centered perfectionism into a *variant* of Rawls's theory of justice, the re-elaboration would turn it into Rawls's theory *simpliciter* or into something extremely close to Rawls's theory".[55] After all "when we ask what principles of justice can appropriately distribute opportunities for autonomy in order to foster the attainment of autonomy as a primary natural good, we could plausibly conclude that Rawls's principles (or principles very close to Rawls's) serve that purpose admirably".[56] Thus, "somebody developing an autonomy-centered account of justice might well fix upon Rawls's principles".[57]

However, there are two main problems with Kramer's worry about indistinguishability.[58] First, it is not clear why deontic edificatory perfectionists would need or want to present their view in terms of primary goods. Kramer's argumentation is thus curious here because he himself states that deontic edificatory perfectionism "could most promisingly be advanced in a Rawlsian guise" before going on to critique such a view for being indistinguishable from Rawlsianism. But edificatory perfectionists such as Raz and Wall have not gone

[54] Ibid., pp. 292-3.
[55] Ibid., p. 294 (emphasis added).
[56] Ibid., p. 295.
[57] Ibid.
[58] See also the discussion of indistinguishability in Section 2.3 of Chapter 2.

about defending their view in Rawlsian terms. Behind Wall's endorsement of edificatory-perfectionist duties, for instance, is not a Rawlsian picture in which autonomy is a primary good to be distributed in accordance with some distributive principle but rather a fairly un-Rawlsian picture in which autonomy is a component of an objectively valuable and flourishing human life and the role of the state is to foster human flourishing:

> [Perfectionism] is the thesis that political authorities should take an active role in creating and maintaining social conditions that best enable their subjects to lead valuable and worthwhile lives. If autonomy is a central component of a fully good life, and if we ought to accept the general thesis of perfectionism, then it would follow that political authorities should take an active role in creating and maintaining social conditions that help their subjects realize this ideal.[59]

Once deontic edificatory perfectionism is presented in its natural form, and not forced into a Rawlsian mold, it is difficult to see how this view has lost its distinctively perfectionist character just because it makes claims about what the state "ought" (as opposed to "may") to do. Deontic edificatory perfectionism, after all, asks a very different question from that asked by liberal neutralism (and aspirational perfectionism). It asks: What do people in general need to flourish, to live a life of meaning and worth? Which conditions generally conduce to, and which detract from, the moral, intellectual, cultural and artistic excellence of the citizenry? In answering these full-bloodedly perfectionist questions, we are unlikely to conclude that Rawls's principles would "serve [our] purpose[s] admirably".

The second problem with Kramer's move here is that, even if (contra the preceding point) the best way to depict the difference between Rawlsians and deontic edificatory perfectionists were in terms of a disagreement about the index of primary natural and social goods, it would not follow that deontic

[59] Wall, S., *Liberalism, Perfectionism and Restraint*, p. 131. It is worth noting that Kramer does consider this passage from Wall but his only comment on it is that its "strongly perfectionist understanding of the role of government" is "redolent of the quidnunc mentality" (Kramer, M., *Liberalism with Excellence*, p. 293). But this takes us back to the argument in Section 2. For what makes it quidnunc? It cannot be that Wall's conception of the role of government goes beyond the proper responsibilities of government because, again, whether or not the promotion of valuable and worthwhile lives is a proper responsibility of government is the question at issue here.

edificatory perfectionism is indistinguishable from Rawlsianism or that it endorses liberal neutrality. This becomes particularly clear when we consider what Wall calls "non-autonomy-based perfectionist political action" — that is, when we consider not just *autonomy-centered* deontic edificatory perfectionism but also other ideals, distinct from autonomy, that edificatory perfectionists have sought to promote.[60]

To see this, consider some autonomy-unrelated examples from Wall. Suppose that an edificatory-perfectionist state "grants public subsidies to artists" and museums in "an effort to promote and stimulate appreciation of high art".[61] Suppose next that an edificatory-perfectionist state "implements regulations and restrictions designed to discourage the consumption of a class of drugs" as a way of "taking steps to promote the well-being of its subjects".[62] Suppose, finally, that an edificatory-perfectionist state "adopts an immigration policy that favors those who are familiar with the national culture" as part of an effort to ensure "the preservation of its distinctive national culture".[63] The most natural way to understand Wall's point here, it seems to me, is that the role of the state is to promote human flourishing and thus the state should promote these elements of the flourishing life (i.e., appreciation of high art, well-being, and a distinctive national culture). But even if, somewhat contrivedly, we read Wall as suggesting that appreciation of high art, well-being, and a distinctive national culture should be understood as primary goods, as "distribuenda covered by the principles of justice", it does not follow that Wall's view is no longer edificatory-perfectionist. A disagreement about the index of primary goods, in short, is not just a minor intramural matter for Rawlsians: it can conceal deep disagreements between Rawlsians and edificatory perfectionists.

5 Conclusion: in search of new mentalities?

I have argued that Kramer's account of the quidnunc mentality objection begs the question against edificatory perfectionism by assuming that efforts to

[60] As Wall also puts it, "autonomy-based political action is one type of perfectionist political action. But it is not the only type." See Wall, S., *Liberalism, Perfectionism and Restraint*, p. 213.
[61] Wall, S., *Liberalism, Perfectionism and Restraint*, p. 213.
[62] Ibid., p. 219.
[63] Ibid., p. 228.

promote the edification of the citizenry are outside of the proper remit of government.

However, while I ultimately do not think that Kramer's quidnunc mentality objection succeeds, let me close by commenting on a sense in which I think Kramer's discussion nonetheless contains an underlying insight that opens up a fascinating and novel set of questions. The insight I have in mind is Kramer's "change of perspective" — his shift from evaluating the treatment of the citizen to evaluating the mentality of the state. After all, one might accept Kramer's general contention that the deeply objectionable feature of edificatory perfectionism resides in its mentality (or more precisely: in the mentality credibly ascribable to it), even while rejecting (as I have done) his particular contention that the quidnunc mentality with its attendant self-abasement and self-aggrandizement is the relevantly objectionable mentality that sets edificatory perfectionism apart from its rivals.

One interesting line of further inquiry, then, is to try to understand what other mentalities might be reasonably ascribable to states, perfectionist and otherwise. Of course, this search for new mentalities, as we might call it, need not be undertaken only with the aim of criticizing perfectionism and in fact it might end up providing grist for perfectionism's mill. For perhaps the mentality that at a high level of generality credibly accounts for edificatory perfectionism is not quidnunc meddlesomeness but rather an attractive form of solicitude and benevolence and a genuine desire to see other people flourish — an ethos perhaps most powerfully affirmed by Martin Luther King:

> We are tied together in a single garment of destiny, caught in an inescapable network of mutuality. And whatever affects one directly affects all indirectly. For some strange reason I can never be what I ought to be until you are what you ought to be. And you can never be what you ought to be until I am what I ought to be…John Donne caught it years ago and placed it in graphic terms — 'No man is an island entire of itself. Every man is a piece of the continent — a part of the main.' And he goes on to say, 'Any man's death diminishes me because I am

> involved in mankind. Therefore never send to know for whom the bell tolls; it tolls for thee'".[64]

Similarly, I wonder whether the high-level mentality most plausibly imputable to liberal neutralism is not some admirable ethic of self-restraint but rather an unappealing and more Scrooge-like attitude of indifference, frostiness, alienation and self-absorption. So where perfectionism embodies the ethos evoked by Martin Luther King, perhaps neutralist liberalism channels the spirit of Cain who asks, "Am I my brother's keeper?" Naturally, many questions arise here, including that of how to adjudicate which of a range of potentially ascribable mentalities is predominant in any given system of governance. But the interesting point for our purposes is that the legacy of Kramer's quidnunc mentality objection might thus be to initiate a shift towards the identification and evaluation of the mentalities of perfectionist and neutral states — a shift that would produce new theories and new language and that would shed new light on what is at stake in the neutrality-perfectionism debate.

[64] Martin Luther King Jr., "Remaining Awake Through a Great Revolution" in J. Washington (ed.), *A Testament of Hope: The Essential Writings and Speeches of Martin Luther King, Jr.* (New York: Harper Collins, 1986), pp. 269-70. (I thank Mark Reiff for drawing my attention to this passage.)

Perfectionism: Political not Metaphysical

1 A perfectionism come of age?

It is a central tenet of much contemporary liberal political philosophy that the state should be, in some sense, neutral among the competing conceptions of the good life that exist within society. A neutral state, it is often said, is one that does not impose or even promote any particular ethical or religious doctrine. By contrast, perfectionists reject the principle of state neutrality and hold that the state may legitimately promote some vision of the good or flourishing human life. For perfectionists, the state "should act with discrimination to encourage the good and the valuable and to discourage the worthless and the bad".[1]

It is, by now, widely recognized that there are two importantly different ways of upholding neutrality: comprehensive liberalism and political liberalism.[2] Comprehensive liberalism — held historically by Kant and Mill,[3] and more recently by early Rawls,[4] late Dworkin,[5] Pallikkathayil[6] and others — is the traditional view and holds that neutrality is entailed by, or flows from, some comprehensive moral or philosophical doctrine. For instance, in *On Liberty* Mill presents the view that individuality and "experiments in living" are central components of the good life and, therefore, that some kind of liberal neutrality is justified on the grounds that only a neutral state would create the conditions necessary for such autonomous and experimental living and would prevent individuals being channeled into stultifying, conformist ways of life that are not of their own choosing.[7]

---

[1] Raz, J., "Liberalism, Skepticism, and Democracy", *Iowa Law Review* 74 (1989), p. 785.
[2] See, e.g., Quong, J., *Liberalism Without Perfection* (Oxford: Oxford University Press, 2011), pp. 15-21.
[3] This, at least, is how the canonical story goes. See, e.g., Rawls's references to Mill and Kant as comprehensive liberals in his *Political Liberalism* (New York, NY: Columbia University Press, 1996), p. 78, 145, 159, 199.
[4] Rawls, J., *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971).
[5] Dworkin, R., "Foundations of Liberal Equality" in *The Tanner Lectures on Human Values* (Salt Lake City: University of Utah Press, 1990).
[6] Pallikkathayil, J., "Neither Perfectionism nor Political Liberalism", *Philosophy & Public Affairs* 44 (2016), pp. 171-96.
[7] Mill, J. S., *On Liberty* (London: Penguin Books, 1985).

Against this, *political* liberals argue that comprehensive liberalism's defence of neutrality is, in an important sense, *not neutral enough.* In modern societies, they stress, many citizens will reasonably reject Mill's comprehensive story about the good life, and so comprehensive liberalisms of this kind will not command the acceptance of all reasonable citizens and become the basis of a just and stable order. For Rawls, a liberalism based on Millian individuality or on some other "comprehensive moral ideal" would be "but another sectarian doctrine", but another attempt to "embody the whole truth in politics".[8] Similarly, Larmore argues that, since in conditions of modernity there is no systematic agreement about the ends and purpose of human life, a "liberalism come of age" will not be grounded in some general "philosophy of man": liberals, he argues, must instead provide a "neutral justification of political neutrality".[9] Political liberals therefore typically defend neutrality in morally minimal, broadly contractualist terms — by reference to the fact of reasonable disagreement, to a principle of public justification, and to a norm of equal respect — rather than by reference to some comprehensive moral ideal.

However, while liberals have paid close attention to the distinction between comprehensive liberalism and political liberalism, perfectionists have hardly availed themselves of this distinction. Almost all perfectionists are *comprehensive* perfectionists: that is, perfectionists who base their perfectionism in some comprehensive moral doctrine. Thus Hurka and Sher ground their perfectionisms in broadly Aristotelian ideas about human nature and human potentiality,[10] Raz and Wall ground their perfectionisms in ideas about personal autonomy,[11] Jiang grounds his perfectionism in Confucianism,[12] and Yurako grounds her perfectionism in a particular version of feminism.[13] And, of course, the paradigmatic historical example of perfectionism, namely

---

[8] Rawls, J., "Justice as Fairness: Political not Metaphysical", *Philosophy & Public Affairs* 14 (1985), pp. 245-6; Rawls, J., *Political Liberalism*, p. 442.
[9] Larmore, C., *Patterns of Moral Complexity* (Cambridge: Cambridge University Press, 1987), p. 51, 53, 129.
[10] Hurka, T., *Perfectionism* (Oxford: Oxford University Press, 1990); Sher, G., *Beyond Neutrality: Perfectionism and Politics* (Cambridge: Cambridge University Press, 1997).
[11] Raz, J., *The Morality of Freedom* (Oxford: Oxford University Press, 1986); Wall, S., *Liberalism, Perfectionism and Restraint* (Cambridge: Cambridge University Press, 1998).
[12] Jiang, Q., *A Confucian Constitutional Order*, eds. D. Bell and R. Fan and trans. E. Ryden (Princeton: Princeton University Press, 2013).
[13] Yuracko, K. A., *Perfectionism and Contemporary Feminist Values* (Bloomington: Indiana University Press, 2003).

"the Catholic and Protestant states of the early modern period", was based squarely in Christian doctrine.[14]

But this neglects an important and plausible position in the conceptual space, namely *political perfectionism*. According to political perfectionism, states may promote some ideals of the good or flourishing human life (hence the *perfectionism*) but must base their view in the same sort of morally minimal, broadly contractualist framework as political liberals (hence the *political*).[15] Thus, for political perfectionists, the conception of the good life that states may promote does not flow from some comprehensive ethical or religious doctrine (be it Millian, Aristotelian, Confucian or Christian) but is instead determined by contractualist considerations about what is reasonably acceptable to suitably idealized citizens. To put the view in a Rawlsian key: the state may act on the basis of a freestanding political conception of the good life which draws on fundamental ideas implicit in the society's public ethical culture, which enjoys overlapping consensus in the sense that it fits into citizens' various comprehensive doctrines in the manner of a module, and which all citizens may reasonably be expected to endorse.

Political perfectionism has generally been either overlooked or considered faintly incoherent. In their two-by-two grid (which includes comprehensive liberalism, political liberalism, comprehensive perfectionism, and political perfectionism), Mulhall and Swift leave the political perfectionism quadrant empty, saying "we know of no theorist, contemporary or otherwise, who could be described as a purely political perfectionist".[16] Indeed, they seem to find the

[14] Rawls, J., *Political Liberalism*, p. 195.

[15] There are different interpretations of what the "political turn" within liberal theory — a turn I here recommend to perfectionists — consists in. Following the interpretation that informs much contemporary political liberalism and public reason liberalism, I understand a view to be "political" when it is grounded in a morally minimal and broadly contractualist framework, as opposed to a comprehensive doctrine. Others interpret political views as those that are restricted not only in their grounds but also in their domain. A conception of the flourishing life that is political in this twofold sense (perhaps civic republicanism is an example) would also naturally count as a form of political perfectionism in my sense. But, to do so, it need not be political in domain: an ideal of the good or flourishing human life that applies to individuals not only in their lives as citizens but also in their personal lives, and that applies not only to the "basic structure" of society but also to institutions and associations generally, would be political in my sense, so long as this ideal is publicly justifiable. I thank a reviewer for pressing me to clarify this point.

[16] Mulhall, S. and Swift, A., *Liberals and Communitarians* (Oxford: Blackwell Publishing, 1996), p. 252.

position — which, in effect, allows governments to promote the good life but does *not* allow them to do so on the basis of any general theory of the good life — rather paradoxical, saying: "it is difficult to see what might justify such a schizophrenic or masochistic position, in which the theorist denies herself the resources that she allows to the politician".[17] Quong, returning to the same grid, mentions a few philosophers who might qualify as political perfectionists before quickly adding, "but it is not clear that any of them…would endorse the political perfectionist label".[18] And while Quong himself does not find political perfectionism incoherent (as Mulhall and Swift seem to do), he does regard it as a kind of intellectual curiosity, saying in a recent paper:

> Political liberals who take the fact of reasonable pluralism as an essential premise in political philosophy have a clear rationale for seeking to construct a freestanding conception of political justice that could be the subject of agreement amongst an idealized constituency of persons who agree only on some abstract political values. But it is unclear what would motivate a perfectionist to pursue a contractualist strategy of this sort. Once particular claims about what is intrinsically valuable are assumed to be true and to be the legitimate source of political reasons, the main impetus behind seeking a freestanding conception — the fact of reasonable pluralism — seems to have been jettisoned. I don't see why such perfectionists wouldn't simply prefer to make direct appeals to what they take to be true claims about the good life or human flourishing (which is, of course, what most contemporary perfectionists do).[19]

However, as I shall go on to argue, there are good reasons for perfectionists to pursue a political rather than comprehensive approach. Far from "jettisoning" it, perfectionists can and should take seriously the fact of reasonable pluralism. In this way, perfectionists should candidly acknowledge that political liberalism has taught us much about the fact of reasonable pluralism and its significance for political philosophy, and they should concede that perfectionist politics

---

[17] Ibid.
[18] Quong, J., *Liberalism Without Perfection*, p. 20.
[19] Quong, J., "*Liberalism Without Perfection:* Replies to Lister, Kulenovic, Zoffoli, Zelic, and Baccarini", *Philosophy and Society* 25 (2014), p. 101. See also Quong, J., *Liberalism Without Perfection*, p. 217.

based on comprehensive religious or philosophical systems is no longer morally viable today, if indeed it ever was. Thus political perfectionism can actually give us the best of both worlds by incorporating the central insights of political liberalism — that the terms of our common political life should as far as possible be justifiable to all, and that political philosophers must show greater sensitivity to reasonable pluralism — whilst also avoiding its rather less appealing requirement that we "cordon off political morality from our best understanding of human flourishing".[20] In this sense, the lessons of political liberalism are well taken, and even though political perfectionism does not follow political liberalism in demanding neutrality on the good life, it is still sensitive to the underlying anxieties and aspirations that gave rise to talk of neutrality in the first place. Far from being paradoxical or unmotivated, then, it seems reasonable to say, adapting the phrase of Larmore, that political perfectionism represents a perfectionism come of age.

In the rest of this paper, I seek to motivate and defend political perfectionism. I first defend, in Sections 2 and 3, its *political* aspect against comprehensive perfectionists who see no need to adhere to public justification and against political liberals who question why perfectionists would be motivated to pursue public justification. I then defend, in Section 4, the *perfectionist* aspect of political perfectionism against political liberals who doubt that much perfectionist policy would pass the public justification test. Finally, I address in Section 5 the concern that political perfectionism is too mild a position in this debate to be of any interest.

2 A suspicious tendency

Although, as I indicated, political perfectionism is a neglected view, there are a few thinkers whose views can be understood as versions of political perfectionism, namely the so-called "moderate perfectionists" Caney, Chan and Mang.[21] I largely agree with the general approach of these thinkers, though I

[20] Wall, S., *Liberalism, Perfectionism and Restraint*, p. 123.

[21] Caney, S., "Liberal Legitimacy, Reasonable Disagreement, and Justice", *Critical Review of International Social and Political Philosophy* 1 (1998), pp. 19-36; Chan, J., "Legitimacy, Unanimity, and Perfectionism", *Philosophy & Public Affairs* 29 (2000), pp. 5-42; Mang, F., "Liberal Neutrality and Moderate Perfectionism", *Res Publica* 19 (2013), pp. 297-315. Billingham's "perfectionist internal conception" is also a form of political perfectionism, though he does not actually endorse the perfectionist internal conception and only posits it as a device for

would take issue with a tendency in moderate perfectionist thinking to date — one that I seek to redress in subsequent sections.

The tendency I have in mind is that moderate perfectionists have generally developed their views in quite a dialectically reactive way. Let me explain. Political liberals standardly argue that the state may not promote any particular conception of the good life because questions of the good and of human flourishing are subject to reasonable disagreement. In response, moderate perfectionists have been quick to point out that, while some aspects of the good life (e.g. religion) are indeed subject to reasonable disagreement, many others are not. Who, after all, would seriously deny that friendship and love, science and wisdom, amusement and integrity are part of a good life? Who could reasonably doubt that music and beauty, knowledge and virtue are valuable for their own sake? Surely any reasonable person would agree that an individual consumed by drug addiction — who is constantly chasing after his next fix — leads a life that is lacking in meaning and worth?

This, of course, is one half of the so-called asymmetry objection to political liberalism — the objection that, at a deep level, political liberals treat the "right" and the "good" in an arbitrarily asymmetric way.[22] And, in presenting this objection, moderate perfectionists have generally piggybacked on the contractualist material of political liberals, arguing that, however the notion of reasonable rejectability is understood by political liberals, the right and the good will come out having roughly equivalent ("symmetrical") levels of controversy and rejectability. For instance, Chan states that the "aim" of his article is to show that "Nagel's contractualist argument not only fails to justify state neutrality" but "can actually be turned around to support…perfectionism"[23]; Caney has in separate articles argued that the contractualist theories of Rawls, Larmore and Barry in fact vindicate

---

critiquing Quong's internal conception of political liberalism. See Billingham, P., "Liberal Perfectionism and Quong's Internal Conception of Political Liberalism", *Social Theory and Practice* 43 (2017), pp. 79-106.

[22] The other half points out that many questions of the right (e.g. questions concerning distributive justice, capital punishment, immigration, just war) *are* subject to reasonable disagreement. So, in sum, the asymmetry objection takes off from the observation that the good/right distinction does not track the distinction between what is and what is not subject to reasonable disagreement.

[23] Chan, J., "Legitimacy, Unanimity, and Perfectionism", p. 36, 42.

perfectionism rather than political liberalism[24]; and Mang opens his paper by saying "this article defends a moderate version of state perfectionism by using Gerald Gaus's argument for liberal neutrality as a starting point for discussion".[25]

Taking off, in this way, from influential political liberals — Barry, Gaus, Larmore, Nagel, Rawls, Quong — has resulted in moderate perfectionists not generally examining the contractualist details that political liberals have agonized over, such as how far the justificatory constituency of reasonable citizens should be idealized or whether perfectionist policies must be publicly justified in the consensus sense (whereby citizens accept a policy on the basis of shared reasons) or the convergence sense (whereby citizens accept a policy on the basis of reasons that may differ from citizen to citizen).

My worry with this dialectically reactive approach, then, is that it feeds the suspicion that political perfectionism is not well-motivated: that political perfectionists are not serious about their contractualism, that they use the contractualist framework in a kind of intellectually opportunistic way, and that perfectionists should instead just make direct and upfront appeals to the truth, as they see it, about the good life and human flourishing (i.e. should just adopt comprehensive perfectionism).

3 Why perfectionists should take a political turn

To dispel the concern that political perfectionism is not well motivated, perfectionists need to provide a positive rationale for their taking a political turn. This is the goal of this section. To achieve it, it will be helpful first to get clearer about what perfectionism is and why it has been seen as incompatible with, or at least as ill-fitting with, contractualism.

Perfectionism can be helpfully understood in terms of two claims or commitments.[26] The first is a claim in *value theory*, namely that a good human

[24] Caney, S., "Impartiality and Liberal Neutrality", *Utilitas* 8 (1996), pp. 273-93; Caney, S., "Liberal Legitimacy, Reasonable Disagreement, and Justice".
[25] Mang, F., "Liberal Neutrality and Moderate Perfectionism", p. 297.
[26] Here I follow Wall: see his "Perfectionist Politics: A Defense", in T. Christiano and J. Christman (eds), *Contemporary Debates in Political Philosophy* (Oxford: Blackwell, 2009), pp. 99-117.

life involves the pursuit or achievement of objectively valuable goods and ideals. This claim is to be contrasted with a subjectivist view according to which the goodness or value of an individual's life is determined by whether it is satisfying to her or by whether it fulfills her actual or idealized desires. These objective human goods are often presented in the form of a list and commonly include items such as virtue, wisdom, knowledge, friendship, autonomy, and appreciation of true beauty. And naturally different perfectionists — Plato, Confucius, Aquinas, Marx, Nietzsche historically, and Raz, Hurka, Sher, Wall in more recent times — include different items in their respective accounts of the objectively valuable human goods. Second, perfectionists make a claim in *political theory*, namely that there is no special constraint against the state's promotion of these objectively valuable goods and ideals. So while perfectionists do of course disagree about how these goods may be promoted (e.g. may the good be promoted coercively, or only by non-coercive means?), they all agree that objectively valuable moral ideals may in principle play a role in guiding the activity of the state.

Having characterized perfectionism, we can now see the source of the worry about the coherence of, and motivation for, political perfectionism. The worry comes from the objectivist strand of perfectionism. After all, if one is an objectivist, why would one also be a contractualist? If one thinks there are objective truths about what is and is not a flourishing life, why not simply make direct appeal to these truths rather than going through the trouble of putting would-be laws before the tribunal of citizens' reasonable judgement?

There is, of course, an obvious answer to these questions. The answer involves piggybacking once again on the arguments that political liberals have given in response to similar questions asked of their own view. After all, as Rawls and others are keen to emphasize, political liberals do not have to be skeptics or relativists about the good. So, like objectivist political perfectionists, objectivist political liberals also have to explain why they care about public justification. In response, political liberals have appealed to ideas such as legitimacy, stability and respect. Along similar lines, then, political perfectionists could simply import the political liberals' argument that a law or policy that fails to secure public justification is not legitimate, or would not be stable in the right way, or does not respect citizens as free and equal. This would give political

perfectionists a clear basis for attending to the fact of reasonable disagreement.[27]

The worry with this approach is that most perfectionists have not generally seen any problem — from the point of view of respect, stability, legitimacy and so on — with promoting laws and policies that are not publicly justifiable. That, after all, is what has made most perfectionists *comprehensive* perfectionists. And so while there is, in theory, nothing stopping someone from signing up both to perfectionism as well as to some further views about stability, respect and legitimacy, the existence of this possibility hardly dispels the suspicion that perfectionism and contractualism are ill-fitting because it is not a well-integrated combination of views that perfectionists have generally taken up. So it would be preferable if it could be shown that perfectionists *qua perfectionists* have reason to be invested in the public justification apparatus. This kind of distinctively perfectionist reason would speak both to comprehensive perfectionists who have not seen any need to seek public justification for their favoured policies as well as to political liberals who struggle to see why perfectionists would choose to present their view in contractualist terms.

Putting the problem in this way helps us to see that what we want is an argument that constitutively connects the pursuit of public justification with some aspect of an objectively worthwhile or flourishing human life. For perfectionists to show that, qua perfectionists, they have internal reasons to seek public justification, they have to show that public justification is grounded in ideas of the "good", as opposed to grounding it, as political liberals have generally done, in ideas of the "right" such as respect,[28] stability,[29] justice,[30] legitimacy,[31] and fundamental reason-giving features of social morality.[32]

---

[27] In rare cases where political perfectionists have touched on the question of foundations, they have gone down this line. For instance, Chan appeals to legitimacy ("Legitimacy, Unanimity and Perfectionism", p. 35) and Billingham to legitimacy and respect ("Liberalism Perfectionism and Quong's Internal Conception of Political Liberalism", p. 100). See also Quong, J., *Liberalism Without Perfection*, p. 20.

[28] Larmore, C., "The Moral Basis of Political Liberalism", *The Journal of Philosophy* 96 (1999), pp. 599-625.

[29] Weithman, P., *Why Political Liberalism?* (Oxford: Oxford University Press, 2010).

[30] Quong, J., "On the Idea of Public Reason", in J. Mandle and D. Reidy (eds), *The Blackwell Companion to Rawls* (Oxford: Wiley Blackwell, 2013), pp. 265-80.

[31] Freeman, S., *Justice and the Social Contract* (Oxford: Oxford University Press, 2007), p. 221.

[32] Gaus, G., *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World* (Cambridge: Cambridge University Press, 2011).

Fortunately, in recent years some theorists have begun to develop a broadly communitarian defence of public reason, arguing that the practice of public justification makes possible a valuable form of community in which, despite profound and widespread disagreements on a range of important matters, citizens exhibit *civic friendship* in their relations with one another.[33] Civic friendship could serve very naturally as a foundation for political perfectionism for two reasons. First, unlike the other moral foundations on which public justification is customarily based — respect, stability, justice, legitimacy, and so on — civic friendship seems to belong more to the category of the good than to that of the right. It is plausible to think, that is, that human life goes better to the extent that it involves more relations of civic friendship and fewer relations of civic enmity. And second, civic friendship is a specification of a general ideal, namely friendship, that already features in most perfectionists' lists of the objective goods in a flourishing life and that finds a place in many religious, ethical and philosophical traditions.

Civic friendship is a relationship that can obtain between citizens and that shares some of the valuable features of ordinary interpersonal friendship. Different theorists have different accounts of what feature civic friendship shares with ordinary friendship. Some hold that *cooperation* is fundamental to both forms of friendship; others say that the valuable feature that civic friendship has in common with ordinary friendship is that of *acting in concert and constituting a collective agent*; and yet others say that the relevant feature is *non-prudential concern* for others.[34] I have sympathy with each of these accounts, but would add that, like ordinary friends, civic friends have a fairly thick set of shared substantive values and commitments.[35] Thus, on this

[33] See Ebels-Duggan, K., "The Beginning of Community: Politics in the Face of Disagreement", *The Philosophical Quarterly* 60 (2010), pp. 50-71; Lister, A., *Public Reason and Political Community* (London: Bloomsbury, 2013); Leland, R. J. and van Wietmarschen, H., "Political Liberalism and Political Community", *Journal of Moral Philosophy* 14 (2017), pp. 151-75.

[34] See, respectively, the accounts of Ebels-Duggan, Lister, and Leland and van Wietmarschen.

[35] On the surface, this might look similar to Leland and van Wietmarschen's claim that friendship (both civic and ordinary) requires "an at least partially shared conception of what is in one another's interest" ("Political Liberalism and Political Community", pp. 158-9). However, Leland and van Wietmarschen cash out this notion in a very thin and procedural way, such that civic friends' shared conception can consist merely in each citizen recognizing "the good of making one's own choices" (ibid., p. 159). As I go on to explain, I am doubtful that this account of what must be shared amongst civic friends is thick enough to ground an argument for public justification because a shared conception of this rather procedural kind could be achieved

account civic friendship is a cooperative relationship between citizens who have non-prudential care and concern for each other and who act in concert on the basis of a fairly thick set of shared substantive values and commitments. While this account does not retain all of the features of ordinary interpersonal friendship — most obviously, civic friends lack the intimacy and emotional closeness of ordinary friends — the thought is that it includes enough of what we care about in friendship that civic friendship thus understood represents a desirable and morally significant relation that should matter to perfectionists.

To support the claim that civic friendship involves a fairly thick set of shared substantive values and commitments, it is helpful to consider the role that shared substantive values and commitments play in ordinary interpersonal friendship. Many philosophers are attracted to the idea that friendship involves shared values and commitments.[36] Friends spend time together engaging in, and talking about, shared interests, hobbies, enthusiasms, hopes and fears. Through speaking and deliberating together, they form a partially shared set of values and commitments. They make plans and undertake projects on the basis of this partially shared conception, and when one friend senses that the other is going off in the wrong direction, she will intervene: she will remind her friend of all those conversations they've had about what really matters in life. This joint perspective and common evaluative ground forms an important part of the bond between friends.

This is not, of course, to say that friends must agree about everything or that friends cannot hold certain divergent values and commitments. Indeed, this is just as well for the analogy with civic friendship, because agreement about everything or even most things is hopelessly unrealizable among the millions of citizens of modern pluralistic societies. But some kind of shared horizon of substantive values does seem necessary for ordinary interpersonal friendship, and so disagreements in some domains will need to be offset by agreements in other domains. If Adam and Betty purport to be friends but are at odds about everything — if he is a factory farmer, she an animal rights activist; he a self-

---

within alternative democratic theories that do not include a public justification principle. (The same goes for the accounts of Ebels-Duggan and of Lister.) For an argument from civic friendship to public justification to go through, then, some more substantial kind of shared ground must, it seems to me, be a necessary feature of civic friendship.

[36] For an account of love that emphasizes shared ends, see Ebels-Duggan, K., "Against Beneficence: A Normative Account of Love", *Ethics* 119 (2008), pp. 142-70.

indulgent hedonist, she an effective altruist — then one is inclined to doubt the sincerity either of their commitments or of their friendship. It is true that Adam and Betty, despite their highly divergent ends, could still enjoy some features associated with friendship, such as showing kindness to each other. But friendship is not just about doing things *for* each other; it is also, crucially, about doing things *with* each other. And even if Adam and Betty were to make efforts to accommodate themselves to, or involve themselves with, each other's activities and pursuits, the worry is that they are only "humouring" each other or otherwise engaging in some subtle form of condescension that is at odds with the spirit of friendship.[37] So insofar as individuals lack shared substantive values and commitments — insofar as individuals cannot make each other's ends their own — they miss out on an essential part of what it is to be a friend.

Similar considerations apply in the political arena in the case of civic friendship. It is often said that, given that pluralism about religious and philosophical doctrines is "the inevitable long-run result of the powers of human reason at work within the background of enduring free institutions" in liberal democratic societies, we should not expect a form of political community in which all citizens share a worldview, religion or common way of life.[38] Yet, while it is no doubt true that we should not expect citizens to *fully* agree on their substantive values and commitments, we should also be careful to avoid the opposite mistake of conceptualizing the relation between citizens as being one of antagonism or mutual incomprehension — of seeing citizens as grouped by their disparate doctrines and unable to freely endorse the terms that govern their common life. For on this latter picture "there is only a *modus vivendi*, and society's stability depends on a balance of forces in contingent and possibly fluctuating circumstances".[39] It is thus necessary to view citizens as possessing some fairly thick set of shared substantive values and commitments. As political liberals have argued, such shared substantive values and commitments include abstract political ideals such as freedom, equality, fairness and justice (though, as I will argue in Section 4, some perfectionist ideals such as moral, intellectual and artistic excellence also belong to this set). When citizens express their non-prudential concern for each other by acting in concert on the basis of these shared substantive values and commitments, they

[37] See ibid., p. 160.
[38] Rawls, J., *Political Liberalism*, p. 4.
[39] Ibid., p. 392.

realize civic friendship and thus achieve what Rawls would call "the deepest and most reasonable basis of social unity available to us in a modern democracy".[40]

Having spelled out an account of civic friendship and explained how it relates to ordinary friendship, let us consider the connection between civic friendship and public justification. The idea I want to defend is that civic friendship requires public justification — that relating to others as civic friends requires that we do not impose on them laws and policies that they might reasonably reject. When we ensure that the terms of our common life are acceptable to all reasonable citizens we express the kind of cooperativeness, non-prudential concern, and commitment to shared values that is constitutive of civic friendship. We act and reason *with* others, rather than merely *for* them.[41] By contrast, a politics in which citizens fight it out, individually and in groups, to try to impose their favoured sectarian doctrines on others — in which citizens seek to "embody the whole truth in politics", as opposed to seeking overlapping consensus — would make citizens seem more like civic enemies than civic friends.

Against this claim that civic friendship requires public justification, however, Billingham has argued that there are alternative models of political discourse and decision-making that realise civic friendship yet that do not involve public justification.[42] Of course, if we take, say, inflammatory and sectarian religious fulminations as the paradigm of non-public reason, then it might seem that civic friendship requires public justification.[43] But if we focus on more plausible and attractive alternatives to public justification, it becomes less clear why *all* alternative ideals short of public justification must fail to realise civic friendship. One such sympathetic alternative is Eberle's "ideal of conscientious engagement", which requires citizens to favour only policies for which they have a high degree of rational justification, to deliberate openly and honestly, and to refrain from supporting laws that violate the dignity of others, but which, importantly, does not require citizens to withhold support for policies

[40] Ibid., p. 391.
[41] See Ebels-Duggan, K., "The Beginning of Community", p. 56.
[42] Billingham, P., "Does Political Community Require Public Reason? On Lister's Defence of Political Liberalism", *Politics, Philosophy & Economics* 15 (2016), pp. 20-41.
[43] See Eberle, C., *Religious Conviction in Liberal Politics* (Cambridge: Cambridge University Press, 2002), pp. 109-15.

that are not publicly justified.[44] Billingham refers to these alternative models as "argumentative democracy" and argues that civic friendship does not in fact require public justification because argumentative democracy also enables citizens to realise civic friendship.

However, an important difference between a society governed by a principle of public justification and a society governed by the ideal of conscientious engagement is that, in the former, citizens act on the basis of shared substantive values and commitments such as freedom and equality. Thus, when members of a society governed by public justification advance laws and policies, they are acting in concert on the basis of shared substantive values and commitments and so they realize civic friendship. By contrast, mere argumentative democracy does not realise civic friendship because its outcomes will not necessarily be seen by all citizens as advancing a shared conception of their values and commitments. Argumentative democracy, in other words, allows the political community to pursue and promote values that some citizens regard as fundamentally alien.

Perhaps Billingham would argue here that we should not exaggerate the difference between the values and commitments shared by citizens of argumentative democracies and those shared by citizens of societies regulated by public justification. After all, argumentative democracy involves a range of shared commitments, including, notably, a commitment to following democratic procedures. However, citizens in argumentative democracy do not have a thick set of shared *substantive* values and commitments — those, that is, beyond the commitment of each to follow democratic procedures and to abide by norms of honesty and openness. The crucial question is therefore what *kind* of set of shared values and commitments is necessary in order to realise civic friendship. Is argumentative democracy's rather thin and procedural shared conception enough to realise civic friendship, as Billingham suggests, or is something more robust and substantive required?[45]

---

[44] Ibid., pp. 84-108.

[45] The labels "procedural" and "substantive" are not perfect here because democratic procedures also realize substantive values such as equality and fairness. (I thank a reviewer for drawing my attention to this point.) But the value of equality, fairness and other substantive goods is not exhausted by democracy, and so citizens in argumentative democracy will have a thinner set of shared substantive values and commitments than that of citizens in a society regulated by public justification. Ultimately, nothing hangs on the terminology: the key question is whether

In my view, it is doubtful whether the largely procedural shared values and commitments possessed by citizens in argumentative democracy will be sufficient to realise civic friendship. To see that civic friendship requires agreement not just on procedure but on substance too, it can be helpful to return again to the analogy of ordinary interpersonal friendship. Consider two flatmates, Colin and Dan. Colin and Dan have nothing in common — no shared hobbies, no shared sense of humour, no shared values — except a shared procedural endorsement of the good of making choices for oneself.[46] Each has an attitude of benign neglect towards the other, and they live parallel lives in separate, non-overlapping social circles. In this case, it is hard to see how the relationship between Colin and Dan, who live like ships passing in the night, can be said to realise the good of friendship. Colin and Dan might be flatmates, but they are surely not mates.

It might be objected here that the Colin and Dan case is unfair because its decision-making procedure does not closely approximate that of argumentative democracy, which generally has citizens being more engaged with each other's lives. But now consider two other flatmates, Eve and Fred. Like Colin and Dan, Eve and Fred agree on nothing, substantively speaking, but do agree to a decision-making procedure. Where Colin and Dan agree in effect to live "together yet apart", Eve and Fred's compromise procedure involves taking it in turns when deciding how to spend weekends. Every first week, then, Eve drags a reluctant Fred to church, and every second week Fred drags a reluctant Eve to ballet. Fred finds church unpleasant and boring, and Eve thinks much the same of ballet. In this case — however much both parties are committed to the "alternate weekends" decision-making procedure — it is doubtful that Eve and Fred's relationship realizes the good of friendship since they do not manifest mutual concern by acting on a fairly thick set of shared substantive commitments and values. Given their mutual disdain for each other's values and pastimes, it would seem conceptually confused to describe Eve and Fred's rather strange arrangement as a friendship.

---

citizens' shared commitment (be that "procedural" or "thinly substantive") to democratic decision-making is sufficient to realize civic friendship.

[46] This example is meant to challenge Leland and van Wietmarschen's claim that civic friends' shared conception can consist in something as thin as simply recognizing "the good of making one's own choices". See footnote 35.

Perhaps there are special cases in which friendships can involve these kinds of procedural compromises. Consider Gwen and Hanna: very old childhood friends who have grown apart but who still care deeply about one another. Gwen and Hanna might plausibly go along with what the other enjoys doing simply because they like spending time together. Indeed, this willingness to compromise — to go along with what the other person enjoys, despite not enjoying it oneself — can express an important kind of devotion, loyalty and commitment between friends.[47]

But, for two reasons, this observation does not undermine the argument from civic friendship to public justification. First, it is doubtful that this kind of case can support argumentative democracy because it trades on a specific feature of ordinary interpersonal friendship — namely, emotional connection and intimacy — that has no analogue in argumentative democracy. Once this feature is removed, the terms on which the relationship is conducted begin to look far less appealing. It would be quite strange, for instance, if acquaintances or less intimate friends had nothing in common and if they endlessly made compromises about how to spend their time together. Second, and even granting that old childhood friendship does analogize well to argumentative democracy, it is important to stress that I need not deny that friendships based on compromises can realise and express important values such as devotion, loyalty and commitment. All I claim is that something is missing in such friendships. Compromises, in other words, compromise *something*, and what they compromise — what is missing in a case of childhood friends such as Gwen and Hanna who still care about each other but who have grown apart, as compared with similarly longstanding and affectionate friendships in which the friends' values and interests have not diverged — is the specific value of sharing dreams and values, of engaging in mutually rewarding activities and hobbies, and of choosing how to spend time together on the basis of a joint understanding, developed over many years, of what really matters in life.

Similarly, and moving back to politics, consider Ian, who, though honest and well-intentioned, coercively imposes upon Jenny views that she reasonably rejects. The cases of Colin and Dan, and Eve and Fred, are intended to support

[47] I thank two reviewers for putting to me a point along these lines.

by analogy the thought that, in such a case, Ian and Jenny do not relate to each other as civic friends, even when these reasonably rejected impositions issue from a decision-making procedure that Jenny accepts and even though Ian and Jenny are committed, in all the ways Billingham stresses, to deliberating honestly and openly with one another. Consideration of the analogy with ordinary friendship thus suggests that there must be some robust substantive agreement between citizens if they are to relate as civic friends. Shared endorsement of procedures, personal or democratic, is not enough. So, insofar as argumentative democracy does not secure civic friendship, the availability of this alternative fails to undermine the claim that civic friendship requires public justification.

In light of this, there need be nothing curious or unmotivated about political perfectionism. Since most perfectionists are already committed to the ideal of friendship, since civic friendship is one plausible specification of that ideal, and since under conditions of pluralism realizing a relation of civic friendship requires public justification, perfectionists have internal reasons qua perfectionists to take a political turn and to attend to the fact of reasonable disagreement.

Let me close this section by addressing two further objections to this argument. First, one might worry that grounding public justification in the good of civic friendship does not generate a principle with the requisite strength and moral status. The public justification principle is usually conceptualized not merely as a valuable ideal but as a strict constraint — one of "very great and normally overriding weight", in Rawls's words.[48] The problem is that perfectionists are typically committed to a list of objective values, only one of which is friendship. While adherence to public justification may advance civic friendship, it may also compromise other objective values that the perfectionist considers important, such as science, music, art, knowledge and virtue. So while the value of civic friendship might give a perfectionist a weak pro tanto reason to commit to public justification, it is not clear how it would generate a binding and overriding public justification principle.

[48] Rawls, J., *Political Liberalism*, p. 241.

A public justification principle grounded in some aspect of the good life will admittedly be weaker than one grounded in some generally inviolable deontological notion such as respect for persons. But I do not regard this as a defect. For many political liberals have greatly overstated the strength of the public justification principle, which may indeed be better conceived of as one desideratum or ideal among others. Since not every valuable perfectionist policy will be accepted by all reasonable citizens, there is a trade-off between perfectionist value and public justification. So a weaker public justification principle — one that makes space for the realization of other weighty perfectionist values — can be seen as a virtue rather than a vice.

However, even if it is conceded that a plausible public justification principle would be *sometimes* overridable, one might still worry that a public justification principle grounded on the value of civic friendship would be *too easily* or *too routinely* overridden. When the demands of civic friendship come into conflict with the promotion of other perfectionist values, why would we expect civic friendship even to generally trump? Even if it does not function as a strict constraint, can a public justification principle grounded on civic friendship at least be expected to hold sway in the normal run of cases?[49]

For three reasons, it seems unlikely that the principle would be routinely overridden by countervailing perfectionist considerations. First, and most obviously, many of the core activities of a perfectionist state will be non-perfectionist in character: they will involve the pursuit of justice, fairness, equality, and other such neutral values. So the frequency of conflict between civic friendship and other perfectionist values should not be exaggerated because in many instances there will be no other perfectionist values in play to create any such conflict.

Second, it is important to be clear that, where conflicts do indeed arise, the conflict is not between civic friendship and other perfectionist goods but rather between civic friendship and the *marginal* value of other perfectionist goods. Consider the case of art. In deciding whether a civic-friendship-based public justification principle would be robust enough, the relevant question is not: Does the value of civic friendship generally outweigh the value of art? After all,

[49] I thank Jonathan Quong and a reviewer for putting this objection to me.

a key contention of this paper — one that I defend at length in Section 4 — is that much state-supported art will survive the public justification test. Moreover, much art will be promoted by private actors within civil society, and thus will not even be subject to the public justification test. So a society regulated by public justification will already realize much artistic value. What conflicts with civic friendship is thus not art simpliciter but rather those forms of art that would not be publicly justifiable and that would not be promoted by civil society. Perhaps fumage — a surrealist art form that involves painting with smoke — would be an example. In considering the robustness of a civic-friendship-based public justification principle, then, the relevant question is: Does the value of civic friendship generally outweigh the marginal value that some publicly non-justifiable art form (such as fumage) adds to an already-rich and already-valuable artistic culture? Unlike the former question, which sets civic friendship against art simpliciter, this question can quite plausibly be answered in the affirmative.

Third, the value of civic friendship should not be underestimated. The worry about the weakness of a civic-friendship-based public justification principle is tempting if civic friendship is viewed as though it were merely a matter of common courtesies such as saying "please" and "thank you" or holding the door open for strangers.[50] But such a picture sells civic friendship short. In reality, the realization of civic friendship is a tremendously valuable and impressive achievement: when citizens, despite their myriad disagreements about how to live, and despite sharing no bonds of blood, forego the impulse to impose their preferred sectarian doctrines and instead express non-prudential concern for each other by seeking to act in concert on the basis of a shared set of substantive values and commitments, they realize the deepest and most reasonable form of community and social unity available in modern democracy. This achievement is not something to be taken lightly.

All of this, to repeat, is not to say that there will be no case in which the marginal value of perfectionist policies that breach the public justification principle outweighs the cost in terms of civic friendship. The point is just that, once we clarify when such conflicts arise (only with respect to certain state actions), what is at stake in such conflicts (only the marginal value of

[50] See Lister, A., *Public Reason and Political Community*, p. 129.

perfectionist goods), and why civic friendship is an especially weighty value (it impressively realizes community in the most improbable of circumstances), it seems unwarranted to insist that civic friendship will regularly and routinely be outweighed. So, while it is impossible to say anything highly precise here because we are simply dealing with competing pro tanto reasons, my sense is that grounding public justification in civic friendship will generate an interesting principle of political morality.

The second objection is that I have not provided much in the way of argument for the claim that civic friendship is an aspect of the good life. One might therefore expect at this point an account of the good/right distinction, an account that makes it clear why civic friendship belongs to the "good" side and respect, legitimacy, justice and so on to the "right" side.

Giving a basis for the good/right distinction turns out to be rather tricky, however, and is not something that political liberals have always been especially clear about. As Sher rightly states, "the extension of [the good/right distinction] is better understood than its intension".[51] We have a good intuitive grip on the distinction, recognizing that the *good* includes "religious doctrines, ideals of character and virtue, aesthetic and cultural values, and norms of sexual behavior" and that the *right* includes "claims about health, economic prosperity, security, freedom, equality", even if we cannot explain the basis for this categorization.[52] The lack of an underlying account here need not therefore be as damning as it might at first seem: since friendship is often taken as a paradigmatic ideal of the good, and since civic friendship is a plausible specification of this ideal, it is reasonable to think that any account of the good/right distinction would locate civic friendship on the good side of that distinction.

One might worry here about the strength of the analogy between friendship and civic friendship. Even if ordinary interpersonal friendship would always fall on the good side of the good/right distinction, can we be so sure that *civic* friendship — which is a political rather than personal relationship — would do so? By the time friendship has become civic friendship, is it not so watered-

[51] Sher, G., *Beyond Neutrality*, p. 38.
[52] Ibid.

down that it can no longer claim a place in the perfectionist pantheon of human goods?

However, my account of civic friendship is intended to preserve many of the valuable features of ordinary interpersonal friendship. Thus, although it is true that some elements of interpersonal friendship (e.g. emotional closeness) do indeed disappear in the transition to civic friendship, many other elements (e.g. sympathetic, mutual, non-prudential concern for another based on shared substantive values and commitments) do not. Many of the norms that give friendship its meaning and richness, such as goodwill and sincerity, can be preserved in civic friendship. There are also civic analogues of the shared history that characterizes personal friendship. The move from friendship to civic friendship is therefore not best conceived as a move from goodwill, care and fellow-feeling to their absence but rather as a move from one *form* of goodwill, care and fellow-feeling to another.[53] Civic friendship thus seems to bear on human flourishing, just as ordinary friendship does. In short, since the right is often thought to be about the fair distribution by social institutions of rights, duties, burdens and benefits, and not primarily about the qualitative character of the relationship between people, civic friendship is more aptly viewed through the lens of the good than the right.

4 Would much perfectionism survive the public justification test?

So far, I have sought to defend the *political* aspect of political perfectionism. In this section, I turn to defending the *perfectionist* aspect of political perfectionism. There are various important objections to perfectionism, and it is not possible to respond to them all here. But an important and recurring charge has been that since perfectionist ideals of the good life and of human flourishing are controversial, perfectionist politics would inevitably involve the promotion of ends that some citizens reasonably reject. A natural worry at this point, then, is that a political perfectionism would not be robustly and recognizably perfectionist. After all, couldn't standard perfectionist policies such as subsidies for art galleries be reasonably rejected by those citizens who endorse comprehensive doctrines that are opposed to the relevant kind of art?

[53] See Schwarzenbach, S., "On Civic Friendship", *Ethics* 107 (1996), p. 113.

Taking the fact of reasonable pluralism seriously, how many aspects of the flourishing life can be expected to survive the public justification test?

While this is an entirely natural worry to have, I should say, first of all, that it is unfair in one respect. Political perfectionism is not meant to be a specific political theory; rather, it is a *category* of political theories, just as political liberalism, comprehensive liberalism and comprehensive perfectionism are categories of political theories. There are many possible varieties of political perfectionism, depending on the moral foundation of the view, the degree to which the justificatory constituency is idealized, the structure and scope of the public justification principle, and other related questions. So the question of how much perfectionism would be publicly justified according to political perfectionism cannot be answered at this level of abstraction, just as the answer to the question of what would be publicly justified according to political liberalism depends on whether one is asking, say, Rawls or Gaus.

Having said this, if there were no positive reason to think that at least an appreciable amount of perfectionism would pass the test of public justification, then the entire category of political perfectionism would fail to even get off the ground. We thus need some sense of how a political perfectionist might go about publicly justifying perfectionist ideals. To this end, let me spell out my general way of thinking about public justification. This way of theorizing public justification, I shall argue, vindicates an appreciable amount of perfectionism. But, again, to be clear: the reason for offering a deeper and more fleshed-out account of public justification at this point is to illustrate how political perfectionism *could* get up and running; since political perfectionism is a family of views, it is not wedded to this (or to any other) specific account of the public justification principle.

The public justification principle requires that laws and policies be in some sense acceptable to, or justifiable to, all reasonable citizens. An immediate and formidable problem for those on board with the public justification project, however, is this: Given the depth of disagreement amongst sincere, good-willed, intelligent citizens in modern pluralistic societies, will anything pass this test? Isn't the public justification requirement a recipe for libertarianism at best and anarchism at worst?

Some are willing to embrace this conclusion. Gaus, for instance, notes that the anarchist challenge "seems like the right sort of challenge for a liberal theory".[54] It reminds us, he says, of an important insight familiar to classical social contract liberals such as Hobbes and Locke but often forgotten by contemporary egalitarian liberals whose views often move in a socialist direction: namely that "liberalism is at the edge of anarchy" and that, for liberals, the question must always be that of why state coercion is justified.[55] On this view, public justification is a radical principle which, when faithfully applied, tilts towards a "purer", more minimal kind of liberalism. Others, however, have tried to resist this implication and to show that a more substantial range of public policies could survive the public justification test. The standard way of doing this is to idealize the constituency of reasonable citizens to whom justification is owed. On Quong's view, for instance, acceptance of certain liberal axioms such as freedom, fairness and equality is part of the very definition of what it means to be reasonable. By building more content into the notion of reasonableness in this way, it becomes clearer how public justification could vindicate a more substantial egalitarian-liberal political order.

I favour the second approach to conceptualizing public justification: I believe, that is, that any adequate public justification principle will need to employ very strong idealizations. So the crucial question for a public justification view is this: How should the justificatory constituency be idealized, and what justifies this choice of idealizing conditions?

This way of thinking about public justification leads to political perfectionism because while endorsement of certain core liberal axioms (namely, freedom, fairness and equality) should be part of the criteria of reasonableness, there is no principled reason for stopping there as opposed to going on to include certain perfectionist axioms too.[56] Three plausible candidates for such

[54] Gaus, G., "Liberal Neutrality: A Compelling and Radical Principle", in S. Wall and G. Klosko (eds), *Perfectionism and Neutrality: Essays in Liberal Theory* (Lanham, MD: Rowman and Littlefield, 2003), p. 155.
[55] Ibid. See Simmons, J., *On the Edge of Anarchy: Locke, Consent and the Limits of Society* (Princeton: Princeton University Press, 1993).
[56] The thought here is similar to that in Billingham, P., "Liberal Perfectionism and Quong's Internal Conception of Political Liberalism". An important difference, however, is that

perfectionist axioms are: moral, intellectual and artistic excellence. Contrary to what political liberals often suggest, perfectionist ideals such as moral, intellectual and artistic excellence are not radically indeterminate. They function in roughly the same way as neutral ideals such as freedom and fairness insofar as there is a fairly stable set of core or paradigmatic cases of moral, intellectual and artistic excellence as well as a range of borderline cases in which it is unclear whether or not these ideals are exemplified.

Much more could be said about these perfectionist axioms. What do they each involve? Why these particular axioms rather than some other perfectionist axioms? How do the perfectionist axioms relate to the liberal axioms? Do they all have equal weight? Or should, say, a policy that promotes equality at the cost of some artistic excellence take precedence over a policy that promotes artistic excellence at the cost of some equality? These are some of the kinds of questions that specific versions of political perfectionism will need to answer. But for our purposes, which are more general and illustrative, the important point is this: someone who is reasonable accepts the value and importance of moral, intellectual and artistic excellence. Perfectionist laws and policies such as state subsidies for art galleries would thus be expected pass the test of public justification because a person who rejects the intrinsic value of artistic excellence would count as unreasonable. Her veto has no normative force here — just as the veto of someone who rejects freedom, fairness or equality has no normative force — because public justification is a matter of justification not to actual citizens but to suitably idealized (i.e. reasonable) citizens.

Billingham gives no positive reason for thinking that there is perfectionist content to the idea of the reasonable. He does not endorse or argue for the "perfectionist internal conception" and only postulates it as a critical tool to contend that Quong lacks grounds for preferring antiperfectionist over perfectionist definitions of reasonableness. By contrast, in the following paragraphs I seek to develop a positive reason for thinking that reasonableness includes perfectionist content. This is important because (as I go on to explain in the main text) until some reason has been given for thinking that the perfectionist definition of reasonableness is the right definition, antiperfectionists can legitimately say in response to Billingham that pointing to the *mere possibility* of a "perfectionist internal conception" is too undermotivated and ad hoc an argumentative strategy to cast real doubt on the more deeply theorized definitions of Rawls and Quong. A second and related (though more cosmetic) difference is that Billingham makes reasonableness conditional on acceptance of fairly specific perfectionist judgements and claims (e.g. "the existence of duties of autonomy-promotion"; ibid., p. 105), whereas it strikes me as more systematic and plausible to think about the criteria of reasonableness at a higher level of generality and thus in terms of a handful of core axioms.

The vital question that arises here is: Is this the right justificatory constituency? What justifies the claim that reasonableness is conditional not only on acceptance of liberal axioms but on acceptance of perfectionist axioms too? In particular, should those who reject the value of artistic excellence really be excluded as "unreasonable"? Without a clear answer to these questions, the strategy of baking perfectionist content into the justificatory constituency looks unmotivated and gerrymandered — and thus unconvincing in comparison to the more deeply theorized accounts of Rawls and Quong.[57]

To answer this question, I propose a novel perfectionist conception of society. Before developing this conception, though, some context is necessary. For Rawls, the bounds of reasonableness are set by "a fundamental organizing idea within which all ideas and principles can be systematically connected and related. This organizing idea is of society as a fair system of social cooperation between free and equal persons".[58] But what, in turn, justifies this organizing idea of society? There are disagreements about this. On some interpretations of Rawls, this conception of society is justified in virtue of its embeddedness within the public political culture of liberal democratic societies. Thus our theorizing "starts from this idea [of society as a fair system of social cooperation between free and equal persons] as one of the basic intuitive ideas which we take to be implicit in the public culture of a democratic society".[59] Quong, however, and rightly in my view, rejects this way of proceeding (as well as this way of interpreting Rawls) because "political liberalism is a justificatory project, not an interpretive one, and so it need not take certain facts about our public culture as static or given".[60] Why should we accord normative authority to widely shared but potentially mistaken beliefs and understandings simply because they happen to be firmly embedded within the public political culture of liberal democratic societies? Instead, Quong suggests (and interprets Rawls as suggesting) that we should view society as a fair system of social cooperation between free and equal persons because doing so leads to the most coherent and compelling overall view — the view that is most "congenial to our most

[57] I thank a reviewer for pressing me to address this concern.
[58] Rawls, J., *Political Liberalism*, p. 9.
[59] Rawls, J., "Justice as Fairness: Political not Metaphysical", p. 231.
[60] Quong, J., *Liberalism Without Perfection*, p. 155.

firmly held convictions" — and thus because this conception of society is recommended by the method of reflective equilibrium.[61]

I agree that the bounds of reasonableness are to be set by some fundamental organizing conception of society and that this conception of society must be selected on the basis of reflective equilibrium. But I am doubtful that the political liberal conception of society — a fair system of social cooperation between free and equal persons — is the only, let alone the best, such conception available. In particular, a plausible and arguably superior alternative conception of society is that of society as a fair scheme of social cooperation and a striving for human flourishing between free and equal persons.

This perfectionist conception of society is appealing in a number of respects. First, and like the Rawlsian conception of society as a fair system of social cooperation, the perfectionist conception is articulated at a sufficiently highly level of generality that it could serve as a fundamental organizing idea within which other more specific ideas and principles can be systematically connected and related. Second, and relatedly, the perfectionist conception of society as a fair scheme of social cooperation and a striving for human flourishing between free and equal persons has internal plausibility. Even before we consider whether the perfectionist conception of society yields verdicts that match our considered judgements, that is, it expresses an ideal that is plausible and attractive in its own right. Third, something akin to this perfectionist conception of society has pedigree within perfectionist thought and has been defended by historically significant perfectionists such as Aristotle, who holds that "the political fellowship must…be deemed to exist for the sake of noble actions, not merely for living in common".[62]

Indeed, the perfectionist conception of society is arguably superior to the Rawlsian conception because it coheres with more of our most firmly held convictions. For while both conceptions recognize that citizens are free and equal and that society is a just and cooperative endeavor, the conception being proposed here is in addition sensitive to certain distinctly perfectionist convictions shared by a good many citizens in modern liberal democracies.

[61] Rawls, J., *Political Liberalism*, p. 8.
[62] Aristotle, *The Politics*, trans. H. Rackham (Cambridge, MA: Harvard University Press, 1990), p. 219.

Most ordinary people, after all, are not antiperfectionists. They see nothing wrong with state support for the arts. Indeed, if the recent reaction to the burning of Notre Dame is anything to go by, most people believe that it is *incumbent* on states to take decisive steps to preserve a society's historic treasures, architectural wonders, and cultural heritage. Many deeply held perfectionist convictions also surface in the form of fears and lamentations. It is fairly common, for instance, to hear individuals decry the coarsening and degradation of public culture — a public culture, so the complaint goes, that abounds in advertisements that foster acquisitiveness and direct energies towards the accumulation of material possessions, in images and representations of women that define beauty in terms of sexual availability, and in vapid reality television that glamorizes vulgarity and consumerism. Of course, and as has just been noted, the mere fact that these perfectionist convictions are common among citizens in liberal democracies does not by itself bestow them with any particular normative authority. But insofar as these are important and compelling ideas that survive the auditing process of reflective equilibrium, it is an advantage of the perfectionist conception of society as a fair scheme of social cooperation and a striving for human flourishing that it coheres with and accommodates these distinctly perfectionist aspects of ordinary moral and political thought. In this sense, the Rawlsian conception of society is not so much wrong as incomplete.

One feature of this perfectionist conception of society that might raise eyebrows amongst Rawlsians is that it takes the promotion of human flourishing to be an end or goal of society. Rawls, by contrast, stresses that within his theory "society has no final ends and aims in the way that persons or associations do" — no final ends, that is, other than "constitutionally specified" neutral ends such as the blessings of liberty and the common defence.[63] In this, political liberalism distinguishes itself from perfectionist societies of the past, which "pursued as final ends religion and empire, dominion and glory".[64] The examples of perfectionist ends that Rawls provides, however, are unfortunate because they give the misleading impression that in order to be a perfectionist one must be a theocrat or an imperialist. Political perfectionists can readily accept that society has no final ends or aims of the kind that Rawls mentions.

[63] Rawls, J., *Political Liberalism*, p. 41.
[64] Ibid.

Their only contention is that this does not rule out some other perfectionist ideal that has broader appeal and that is more at home in modern thought, such as human flourishing, from serving as one of the final ends of a liberal democratic society.

We are now in a position to consider the citizen who rejects art. Is such a citizen really unreasonable? A crucial point to note is that, in order to be reasonable, citizens need not organize their life around all or even any of the three perfectionist axioms of moral, intellectual and artistic excellence. Just as one might accept that physical health is good but choose not to go to the gym, a reasonable citizen might accept the validity of the perfectionist axioms but choose for whatever reason not to pursue them in her personal life. But what *would* be unreasonable is to actively reject moral, intellectual and artistic excellence as reason-giving in politics and thus to seek to veto the entry of these ideals into our political life. To do so would run counter to the organizing perfectionist conception of society as a fair scheme of social cooperation and a striving for human flourishing between free and equal persons. The exclusion of the art-hater from the constituency of reasonable citizens, then, is not an ad hoc fix but rather flows from deeper ideas about the nature of society which set the bounds of reasonableness — just as a Rawlsian would justify the exclusion of free riders from the constituency of the reasonable on the grounds that they fail to honour the idea of society as a fair system of social cooperation.

Naturally, there will be various objections to this approach to public justification in general and to the public justification of perfectionism in particular. But perhaps the deepest worry is that, by idealizing so heavily and by building so much content into the definition of reasonableness, the public justification apparatus simply becomes redundant.[65] After all, if laws and policies are permissible when they are acceptable to reasonable citizens, and if

[65] For this charge as applied to Quong's internal conception of political liberalism, see Van Schoelandt, C., "Justification, Coercion and the Place of Public Reason", *Philosophical Studies* 172 (2015), pp. 1037-41. As this illustrates, the redundancy objection is a general objection to heavily moralized accounts of reasonableness, not a specific objection to accounts of reasonableness that are heavily moralized along perfectionist lines. In other words, there is nothing *distinctively* redundant about moving from perfectionist (as opposed to antiperfectionist) axioms to perfectionist (as opposed to antiperfectionist) outcomes. So at worst, if my reply to the redundancy worry fails, (this version of) political perfectionism is in the same boat as (Quong's version of) political liberalism — which is sufficient to show that (this version of) political perfectionism is an important and interesting position in these debates.

reasonable citizens are defined in terms of endorsement of certain perfectionist axioms, then all the moral work is being done by those perfectionist axioms (or rather by the perfectionist conception of society that supplies content to the axioms). But in this case hasn't the distinctive idea of "justification-to" or "acceptability-to" simply dropped out of the picture? What is the point of the constituency move? Why not cut out the middleman and just derive perfectionist outcomes directly from the perfectionist axioms (or from the perfectionist conception of society)?

At some level, I think this charge is accurate. But it is unduly pessimistic. For while it is indeed possible to frame this objection in terms of redundancy — public justification is a "spare wheel"; this is a "victory by stipulation"; we have "definitional fiat" from start to finish — there are also more charitable ways of understanding and describing what is going on. In particular, I view public justification essentially as a heuristic device — a framework for setting out in an attractive and systematic way certain basic thoughts and commitments, and not something that was ever expected to exist, let alone "do moral work", at the morally fundamental level. Public justification is meant to model or represent certain deep normative commitments, and thus to say that some policy is justified to all reasonable persons is no more, and no less, than to say that this policy can be derived from the moral content internal to idea of society as a fair scheme of social cooperation and a striving for human flourishing between free and equal persons. Since derivations in political philosophy, unlike those in mathematics, are not an exact science, the point of introducing the constituency move is to make things as vivid and pellucid as possible. So, while strictly speaking dispensable, the process of considering whether a given policy could be reasonably rejected helps us to focus intuitions, to organize ideas, and to make more transparent derivations and extrapolations.

Of course, there are limits to how far this strategy can go whilst remaining faithful to the basic public reason motivations.[66] It would clearly make a mockery of the public justification apparatus if we held a view such as "Christianity: Political not Metaphysical", according to which reasonableness is conditional on acceptance of Christian doctrine, where this conception of the constituency of reasonable citizens is in turn justified by appeal to a Christian

[66] I thank Tom Sinclair for pressing me to be clearer about this.

conception of society. A view that defines "reasonable disagreement" in a way that bears no relation at all to the folk concept is not aptly characterized as a "public reason" view. But I see no reason to think that these limits of fidelity are necessarily coextensive with liberal neutrality. A moderately perfectionist view of the kind that I propose — according to which members of the justificatory constituency endorse both liberal axioms (such as freedom, equality and fairness) and perfectionist axioms (such as moral, intellectual and artistic excellence) — would still leave room for the fact of reasonable disagreement to play a role (such as with respect to religious matters), albeit a more circumscribed role than it plays within political liberalism, and so such a view could plausibly be regarded as true to the spirit of the public reason project and faithful to its underlying motivations.

The foregoing analysis is somewhat involved. It sides with Rawls and Quong on some points: that public justification must involve considerable idealization lest it end up being "political in the wrong way".[67] But it diverges from them on other points: on the view presented here, the bounds of reasonableness are set by the fundamental perfectionist conception of society as a fair scheme of social cooperation and a striving for human flourishing between free and equal persons, which in turn implies that reasonableness is conditional on acceptance not just of liberal axioms (namely, freedom, fairness and equality) but also perfectionist axioms (namely, moral, intellectual and artistic excellence).

But, as stressed at the beginning of this section, this analysis is only illustrative: it is intended to give a sense of how a political perfectionist might go about publicly justifying perfectionist ideals. Political perfectionism can come in many stripes and can be plausibly combined with *other* accounts of public justification. For instance, some favour a more empirical approach to public justification, one which ties the normativity of public justification more closely to the normativity of consent.[68] On these views, public justification requires justification to citizens pretty much as they actually are, with only minimal idealization. Now, as suggested above, I am sceptical about such an approach because it unduly accords normative authority to the potentially unjust beliefs

[67] Rawls, J., *Political Liberalism*, p. 40.
[68] See, e.g., Klosko, G., "An Empirical Approach to Political Liberalism", in S. Young (ed.), *Political Liberalism: Variations on a Theme* (Albany, NY: State University of New York Press, 2004), pp. 129-48.

that prevail in existing societies and thus is political in the wrong way. But what does seem true is this: *if* the empirical approach is the way to theorize public justification, *then* a good deal of perfectionism will pass the public justification test.[69] To see this, note first that since it is doubtful that there is anything (even at a high level of generality) that *all* real-life citizens would endorse, this kind of view (if it is to avoid the anarchist challenge) is best understood as requiring widespread endorsement rather than unanimity.[70] But then, while many abstract neutral values such as freedom and equality are widely endorsed by citizens in modern democratic societies, the same is true of many abstract perfectionist ideals such as human flourishing and moral, intellectual and artistic excellence. Descending to a more specific level, it seems empirically likely that very many people in actual societies — at least as many as would endorse any comparably specific neutralist claim about, say, the level of taxation — would agree that an individual wholly consumed by drug addiction does not live a meaningful and worthy life. Thus alternative accounts of public justification, according to which the content of public reason is a matter of empirical investigation rather than moral construction, also seem prima facie congenial to political perfectionism.

The question of how much perfectionism passes the public justification test thus depends largely on how that test is understood. But while we cannot settle in advance the question of just how perfectionist political perfectionism would be — because it is a line of inquiry that emerges only once political perfectionism is recognized as occupying an important part of the conceptual space and thus it is a line of inquiry that will have to be addressed in due course by specific political perfectionisms and not at this level of abstraction — we do seem to have good enough reason to expect an appreciable amount of perfectionism to pass the test of public justification. Public justification, then, appears to be a versatile and attractive framework for setting out perfectionist thought, and so consideration of political perfectionism helps us to rescue public justification from liberal neutrality.

5 Is political perfectionism too mild?

[69] Klosko concedes this point in "Reasonable Rejection and Neutrality of Justification", in S. Wall and G. Klosko, *Perfectionism and Neutrality*, p. 189 n. 44. See also Quong, J., *Liberalism Without Perfection*, pp. 150-1.
[70] See Klosko, G., "An Empirical Approach to Political Liberalism", p. 132.

At this point, political liberals may be inclined to respond as follows:

> Political perfectionism is too mild to be interesting. For us, neutrality (which says the state cannot promote the good life) was only ever a *corollary* of public justification (which says the state can only promote what can be publicly justified). Neutrality follows from public justification in conjunction with the assumption that questions of the good life are always subject to reasonable disagreement and so can never be publicly justified. The deepest commitment of political liberalism — its "moral lodestar"[71] — is thus actually public justification, not neutrality per se: the aim is less to ensure that laws and institutions are *neutral* and more to ensure that they are *publicly justified*. What political perfectionism helps us to see is the falsity of the assumption that questions of the good life are always subject to reasonable disagreement. So we are perfectly happy to drop neutrality and to introduce mild forms of perfectionism into politics when those perfectionist interventions can be publicly justified. If some anodyne ideals of the good life ("love and friendship") pass the test of public justification, then that is fine by us. Indeed, Rawls even said as much when he affirmed that states could legitimately act on the basis of a *thin* theory of the good that includes certain all-purpose primary goods that "every rational man is presumed to want" regardless of one's life plan — goods such as health, wealth and self-respect.[72] In other words, since political liberals were never really as committed to the *neutralist* (or *antiperfectionist*) strand of their view as to its *political* (or *contractualist*) strand, political perfectionism preserves all that political liberals truly care about and so is not a view that we would find especially troubling or objectionable.[73]

Since this is a semi-sociological claim about what political liberals would and would not find troubling, it is worth noting that neutrality — as distinct from public justification — *does* figure centrally in the thinking and rhetoric of many

---

[71] Macedo, S., *Liberal Virtues: Citizenship, Virtue, and Community in Liberal Constitutionalism* (Oxford: Clarendon Press, 1990), p. 78.
[72] Rawls, J., *A Theory of Justice: Revised Edition* (Cambridge: Harvard University Press, 1999), p. 54.
[73] I thank Paul Billingham, Thomas Sinclair and Anthony Taylor for putting to me, in different ways, the sort of worry contained in this paragraph.

political liberals, and I see no reason not to take these theorists at their word. For political liberals, neutrality is not incidental: it does not play second fiddle to public justification, and is an important commitment in its own right.

Many have taken neutrality to be definitive of political liberalism. According to Dworkin, for instance, "liberalism takes, as its constitutive political morality," the requirement of "official neutrality amongst theories of what is valuable in life".[74] For Ackerman, it is essential to political liberalism that no citizen may assert that "her own ideal of the good life is worthy of special endorsement by the political community as superior to the ideals affirmed by others".[75] Similarly, Quong defines political liberalism in terms of neutrality and states that the "main thesis" of his book is that "the government should remain *neutral* on the issue of the good life, and restrict itself to establishing the fair terms within which citizens can pursue their own beliefs about what gives value to their lives".[76]

Others take neutrality, if not to be definitive of political liberalism, then to be an important and distinct commitment. Rawls holds that states are to refrain from "evaluat[ing] the relative merits of different conceptions of the good…once it is supposed that they are compatible with the principles of justice".[77] To attribute to states the function of "develop[ing] human persons of a certain style and aesthetic grace" — a function that political perfectionists can unapologetically affirm — would thus be, says Rawls, to "*drastically alter*" the shape of his theory, according to which "the human perfections are to be pursued within the limits of the principle of free association".[78] For Gaus, "liberal neutrality" is a "compelling and radical principle", one that prohibits using state power to make others "more perfect in our eyes".[79] The state, he elsewhere says, has no business telling people how to live their lives: citizens "have no reason to submit their disputes about…the good life or good beers to adjudication: on these matters it is entirely reasonable to object that no public judgement is required, and so the use of state power — even in the service of

[74] Dworkin, R., "Liberalism", in S. Hampshire (ed.), *Public and Private Morality* (Cambridge: Cambridge University Press, 1978), p. 127, 142.
[75] Ackerman, B., "Neutralities", in R. B. Douglass et al (eds), *Liberalism and the Good* (London: Routledge, 1990), p. 37.
[76] Quong, J, *Liberalism Without Perfection*, p. 2 (emphasis in original).
[77] Rawls, J., *A Theory of Justice*, pp. 80-1.
[78] Ibid., p. 289 (emphasis added).
[79] Gaus, G., "Liberal Neutrality: A Compelling and Radical Principle", p. 162.

reasonable views — is illegitimate".[80] For Waldron, neutrality is an important ideal with a "rich heritage" within the liberal tradition, stretching back through John Stuart Mill's *On Liberty* to John Locke's *Letter Concerning Toleration* in which Locke argues that the state should concern itself only with civil matters of life, liberty and property, and not with the "care of souls" which should be left to the church.[81]

Beyond these explicit statements of neutrality and antiperfectionism, it is also revealing to note that, when political liberals have worked out the principles and arrangements that they believe would pass the test of public justification, they have always developed *political* principles, about taxation and property rights for example, rather than *moral* and *ethical* principles about virtue and human flourishing. The antiperfectionist strain in political liberal thought thus seems to run deeper than the political liberal imagined above would have us believe, and it exerts an influence on the kind of outcomes that political liberals tend to derive and defend.

In these and other passages, there is overwhelming textual evidence that political liberals have a special commitment to neutrality as such, in addition to, and as distinct from, their commitment to public justification. At the deepest level, then, political liberals still seem to sympathize with the picture of politics according to which "the state should not reward or penalize particular conceptions of the good life but, rather, should provide a neutral framework within which different and potentially conflicting conceptions of the good can be pursued".[82] So, far from being too mild, it appears that most political liberals would find anathema the perfectionist's suggestion that one of the appropriate functions of the state is to favour and promote ideals of the good life that it judges to be excellent, worthy or virtuous and to discourage ways of life that it judges to be debased, shabby or empty — this regardless of whether these value judgements can be publicly justified.[83]

---

[80] Gaus, G., "Reasonable Pluralism and the Domain of the Political: How the Weaknesses of John Rawls's Political Liberalism can be Overcome by a Justificatory Liberalism", *Inquiry* 42 (1999), p. 281.
[81] Waldron, J., "Legislation and Moral Neutrality" in his *Liberal Rights: Collected Papers 1981 – 1991* (Cambridge: Cambridge University Press, 1993), p. 143.
[82] Kymlicka, W., "Liberal Individualism and Liberal Neutrality", *Ethics* 99 (1989), p. 883.
[83] Similarly, political liberals would generally be uncomfortable with religious establishment, even if it could be publicly justified, as in Kramer's case of Convivia in which Buddhism is officially established in order to enhance public-spiritedness and harmoniousness — a

It is true, of course, that Rawls speaks of certain "ideas of the good", such as the "great social good" of "establishing and successfully conducting reasonably just…democratic institutions over a long period of time".[84] He also believes that believes that the state may act in pursuit of "thin" goods such as health, wealth and self-respect — the kind of judgements about the good life that any political theory will need to make in order to get off the ground. Rawls's difference principle, for instance, assumes that it is possible to identify the least well-off members of society, and such identifications will require some thin theory of the good in terms of which well-being can be assessed.[85]

But three important differences between Rawlsian political liberalism and political perfectionism should be noted here.[86] First, not all political perfectionist policies will be as thin as health, wealth and self-respect. Some will be uncomfortably thick for political liberals' liking. That "the ceaseless repetition of the same monotonous task all day six or seven days a week"[87] is a meaningless and unworthy form of employment, that "Scrooge's lifestyle" is emptier and morally inferior to a life of generosity and warmth,[88] that someone who is "wise, upright, talented in music and philosophy, and has a good family and a few very good friends"[89] has an excellent and worthwhile life — these are all judgements about the good life that would be shared by reasonable citizens, yet they are not at all the kind of judgements that make their way into Rawls's thin theory of the good, which is "restricted to the bare essentials"[90] and which includes only very generic all-purpose goods and fundamental human interests. Second, Rawlsian primary goods are generally thought of as *instrumentally* good, since they are all-purpose means that enable citizens to realise their various ends. By contrast, political perfectionism recognizes the role in politics for judgements about the *intrinsic* value of certain goods, activities, excellences

---

discomfort that again suggests a commitment to neutrality per se, as distinct from public justification. See Kramer, M., *Liberalism With Excellence* (Oxford: Oxford University Press, 2017), pp. 23-4. I consider cases of publicly justified yet non-neutral policy in greater depth in Chapter 1.

[84] Rawls, J., *Political Liberalism*, p. 204.

[85] Rawls, J., *A Theory of Justice*, p. 348.

[86] I thank a reviewer for pressing me to explain how political perfectionism differs from Rawlsian political liberalism.

[87] Caney, S., "Impartiality and Liberal Neutrality", p. 278.

[88] Ibid.

[89] Chan, J., "Legitimacy, Unanimity, and Perfectionism", p. 13.

[90] Rawls, J., *A Theory of Justice*, p. 348.

and perfections. Third, Rawls's goods and virtues are confined to the political domain: they are political virtues such as civility and fairness, and political goods such as the good of justice.[91] By contrast, political perfectionism can incorporate "ideals of personal character, as well as ideals of familial and associational relationships, and much else that is to inform our conduct, and in the limit to our life as a whole" — perfectionist ideals that Rawls considered incompatible with political liberalism and that political liberals since him have tended to eschew.[92] For these three reasons, Rawlsian "ideas of the good" are distinct from perfectionist ideals of human flourishing and excellence.

[91] See Rawls, J., *Political Liberalism*, pp. 194-5, 201-6.
[92] Ibid., p. 13.

Do the Reactive Attitudes Justify Public Reason?

It is a central tenet of much contemporary liberal political philosophy that the state must abide by a principle of public justification. According to this principle, the laws and institutions of society must be in some sense justifiable to all reasonable citizens. As Waldron puts this idea, "liberals demand that the social order should in principle be capable of explaining itself at the tribunal of each person's understanding".[1] For Rawls, the principle of public justification primarily implied that important political decisions must be justified in terms of shared public reasons such as fairness and equality, and not by appeal to controversial religious and ethical doctrines.[2] But since then it has been argued that the implications of this principle are even more radical and far-reaching, and that the public justification principle calls into question the legitimacy of much egalitarian liberalism and in fact tilts towards the minimal state.[3]

But the principle of public justification is by no means self-evident. So political liberals and public reason liberals need to provide a justification (or grounding or moral foundation) for public justification. They need, that is, to answer the question: Why reason publicly? Why should we care about justification to all reasonable citizens? Why not simply implement the political rules and institutions that are *correct* — or at least that we think most likely to be correct — regardless of whether those rules and institutions could be reasonably rejected by some citizens?

The traditional answer to these questions was that the public justification principle is grounded in respect for persons. Thus Larmore states that "if we try to bring about conformity to a rule of conduct solely by the threat of force [and not by offering reasons that others recognize as valid], we shall be treating persons merely as a means, as objects of coercion, and not also as ends, engaging directly their distinctive capacity as persons".[4] But, in recent years,

[1] Waldron, J., "Theoretical Foundations of Liberalism", *Philosophical Quarterly* 37 (1987), p. 149.
[2] Rawls, J., *Political Liberalism* (New York: Columbia University Press, 1996).
[3] Gaus, G., *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World* (New York: Cambridge University Press, 2011).
[4] Larmore, C., "The Moral Basis of Political Liberalism", *The Journal of Philosophy* 96 (12), p. 607.

and in light of powerful criticisms of the respect-based grounding,[5] a number of other moral foundations for public justification have been proposed, including stability,[6] justice,[7] reciprocity,[8] and civic friendship.[9]

Of these more recent groundings, one of the most interesting, detailed and original is Gerald Gaus's reactive attitudes argument. This argument makes two key claims. The first is that our everyday reactive attitudes of blame and resentment presuppose public justification because it is appropriate to resent someone who violates a moral rule only if they had sufficient reason to accept that rule. The second claim is that these reactive attitudes are essential to social life and interpersonal relationships and are so deeply embedded within us that to renounce them would be "practically inconceivable" (to use the words of Strawson, who introduced the notion of reactive attitudes).[10] Yet, despite its sophistication and novelty, Gaus's argument has received little critical scrutiny.[11]

In this paper, I argue that the argument from reactive attitudes to public justification is unsuccessful. I first spell out in more detail, in Section 1, Gaus's reactive attitudes argument. I then critique the two key premises in Gaus's argument. I argue, in Section 2, that the reactive attitudes do not presuppose public justification because there are plenty of cases in which we appropriately blame and resent agents for violating a rule despite their not having a reason to accept that rule. The rejection of public justification thus need not also imply the rejection of those reactive attitudes of praise and blame that Gaus takes to be essential to social life. Next I argue, in Section 3, that rejecting the reactive attitudes is not, in any case, so inconceivable. By sketching out an alternative picture inspired by recent writing on moral responsibility, I suggest that social life would still be possible, and perhaps even improved, in the absence of the

---

[5] See, e.g., Eberle, C., *Religious Conviction in Liberal Politics* (Cambridge: Cambridge University Press, 2002), pp. 109-51.
[6] Weithman, P., *Why Political Liberalism?* (Oxford: Oxford University Press, 2010).
[7] Quong, J., "On the Idea of Public Reason", in J. Mandle and D. Reidy (eds), *The Blackwell Companion to Rawls* (Oxford: Wiley Blackwell, 2013), pp. 265-80.
[8] Neufeld, B., "Reciprocity and Liberal Legitimacy", *Journal of Ethics & Social Philosophy* 4 (2010), pp. 1-7.
[9] Lister, A., *Public Reason and Political Community* (London: Bloomsbury, 2013).
[10] Strawson, P. F., "Freedom and Resentment", in G. Watson (ed.), *Free Will* (Oxford: Oxford University Press, 2003), p. 81.
[11] The notable exception to this is Taylor, A., "Public Justification and the Reactive Attitudes", *Politics, Philosophy & Economics* 17 (2018), pp. 97-112.

reactive attitudes. Finally, in Section 4, I delve into Gaus's complex views about the relation between the normative and the empirical in order to query whether the reactive attitudes are even the kind of things that could justify public reason. After all, even if the reactive attitudes presuppose public justification, and even if the reactive attitudes cannot be renounced, how could this show anything more than that we, being who we are and how we are, are stuck — regrettably, perhaps — with public justification? Does Gaus's argument justify public reason or instead (and invoking a distinction that may beg the crucial question here) merely explain it?

1 Gaus's reactive attitudes argument for public justification

In this section, I spell out Gaus's argument from reactive attitudes to public justification. I first explain, in Section 1.1, the specific formulation of the public justification principle that Gaus upholds. I then explain, in Section 1.2, why Gaus believes the reactive attitudes presuppose this public justification principle. And in Section 1.3, I explain why Gaus believes the reactive attitudes cannot be rationally renounced. Together, these two premises make up a strong case for the public justification principle.

1.1 Gaus's public justification principle

Gaus holds the following public justification principle:

> A moral imperative "Φ!" in context *C*, based on a rule *L*, is an authoritative requirement of social morality only if each normal moral agent has sufficient reasons to (a) internalize rule *L*, (b) hold that *L* requires Φ-type acts in circumstances *C* and (c) moral agents generally conform to *L*.[12]

We can abstract away from some of the details of this formulation, for nothing in my discussion hangs on them. The clause that I shall focus on is (b): I shall challenge Gaus's defence of the view that a moral imperative is binding only if agents have sufficient reasons to accept the rule. But one distinctive feature of

[12] Gaus, G., *The Order of Public Reason*, p. 263.

Gaus's principle, and a feature that will be highly relevant to our discussion, is his account of what it means to "have sufficient reasons" to accept a rule.[13]

Here Gaus draws a distinction between internalist and externalist accounts of having a reason. On the externalist view, an agent has a reason whenever there is a reason that applies to her. Gaus rejects the externalist picture because it "misconstrues the relation between having a reason and being a rational agent".[14] The externalist view implies, for instance, that Aristotle had a reason to accept particle physics. But, says Gaus, Aristotle had no such reason; indeed, only by ignoring his evidence and by being irrational could he have come to endorse such a reason. When we move from theoretical reasoning to practical reasoning, externalism might seem more plausible. For surely everyone has a reason to avoid genocide? But consider the case of a benighted tribal society that lacks the concept of an ethnic group. Of this case, Gaus says: "to attribute a reason *to* these peoples, to say that they *possessed* a reason not to eliminate other groups, while acknowledging that any justificatory force of this reason is inaccessible to them as reasoning beings is, I think, not only a misuse of language, but undermines the point of discourse about reasons and rationality".[15] It is one thing to say that *there is* a reason to avoid genocide that applies to these tribal societies. But to say that this reason — this "incomprehensible justifying consideration" — is a reason they *have* is, thinks Gaus, "close to bizarre".[16] As he puts it elsewhere, the proper question to ask of such externalists is: "Reasons for whom?"[17]

Instead, Gaus adopts an internalist account of having a reason. This view, of course, need not say that the reasons we have are those that we would immediately affirm at any particular moment. Some moderate idealization of the reasoner is permitted. The point is just that there must be "some not-too-difficult bridge to cross connecting what reasons a person can be said to possess and what the exercise of her rational faculties can lead her to".[18] Thus:

---

[13] See ibid., pp. 232-57.
[14] Ibid., p. 233.
[15] Ibid., p. 235.
[16] Ibid.
[17] Gaus, G., *Justificatory Liberalism: An Essay on Epistemology and Political Theory* (Oxford: Oxford University Press, 1996), p. 35.
[18] Gaus, G., *The Order of Public Reason*, p. 35.

> 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).[19]

Slotting this into the earlier public justification principle, we see that Gaus's view, abstractly stated, is that one should not make an authoritative demand on others unless those subject to the demand have sufficient reason to accept it, in the sense that they would, after a respectable amount of good reasoning, accept the demand. This is the principle that Gaus takes to have profound and far-reaching implications for social and political life and that he seeks to justify through the reactive attitudes argument.

1.2 Why the reactive attitudes presuppose public justification

The term "reactive attitudes" was introduced by Strawson in his seminal essay "Freedom and Resentment".[20] For Strawson, the reactive attitudes are the "attitudes belonging to involvement or participation with others in inter-personal human relationships".[21] Paradigm examples of such reactive attitudes are resentment, indignation and gratitude. What makes these attitudes *reactive*, says Strawson, is that they are "reactions to the quality of others' wills towards us" — to their good will or ill will.[22]

A crucial point made by both Strawson and Gaus is that these reactive attitudes have certain "appropriateness conditions".[23] For instance, while it makes perfect sense to be indignant if my brother bites me, it seems quite inappropriate to be indignant if my *dog* bites me. Perhaps I can in such moments appropriately experience *non-reactive* attitudes such as frustration and annoyance. But insofar as my dog is incapable of *ill will* towards me, it seems misplaced and inappropriate to direct at it the *reactive* attitudes of resentment, blame and indignation.

---

[19] Ibid., p. 250.
[20] Strawson, P. F., "Freedom and Resentment", *Proceedings of the British Academy* 48 (1962), p. 187-211. Reprinted in G. Watson (ed.), *Free Will*, pp. 72-93.
[21] Strawson, P. F., "Freedom and Resentment", in G. Watson (ed.), *Free Will*, p. 79.
[22] Ibid., p. 83.
[23] Gaus, G., *The Order of Public Reason*, p. 206.

Under what conditions, then, are the reactive attitudes appropriate? Insofar as reactive attitudes are responses to the quality of others' wills towards us, we can put this question in terms of quality of will: under what conditions can individuals be understood to possess a good or ill will?

For Gaus, there are two appropriateness conditions. The first is that an individual must be "capable of caring for a moral rule even when it does not promote her wants, ends, or goals".[24] Following Taylor, we can call this the *Capacity Condition*.[25] This condition explains the common belief that young children, psychopaths, dogs and others who lack minimal capacities for moral motivation should not be the targets of blame and resentment (even if it *is* appropriate for parents, orderlies and dog-owners to experience non-reactive attitudes such as frustration).

I shall not contest the Capacity Condition. My worry is with the second condition for the appropriateness of the reactive attitudes, which, again following Taylor, we can term the *Sufficient Reason Condition*.[26] To introduce this second condition, Gaus notes that sometimes the reactive attitudes are inappropriate even when the first condition is met. Consider the case of systematic indoctrination, such as Winston Smith at the end of *1984*. Even though he is a fully-fledged moral agent, capable of putting aside his own wants and acting on the basis of moral reasons, it is still inappropriate to feel resentment and indignation towards Smith when he embraces Big Brother because the relevant moral reasons "are now beyond his comprehension".[27] The Sufficient Reason Condition, then, says that the reactive attitudes are appropriately directed towards an individual who violates a moral rule only if that individual "has sufficient reasons to endorse the relevant rule".[28] To motivate this condition, Gaus imagines making a moral demand on an individual:

> However, I know that she does not see that it is the moral thing to, and suppose I think her lack of appreciation is quite genuine. I demand "Φ!"

[24] Ibid., p. 211.
[25] See Taylor, T., "Public Justification and the Reactive Attitudes", p. 103.
[26] Ibid.
[27] Gaus, G., *The Order of Public Reason*, p. 218.
[28] Ibid., p. 222.

> and she does not see why she is obligated to Φ. She is puzzled that anyone would think Φ is obligatory. If I think this, then again I cannot reasonably feel resentment or indignation that she fails to Φ, any more than I can feel indignation at a four-year-old who is unable to detach himself from what he most wants to do and so steals some favourite candy. She just cannot see how "Φ!" has any internal authority over her.[29]

In short, the reactive attitudes of resentment and indignation presuppose public justification because these attitudes can be appropriately harboured towards an individual who violates a moral rule only if that individual has sufficient reason to accept the rule — which corresponds to clause (b) of Gaus's public justification principle.

1.3 Why the reactive attitudes cannot be rationally renounced

The second key premise in Gaus's argument is that the reactive attitudes of praise and blame cannot be rationally renounced or repudiated. In making this argument, Gaus follows Strawson, who similarly claims that our commitment to the reactive attitudes is "too thoroughgoing and deeply rooted for us to take seriously the thought" that these attitudes should be repudiated.[30] Such a thought, says Strawson, "is, for us as we are, practically inconceivable".[31] As Gaus put this Strawsonian point:

> We are embedded in certain sorts of practices, with certain beliefs and emotions. They form part of the reasons from which we must judge, criticize, and propose changes. A practice such as social morality is deeply embedded in our view of the world; it affects our understanding of interpersonal relations, including love and friendship, and so of what sort of life is worth living. If the presuppositions of our moral practices are so deep a part of the way we see the world, then to renounce the practice would be to renounce most of what we care for and value. But how could we have reason to do *that*? How can we survey all that matters to us and come to the conclusion that our reasons lead us to

[29] Ibid., p. 219.
[30] Strawson, P. F., "Freedom and Resentment, p. 81.
[31] Ibid.

> give it up, by renouncing the view of the world on which our reasons depend? Where would *that* reason come from? It is, I think, as difficult to argue a moral person out of her moral practices as it would be to argue the psychopath into them; given who they are, they do not have reasons to change their view of the world.[32]

This is a powerful passage, but one that is not easy to interpret. In particular, it is not easy to say precisely what makes it the case that we cannot rationally renounce the reactive attitudes. Is the rational non-renounceability of the reactive attitudes a claim about what is *psychologically* possible — in the same way that some hold that, given our innate selfishness, we just cannot abide by the utilitarian requirement that "as between [our] own happiness and that of others" we be "as strictly impartial as a disinterested and benevolent spectator"?[33] Or is it a claim about what is *conceptually* possible — in the same way that some hold that we cannot prove the legitimacy of the laws of logic without presupposing the laws of logic themselves?[34] (I shall say more on this in Section 3.)

2 Reactive attitudes without sufficient reasons

In this section, I contest Gaus's claim that the reactive attitudes presuppose and entail public justification. To do this, I first present, in Section 2.1, some cases in which it is intuitively appropriate to harbour feelings of resentment and indignation towards individuals for violating some moral rule *even though* those individuals did not have sufficient reason to accept the rule. On further inspection, then, our ordinary moral practice of the reactive attitudes does not seem to presuppose public justification. I then consider and reject two possible responses: denying that the agents in these cases really lack sufficient reason to conform to the violated moral rule (Section 2.2), and biting the bullet and insisting that, despite appearances to the contrary, the reactive attitudes really are inappropriate in these cases after all (Section 2.3).

---

[32] Gaus, G., *The Order of Public Reason*, p. 192.
[33] Mill, J. S., *Utilitarianism*, ed. R. Crisp (Oxford: Oxford University Press, 1998), p. 64.
[34] These two readings of Gaus correspond to the two main ways in which Strawson's original essay has been interpreted. See, e.g., McKenna, M., "Where Frankfurt and Strawson Meet", *Midwest Studies in Philosophy* 29 (2005), pp. 166-7.

2.1 Some cases: genocide, animal cruelty, racism, sexism, littering

The most vivid type of counterexample to Gaus's claim that liability to the reactive attitudes presupposes sufficient reason involves cases of monstrous yet rationalized evil. Consider an individual such as Goebbels, who performs monstrously evil actions yet who, after a respectable amount of deliberation, finds that he sees no sufficient reason to refrain from these actions.[35] He is capable of caring about morality — in this sense he satisfies Gaus's Capacity Condition — and sees that there is *some* moral reason against harming others in general and against genocide in particular. But he assigns this moral reason such little weight that it fails to defeat the more important reason (as he sees it) in favour of genocide. So he simply cannot see how the rule prohibiting the murder of some racial or ethnic group has any authority over him. Yet are we seriously to think that, therefore, Goebbels is off the hook — that we cannot appropriately feel resentment and indignation towards him? Assuming that Goebbels "is puzzled that anyone would think" that refraining from genocide is obligatory, Gaus's view has the implausible implication that "I cannot reasonably feel resentment or indignation that [Goebbels] fails to" refrain from genocide.[36]

It is not only at the heights of vice that the Sufficient Reason Condition misfires, however. Though I will primarily focus on the Goebbels case — because it elicits the relevant intuition most strongly — it is only the most extreme instance of a common kind of structure. Suppose that, at a dinner party, an animal rights activist discovers that she is sat opposite a producer of

[35] This case is briefly mentioned in Baccarini and in Taylor: see Baccarini, E., "Having a Reason and Distributive Justice in *The Order of Public Reason*", *European Journal of Analytic Philosophy* 9 (2013), p. 32; Taylor, T., "Public Justification and the Reactive Attitudes", p. 104.

[36] To be clear: Gaus would of course say that Goebbels must be forcibly prevented from enacting his plans. In response to Quong's Carl case — another "agent-type" challenge involving someone with morally abhorrent views — Gaus says "*in extremis* we must act as we must". (See Gaus, G. "The Good, the Bad, and the Ugly: Three Agent-Type Challenges to *The Order of Public Reason*", *Philosophical Studies* 170 (2014), p. 574; see Quong, J., "What Is the Point of Public Reason?", *Philosophical Studies* 170 (2014), pp. 545-553.) But Gaus would say that, when acting *in extremis* against Goebbels or Carl, the reactive attitudes are inappropriate because we cannot justify our use of preventative force to Goebbels or Carl. And this latter claim about the appropriateness of the reactive attitudes is what I contest in this section. My agent-type challenge thus differs from that of Quong because whereas he is concerned with Gaus's move from the public justification principle to liberal outcomes, I am concerned with Gaus's move from the reactive attitudes to the public justification principle.

foie gras — the luxury food made by inserting a 12-inch steel tube into the throat of a goose in order to force-feed it corn mash (a procedure known as gavage). Aware of the controversy surrounding gavage, the foie gras producer proceeds to spend the evening justifying this horrible procedure by reference to the delicious taste of foie gras and its place in traditional French cuisine. In such a situation, must the animal rights activist suspend her indignation simply because the foie gras producer sees no sufficient reason, from within his system of values and beliefs, not to practice gavage? Consider, next, an individual who often tells racist jokes interspersed with terrible racial slurs. He has been called out on these jokes before, but he considers them to be side-splittingly funny and just fails to see how anyone could possibly be offended by them — "it's only a joke", after all, he tends to think. Does the fact that he sees no sufficient reason not to tell racist jokes make it inappropriate for those around him, or those who overhear, to feel outrage and indignation? Or suppose a woman discovers that her job application was rejected due to the employer's sexist implicit bias against résumés with female names. Insofar as the bias was implicit, the employer would likely not have been able to root it out even after a respectable about of good reasoning. Would it therefore be inappropriate for the woman to feel resentment in this case? Consider, finally, the case of a group of littering youth who genuinely do not think it is a problem to throw litter on the ground. Again, is indignation really an unreasonable reaction in this case just because the youth would be puzzled by the objections of those who oppose littering?

Gaus may, of course, respond to the rhetorical question associated with each of these cases — Goebbels, the foie gras producer, the racist joke-teller, the sexist employer, the littering youth — by saying: "But they don't get it. Your demand makes no sense to them. How can you be outraged at them for failing to conform to moral rules that they do not grasp?" But this typically does not make the reactive attitudes seem any less appropriate. Indeed, a natural reply here is to say: "Yes, and that's the whole point! Their failure to see (variously) genocide, gavage, racism, sexism and littering as something they shouldn't do is precisely *why* I am so indignant!" It is natural to feel indignation towards these individuals, in other words, not *in spite of* their inability to see the reasons that apply to them but precisely *because* of it.

Gaus's account of the appropriateness conditions for the reactive attitudes certainly seems to capture *some* part of our ordinary moral practice. It does a good job, for instance, in explaining what makes the reactive attitudes inappropriate in relation to systematically indoctrinated agents such as Winston Smith in *1984*. But having a sufficient reason not to violate a moral rule cannot be a necessary condition of a rule-violator's liability to the reactive attitudes because we often seem to appropriately harbour reactive attitudes towards those who violate a moral rule that they did not have sufficient reason (even after a respectable amount of reasoning) to endorse. The Sufficient Reason Condition is thus false.

2.2 The agents have sufficient reasons

In response, Gaus might want to argue that these are not, in fact, cases where the transgressors lack sufficient reasons not to transgress. Consider the case of the littering youth. Perhaps, at the moment in which they throw litter onto the ground, the youth see no reason not to litter. But Gaus's account of what it means to have a reason involves some degree of idealization: the reasons we have are not the reasons we actually affirm at any particular moment but rather the reasons we would arrive at after a "respectable amount of good reasoning". So perhaps Gaus could say that, once they spend a respectable amount of time thinking about the ethics of littering, the youth would come to see that it is wrong. The same strategy can of course be used in relation to the other cases, including the central case of Goebbels.

In fact, this move — this insistence that, after a respectable amount of good reasoning, the transgressor would see the light of reason — is the response that Gaus has employed on another occasion when presented with the Goebbels-style objection:

> Do we *really* think that Nazi politicians were models of good reasoners — that they reasoned well, albeit from bad commitments? … Students of ideology typically see fascism as held together more by emotion and myth than by reasoned coherence. An ideology that constituted a revolt

> against reason strikes me, at any rate, as a distinctly inappropriate model of a rationally coherent immoral view.[37]

It is certainly possible that the historical Goebbels would not count as engaging in a "respectable amount of good reasoning" and that, if he did, he would have seen the error of his ways — and that, in this sense, Goebbels *did* have sufficient reason to refrain from genocide.

The problem with leaning on this strategy too heavily, however, is that it can only work for a select group of cases, and so it simply postpones the inevitable. Since Gaus's account of having a reason is procedural rather than substantive, there is — as Quong has pointed out,[38] and as Gaus concedes[39] — no principled constraint on the reasons that an agent can end up having. What an agent "has a reason" to do or believe, on Gaus's account, is not a matter of the objective reasons that apply to her: for this is the externalist account of having a reason, discussed in Section 1.1, which Gaus is at pains to reject, which severs the connection between having a reason and being rational, and which would risk the collapse of public justification into ordinary rational justification. Rather, what an agent has a reason to do or believe is a matter of "what the exercise of her rational faculties can lead her to" after a respectable amount of time.[40] So Gaus's account of having a reason cannot foreclose the possibility of an individual having sufficient reasons to perform bad, even grossly immoral, actions — the kind of actions that seem to appropriately elicit the reactive attitudes.[41]

---

[37] Gaus, G., "On Theorizing About Public Reason", *European Journal of Analytic Philosophy* 9 (2013), p. 78.
[38] Quong, J., "What Is the Point of Public Reason?", *Philosophical Studies* 170 (2014), pp. 546-7.
[39] Gaus, G., "The Good, the Bad, and the Ugly", pp. 570-1.
[40] Gaus, G., *The Order of Public Reason*, p. 35.
[41] It is tempting to wonder at this point: So why doesn't Gaus adopt — or what is stopping Gausians from adopting — a substantive rather than procedural account of having a reason? On this view, no one could have engaged in a "respectable amount of good reasoning" if they do not end up endorsing certain substantive moral values, such as equality and fairness. The possibility of an agent like Goebbels — who, after a respectable amount of good reasoning, concludes that he has an undefeated reason to engage in genocide — is thus ruled out by definition. (I thank a reviewer for pressing me on this point.) Gaus raises at least two problems with substantive idealizations of this kind, however. (1) It risks a "victory by definition" (Gaus, G. "A Tale of Two Sets: Public Reason in Equilibrium", *Public Affairs Quarterly* 25 (2011), p. 310). If public reason liberals say that suitably idealized citizens must have sufficient reason to accept liberal politics, and if they then include substantive moral commitments in those idealizations, then all the work is done by the prior moral commitments and the distinctive public-reason-liberal idea of "justification-to" or "acceptability-to" becomes redundant. Here we have not

So the problem with responding to one counterexample after another by pointing out how, in each case, the agent does not spend enough time reasoning is that, still, nothing in Gaus's theory rules out the prospect of an individual who satisfies all that is required for Gaus's procedural account of having a reason, yet who nonetheless arrives at the same morally abhorrent conclusions. And so we are still left with the question: What should we say about *him*? Would not the reactive attitudes be appropriate here? The short-term strategy, then, of quibbling with details of how cases are described fails in the long term: for it does not face up to the fact that the reactive attitudes seem to be appropriate in the case of individuals who even after a respectable amount of good reasoning have grossly immoral views.

Consider, then, a rationally upgraded version of Goebbels — Goebbels+, say. Like Goebbels, Goebbels+ believes in the permissibility of killing entire racial and ethnic minorities. But, unlike the historical Goebbels, Goebbels+ does not base these views on emotion and myth. Instead, he simply assigns enormous weight to his country's racial purity. The weight he assigns to racial purity is so great that, even though he recognizes pro tanto reasons not to harm others, these reasons against harm are outweighed by reasons to promote racial purity. And so, after a respectable amount of time, the exercise of his rational faculties leads him to formulate and carry out genocide. Goebbels+ knows that some people consider him immoral, but he just cannot understand why they do so. In this case, Goebbels+ lacks a sufficient reason (in the Gausian internalist sense) to desist from killing racial and ethnic minorities. Yet he is surely still liable to the reactive attitudes for doing so.

2.3 The reactive attitudes really are inappropriate here

Even if Gaus can contest the details and rational coherence of the reasoning process of the actual, historical Goebbels, the interesting and difficult question

---

solved the problem of how liberal institutions can be justified in the face of reasonable pluralism so much as defined the problem away. The aim of the public reason project, says Gaus, is "to commence with minimal ideas" that "do not at the outset give us all we seek at the end of the day" (*The Order of Public Reason*, p. 20). (2) Insofar as substantive reasons might be inaccessible to certain agents (like Goebbels), the substantive account of having a reason is vulnerable to the same problem as the externalist account: namely, it implies that agents can have — can *possess* — (substantive) reasons that are absolutely inaccessible to them, and thus it severs the connection between having a reason and being a rational agent.

is this: What should we say about Goebbels+, who engages in a respectable amount of good reasoning yet who nonetheless concludes that he is morally required to perform genocide?

Again, this is only the most extreme case. We can imagine rationally upgraded versions of the other agents mentioned in 2.1 (foie gras producer+, racist joke-teller+, sexist employer+, littering youth+). These are agents who, by stipulation, have engaged in whatever amount of good reasoning Gaus wishes to see, yet who are still repellent in their thoughts and action. What should we say about foie gras producer+, who fails to see what is wrong with the practice of gavage? What should we say about racist joke-teller+, who is puzzled that anyone would think his jokes are wrong and who interprets others' offense as a manifestation of "Generation Snowflake" political correctness? What should we say about sexist employer+ whose implicit biases against women lead him to give preference to the job applications of men? What should we say about littering youth+, who just cannot see what all the fuss is about? Surely the reactive attitudes are still appropriate in these cases?

Here Gaus must, I think, bite the bullet and insist that, despite appearances to the contrary, the reactive attitudes of blame and indignation are inappropriate in these cases, assuming that the agents really do see no sufficient reason to comply with the moral rules that are contravened by their actions. Here, that is, Gaus must simply say: we cannot feel morally outraged at Goebbels+ for failing to conform to a moral rule against genocide that makes no sense to him after rational reflection on the matter and that he regards as a kind of indefensible taboo rather than a strong moral requirement applicable to him.[42] But this is highly counterintuitive and thus rather unsatisfying. If the reactive attitudes of resentment and indignation are appropriate anywhere, they are surely appropriate here in the case of a racist genocidal killer. So is there anything more that Gaus can say in support of the claim that feelings of indignation towards Goebbels+ are inappropriate? In this subsection I consider and reject two ways in which Gaus might support this claim, namely by providing an error theory and by pressing the analogy with Winston Smith.

[42] See ibid., p. 221.

We begin with the error-theory strategy. Gaus could seek to provide an error theory that explains why, if the reactive attitudes are inappropriate in the case of Goebbels+, we had the contrary intuition. To do this, Gaus could point out that there is no reason to deny that Goebbels+ can elicit plenty of *non-reactive* attitudes, such as frustration and annoyance. Nor need he deny that the appropriateness of certain *kinds* of reactive attitude in this case, such as a resentment that is directed at the racialized structure of society and at other related structural injustices that encourage beliefs about racial purity. Indeed, Gaus can even accept a certain kind of reactive attitude towards Goebbels+ himself: not for performing genocide (the moral rule against genocide, after all, is beyond his comprehension), but for allowing himself to become the kind of person who could fail to grasp the wrongness of genocide.[43] So on this error-theoretic proposal, when forming intuitions about whether it is appropriate to feel reactive attitudes towards Goebbels+ for engaging in genocide, we are sensitive to other questions in the picture — for instance, whether it is appropriate to feel non-reactive attitudes about the performance of the genocide, or whether it is appropriate to feel reactive attitudes towards Goebbels+ for allowing himself to be the way he is —which are similar to, and thus conflatable with, but strictly distinct from, the question of whether it is appropriate to feel reactive attitudes towards Goebbels+ for his acts of genocide. Once we realize this, so Gaus might argue, we become less confident that our intuitions are really tracking what they are meant to.

My worry with this strategy is that, like all error theories, it is too speculative to be convincing. For my own part, the relevant intuition seems suitably fine-grained: when contemplating the case of Goebbels+, the indignation I feel is directed not (or not merely) at the structure of society, nor even at Goebbels+ himself for letting himself become the kind of person he is, but rather, and quite specifically, at Goebbels+ *for doing what he does*. It is Goebbels+'s act of genocide, and not the structure of society or Goebbels+'s obliviousness to the wrongness of genocide, that is central to the sense of indignation here. To see this, compare Goebbels+ with Distracted Goebbels+. Distracted Goebbels+ lives in the same structurally unjust society as Goebbels+, and Distracted Goebbels+ similarly sees nothing wrong with conducting acts of genocide. Yet, unlike Goebbels+, Distracted Goebbels+ is too occupied with reading novels to act on

[43] Ibid.

his beliefs. Now, while one might reasonably feel indignation towards both individuals for being the way they are — for being the kinds of people that *would* conduct genocide — is it not reasonable to feel some special additional indignation towards Goebbels+ for actually doing the terrible deeds? It seems highly plausible that Goebbels+, in carrying out acts of mass murder, disgraces himself and makes himself liable to the reactive attitude in a way that Distracted Goebbels+, in virtue of not perpetrating any such acts, is not. This suggests that our reactive attitudes are tracking what they are meant to — namely, Goebbels+'s violation of a moral rule against genocide — and not other factors, such as being the kind of person who sees nothing wrong with genocide, because these other intuition-confounding factors are controlled across Goebbels+ and Distracted Goebbels+.

A second strategy that Gaus could employ in order to support the claim that feelings of indignation are inappropriate in the case of Goebbels+ is to appeal to the Winston Smith analogy.[44] Gaus could, that is, argue that Goebbels+ is relevantly analogous to Winston Smith and so is no more liable to the reactive attitudes for conducting genocide than is Winston Smith for embracing Big Brother. Both, after all, while capable of caring about morality, lack sufficient reasons to endorse the moral rule contravened by their action. If Winston Smith *genuinely* isn't to blame, then maybe Goebbels+ isn't either?

---

[44] An alternative yet related strategy would be to analogize Goebbels+ with a *psychopath*. Isn't Goebbels+ cognitively and emotionally analogous to a psychopath, and aren't reactive attitudes inappropriate in the case of psychopaths? (I thank Henrik Dahlquist and a reviewer for putting this point to me.) However, there are two problems with this strategy. First, there is a morally significant difference between Goebbels+ and a psychopath — namely that the former, but not the latter, satisfies Gaus's Capacity Condition for the appropriateness of the reactive attitudes. Goebbels+ is capable of caring about morality even when it does not advance his ends or goals, whereas a psychopath is not. Unlike a psychopath, Goebbels+ is a moral agent; he just assigns moral considerations insufficient weight and so he reasons his way to immoral conclusions. This explains why we can resent Goebbels+ but not the psychopath, and this is why Winston Smith (who meets the Capacity Condition) is the better analogy here. (The objector might press: But how morally significant is the difference between Gobbels+, who assigns morality little weight, and the psychopath, who assigns morality no weight? Does the bare possession of moral capacity matter enough to make Goebbels+ liable to the reactive attitudes in a way that the psychopath is not? This is an interesting line of thought, but it goes beyond the scope of this paper. Suffice it to say that this response is not available to Gaus — at least not without abandoning or significantly modifying his first appropriateness condition for the reactive attitudes.) Second, for readers who cannot see Goebbels+ as anything but a psychopath, the rationally upgraded versions of the less extreme (and thus clearly non-psychopathic) agents may be the better cases: What should we say about racist joke-teller+, who utters horrible and degrading racial slurs in the course of his jokes? Is it not reasonable to feel indignation at this, despite (or perhaps in part because of) his insensitivity to the wrongness of such utterances?

However, while it is true that a systematically manipulated agent such as Winston Smith and a mass murderer such as Goebbels+ are analogous in the sense of lacking a sufficient reason to endorse some moral rule, they are also disanalogous in a number of morally important respects. The first and most obvious disanalogy is that Winston Smith is subject to intense psychological manipulation by another agent who intentionally and wrongfully causes him to have the views he has, whereas Goebbels+ is not. Perhaps, in other words, Winston Smith is not liable to the reactive attitudes, whereas Goebbels+ is, because manipulation is irreducibly exonerating. Second, Winston Smith's embrace of Big Brother is completely out of character (until the final scene, *1984* sees Winston Smith constantly rebelling against Big Brother), whereas Goebbels+'s genocidal activities are not. More precisely, and as Sher would put this point, Winston Smith's action is not, but Goebbels+'s is, traceable to "the interaction of some combination of his constitutive attitudes, dispositions, and traits".[45] Third, Winston Smith's embrace of Big Brother at the first order is in conflict with his second-order desire to destroy Big Brother, whereas Goebbels+'s first-order desires for racial purity are plausibly aligned with his second-order desires. Fourth, given natural assumptions about the severity of Big Brother's intensive manipulation of Winston Smith, he loses the capacity to be appropriately reasons-responsive, whereas Goebbels+ has such a capacity. Fifth, in arriving at his views Goebbels+ plausibly exhibits "epistemic vices" such as "overconfidence, arrogance, dismissiveness, laziness, dogmatism, incuriosity, self-indulgence, contempt", whereas Winston Smith does not.[46] In light of these five morally significant differences, the analogy between Winston Smith and Goebbels+ is weak and should not dent our confidence that Goebbels+ is liable to the reactive attitudes.

It will not have gone unnoticed that, in sketching out the disanalogies between Winston Smith and Goebbels+, I have gestured towards certain leading theories of moral responsibility.[47] Each of these theories can be understood as an

---

[45] Sher, G., *Who Knew? Responsibility Without Awareness* (Oxford: Oxford University Press, 2009), p. 88.

[46] FitzPatrick, W., "Moral Responsibility and Normative Ignorance: Answering a New Skeptical Challenge", *Ethics* 118 (2008), p. 609.

[47] The five disanalogies in the previous paragraph correspond, respectively, to the following discussions: Mele, A., *Free Will and Luck* (Oxford: Oxford University Press, 2006), pp. 141-44; Sher, G., *Who Knew? Responsibility Without Awareness*; Frankfurt, H., "Freedom of the Will and the Concept of a Person", *Journal of Philosophy* 68 (1971), pp. 5-20; Fischer, J. and Ravizza, M., *Responsibility and Control: A Theory of Moral Responsibility* (Cambridge: Cambridge

account of when the reactive attitudes of praise and blame are appropriate. Here it is important to stress, though, that I am not taking a stand on the question of *which* of these various disanalogies grounds the difference in liability to the reactive attitudes between Goebbels+ and Winston Smith. Another way of putting this is that I am not endorsing any particular *alternative* appropriateness condition for the reactive attitudes to replace the Sufficient Reason Condition (e.g. the reactive attitudes are appropriate only if an agent is unmanipulated; only if an agent's actions are traceable to some interaction of his constitutive attitudes and dispositions; only if the agent's desires mesh in a certain way; and so on) — an alternative condition that would explain why Goebbels+ is, and why Winston Smith is not, liable to the reactive attitudes.[48]

But nor do I need to endorse an alternative condition. For my objection to Gaus's Sufficient Reason Condition is case-based rather than principle-based. The objection is not that the Sufficient Reason Condition is false because some alternative condition (one that refers to manipulation, epistemic vices, the constitutive elements of the deep self, reasons-responsiveness, etc.) better captures and explains our general practices of praise and blame. Rather, the objection is that the Sufficient Reason Condition is false because it implies, highly counterintuitively, that it is not appropriate to resent genocidal mass murderers like Goebbels+ (nor other agents, such as foie gras producer+, racist joke-teller+, sexist employer+, and littering youth+). This counterintuitiveness does not subside once we consider Winston Smith, who is analogous to Goebbels+ in one respect yet disanalogous in many other respects. So, to reject

---

University Press, 1998); FitzPatrick, W., "Moral Responsibility and Normative Ignorance: Answering a New Skeptical Challenge".

[48] My argumentative strategy in this section thus differs from Taylor, who argues against Gaus's appropriateness conditions by pointing to the availability of a superior alternative: the so-called "susceptibility to reason" account of the appropriateness conditions for the reactive attitudes. In a sense, this makes my strategy weaker and less systematic because it relies on intuitions about cases that may not be shared by everyone. But in another sense, my strategy is stronger and more general because it is not tied to the success of any particular alternative theory. For instance, one might plausibly reject Taylor's susceptibility to reason alternative on the grounds that it violates the epistemic condition for moral responsibility (See Vallier, K., *Must Politics Be War?,* manuscript v. 4.0, p. 142). But while this would undermine Taylor's alternative-principle-based objection to Gaus, it would still leave untouched my argument that Gaus's account has the highly counterintuitive implication that we must suspend feelings of indignation and resentment towards genocidal killers, tellers of racist jokes and so on if we discover that such individuals see nothing wrong with their conduct and do not see what all the fuss is about. (My argument also differs from Taylor's argument insofar as he does not consider whether the reactive attitudes can be renounced wholesale, or whether the reactive attitudes have the right normative status to ground public justification — see my Sections 3 and 4.)

the Sufficient Reason Condition, it is enough to show that it implies that Goebbels is off the reactive-attitudes hook; we need not in addition offer, in place of this condition, some alternative theory of liability to the reactive attitudes that explains where among a range of important disanalogies the line between Goebbels+ and Winston Smith should be drawn.[49]

3 Renouncing the reactive attitudes

In the previous section, I argued, contra Gaus, that the reactive attitudes do not presuppose having a sufficient reason. Even if this objection fails, however, there are problems with the second premise of Gaus's argument, which is that the reactive attitudes cannot be rationally renounced. So, to fortify my argument, and to give a fuller picture of the problems with Gaus's reactive attitudes defence of public justification, I now want to challenge Gaus's claim that the reactive attitudes cannot be rationally renounced.

Recall the passage quoted in Section 1.2. This is virtually all Gaus says in defence of the rational non-renounceability of the reactive attitudes. And the heart of argument seems to be that we cannot do without the reactive attitudes because a life without the reactive attitudes would be devoid of meaning and value. Abandoning the reactive attitudes, Gaus suggests, would undermine "our understanding of interpersonal relations, including love and friendship, and so of what sort of life is worth living"; thus "to renounce the practice would be to renounce most of what we care for and value".[50]

However, whether renouncing the reactive attitudes would be disastrous for what we most care about and value is quite contestable. A number of writers on moral responsibility, and in particular Derk Pereboom, have recently argued that many of the reactive attitudes have (i) non-reactive *aspects* that could survive the rejection of the reactive attitudes or (ii) non-reactive *alternatives*

[49] Perhaps Gaus would reply here that each of these five alternative appropriateness conditions also comes with severe costs, in terms of counterintuitiveness, and that this is true of alternative conditions in general, and thus that the Sufficient Reason Condition is the best we've got. But, to make this case, Gaus would need to conduct a detailed cost-benefit analysis of the candidate appropriateness conditions, one that demonstrates the superiority of his Sufficient Reason Condition over all other rival theories of moral responsibility — a tall order, to say the least, and not something that Gaus anywhere does. (I thank Anthony Taylor for discussion of this point.)

[50] Gaus, G., *The Order of Public Reason*, p. 192.

that are just as good as the reactive attitudes themselves.[51] For instance, and to illustrate (ii): the reactive attitude of *resentment* could arguably be replaced by the non-reactive attitude of *sadness* or *disappointment* without thereby threatening anything of value. By drawing on this Pereboomian insight, I suggest that social life would still be possible, and perhaps even improved, in the absence of the reactive attitudes. If on reflection we find this argument to be coherent, if it does not seem to make psychologically impossible demands on us, and if it preserves what we most value and care for in life, then it would seem that Gaus is incorrect to say that the reactive cannot rationally be renounced.

3.1 Pereboom's alternative account

The reactive attitudes, recall, are "reactions to the quality of others' wills towards us" — to their good will or ill will.[52] There are debates about which attitudes are reactive in this sense, but it is possible to categorize the reactive attitudes into two classes and to work from certain paradigmatic cases. On the one hand, there are *negative* reactive attitudes such as resentment, blame and indignation; on the other, there are *positive* reactive attitudes such as gratitude and reciprocal love. Drawing on the Pereboomian strategy of identifying non-reactive aspects of these reactive attitudes, as well as non-reactive alternative attitudes, I shall now argue that against Gaus's claim that renouncing the positive and negative attitudes would so impoverish social life as to make it no longer worth living.

Let us begin with the negative reactive attitudes. As Pereboom argues, resentment and indignation are "suboptimal as modes of communication in relationships relative to alternative attitudes available to us".[53] When someone wrongs me or wrongs a third party, there are other available emotions beyond resentment and indignation: I might feel *hurt* or *shock* or *disappointment* or *sorrow*. These alternative non-reactive attitudes preserve what (if anything) is valuable about resentment and indignation. Disappointment, for instance, can

[51] See, e.g., Sommers, T., "The Objective Attitude", *The Philosophical Quarterly* 57 (2007), pp. 321-41; Pereboom, D., "Free Will, Love, and Anger", *Ideas y Valores* 141 (2009), pp. 5-25; Pereboom, D., *Free Will, Agency, and Meaning in Life* (Oxford: Oxford University Press, 2014).
[52] Strawson, P. F., "Freedom and Resentment", p. 83.
[53] Pereboom, D., *Free Will, Agency, and Meaning in Life*, p. 179.

quite effectively play the communicative role typically played by resentment and indignation. Indeed, Pereboom even argues that these non-reactive substitutes are often preferable because resentment and indignation can "damage or destroy relationships", can "give rise to destructive resistance instead of reconciliation", and can, at the social level, lead to excessively punitive and retributive policies.[54]

One might object that even if the reactive attitudes are often detrimental, there is also a significant positive side to emotions such as anger, resentment and indignation. First, these emotions seem to play an important role in motivating people to combat oppression and injustice. Without the reactive attitudes, would we still fight the good fight with all our might? Second, the reactive attitudes seem essential to our practices of punishment, which in turn seem essential to social cooperation. How would we account for punishment without the reactive attitudes of judgement, resentment and indignation? And, in the absence of the threat of punishment, would people still cooperate and follow rules? So even if Pereboom's approach works for one-on-one interpersonal relations, would it really be able to sustain large-scale social cooperation within highly diverse societies?[55]

These worries, both of which point towards benefits of the practice of the reactive attitudes that would supposedly be lost under Pereboom's alternative practice, raise complex empirical and normative questions that I cannot address fully here. But let me make three brief points. First, regarding the fight against injustice: while anger and indignation do motivate, there are also plenty of powerful non-reactive sources of motivation for resistance and social change, such as compassion, sympathy, sorrow and a sense of justice. Mahatma Gandhi and Martin Luther King, for instance, were both able to successfully resist and protest injustice by appealing to moral principle and not primarily to resentment, indignation and other reactive attitudes that are based on the perceived ill will of others.[56]

Second, regarding cooperation: neither (a) the premise that cooperation requires punishment nor (b) the premise that punishment requires the reactive

[54] Ibid., p. 180.
[55] I thank two reviewers for putting to me the worries about fighting injustice and about sustaining cooperation.
[56] Pereboom, D., *Free Will, Agency, and Meaning in Life*, p. 149.

attitudes is entirely convincing. To take only (b): it is true that punishment is often driven by reactive attitudes such as resentment, blame and indignation; generally, when these emotions are absent, so too is the impulse to punish. But there are various alternative ways of thinking about and justifying punishment — alternatives that do not involve the reactive attitudes. For instance, on Pereboom's "quarantine" account of punishment, the punishment of criminals and cheaters is modeled on the containment of carriers of dangerous diseases: preventative measures are taken to protect society and to preserve commitment to its rules, but these are not accompanied by feelings of blame or resentment towards those who pose the harm.[57]

The third and more general point is about the dialectical bar and about the modesty of my aims in this section. I need not show, and am not seeking to show, that Pereboom's alternative account is perfect — nor even, in fact, that it is as good than our current practice of the reactive attitudes of resentment, indignation, blame and so on. Perhaps, for instance, the quarantine model is on balance worse than the reactive attitudes at driving punishment. Rather, I need merely to show that Pereboom's account is *liveable* or *good enough* — that it leaves intact a solid amount of what we care about. Satisfying this relatively low dialectical bar would be sufficient to undermine Gaus's Strawsonian claim in Section 1.3 according to which renouncing the reactive attitudes would be catastrophic for social life. For Gaus, recall, the practice of the reactive attitudes is "so deep a part of the way we see the world [that] to renounce the practice would be to renounce most of what we care for and value".[58] This language might be plausible if the only alternative to the reactive attitudes were some kind of bleak state of nature devoid of emotion and cooperation. But the availability of a liveable alternative practice such as Pereboom's, even if it is a downgrade from our current practice, shows Gaus's language here to be overstated. Even if there would be *some* costs to giving up the reactive attitudes, that is, Gaus cannot import the enormous Strawsonian costs to which he appeals.

Perhaps some will think that negative reactive attitudes are the easy cases. Abandoning resentment is no great loss. The hard cases are positive reactive attitudes, such as love and gratitude. How could we live without these?

[57] Ibid., p. 169ff.
[58] Gaus, G., *The Order of Public Reason*, p. 192.

However, as Pereboom points out, while love and gratitude are reactive attitudes in some respects, they are also non-reactive attitudes in other respects. So to reject the reactive attitudes one need not reject the attitudes of love and gratitude wholesale; one need only reject the reactive *aspects* of these attitudes. The question is: How much is left? Are enough aspects of love and gratitude in the clear for some recognizable and valuable form of that attitude to remain in place?

Consider gratitude. It might well be true that gratitude has certain reactive aspects — that part of what it means to be grateful is to be responsive to another person's good will. But gratitude also has important non-reactive aspects. For instance, one can be thankful to a young child for some kind gift, despite the child being too young to be able to express good will. As Pereboom says, "gratitude involves, first of all, being thankful toward someone who has acted beneficially" — and insofar as this core sense of gratitude makes no essential reference to the quality of others' wills, it could survive the rejection of the reactive attitudes.[59]

The same goes for love. Again, it might seem undeniable that love is completely or at least primarily a reactive emotion — that it is a response to the quality of another's will. But in fact love has many non-reactive aspects. A mother does not in general love her children because of their good will towards her — this being most obviously true in the case of very young children who are incapable of good will or ill will. Even when mature adults love each other, this love is nourished by intelligence, shared interests, appearance, mannerisms, and other factors that make no essential reference to perceived good will. As Pereboom argues, it is possible to see love in largely non-reactive terms: "love of another involves, most fundamentally, wishing well for the other, taking on aims and projects of the other as one's own, and a desire to be together with the other", none of which is threatened by the rejection of the reactive attitudes.[60]

All of this casts considerable doubt on Gaus's claim that "to renounce the practice [of the reactive attitudes] would be to renounce most of what we care

[59] Pereboom, D., *Free Will, Agency, and Meaning in Life*, p. 190.
[60] Ibid.

for and value".[61] Gaus worries, following Strawson, that to reject the reactive attitudes would be to adopt the dreaded "objective attitude" towards other persons whereby we see them "as an object of social policy; as a subject for what, in a wide range of sense, might be called treatment; as something certainly to be taken account, perhaps precautionary account, of; to be managed or handled or cured or trained".[62] But, while Pereboom's alternative perspective does involve a greater degree of objectivity in interpersonal relations, it does not justify assuming an objective attitude in this sense. Pereboom's alternative account thus shows that it is possible to renounce the reactive attitudes whilst leaving intact much of what we care about in life. In fact, his account may even serve improve our interpersonal relationships by making us calmer, less angry, judgmental and confrontational. And, with determination, education and the right social and cultural environment, there is no reason to think that a gradual adoption of this alternative moral practice is impossible: "the past several centuries have witnessed very significant changes in attitudes toward criminals, the insane, and children, and thus it cannot plausibly be argued that significant emotional change over time is not possible for us".[63]

Of course, for Pereboom, the search for non-reactive analogues and aspects of the reactive attitudes is motivated by the question of how personal relationships and meaning in life could be rationally sustained in the face of hard determinism. But one does not have to subscribe to Pereboom's views on free will to accept that his alternative picture of social life is an interesting and plausible contender and thus to accept that the reactive attitudes can be rationally renounced. So until Gaus engages with this possibility and explains where Pereboom goes wrong, it is not clear that he has grounds for saying that to renounce the reactive attitudes would be to renounce most of what we care for and value.

4 Is the reactive attitudes argument justificatory?

In the previous two sections, I have raised doubts about the claim that the reactive attitudes presuppose public justification and about the claim that the

[61] Gaus, G., *The Order of Public Reason*, p. 192.
[62] Strawson, P. F., "Freedom and Resentment", p. 79.
[63] Pereboom, D., *Free Will, Agency, and Meaning in Life*, p. 185.

reactive attitudes cannot be rationally renounced. But suppose I am wrong about this. Suppose, that is, that reactive attitudes presuppose public justification and that the reactive attitudes cannot be rationally renounced. Still, Gaus's argument would not be in the clear. For all that this would show is that it is impossible to rationally renounce public justification.

The problem here is this: Does this actually constitute a *justification* of public justification? Where does the normativity come in? It shows we are stuck with public justification, to be sure: but is this for good or for bad? Has Gaus provided a justification for why we *should* adhere to public justification, or merely an explanation of why we *do* adhere to public justification?

To feel the force of these questions, it can be helpful to compare Gaus's reactive attitudes argument for public justification with other arguments for public justification, such as the traditional argument from respect. Larmore, for instance, as we saw earlier, argues that respect requires public justification: "if we try to bring about conformity to a rule of conduct solely by the threat of force, we shall be treating persons merely as a means, as objects of coercion, and not also as ends, engaging directly their distinctive capacity as persons".[64] Now some deny that respect really does require public justification; they hold that one can respectfully coerce another person even while offering her reasons unacceptable to her, so long as those reasons are *good* reasons. But what cannot be denied is that respect is genuinely normative: *if* respect *did* require public justification, *then* that would give us strong normative reasons to adhere to public justification. By contrast, it looks as though one can say in response to Gaus's argument: "I see that public justification is entailed by deeply embedded, impossible-to-renounce practices of resentment and indignation; but, still, one cannot derive an 'ought' from an 'is' and so I don't see how empirical practices, even deeply embedded ones such as these, can generate normative reasons to adhere to public justification."

One might worry that this objection is unfair to Gaus. After all, Gaus has complex views about the relation between the normative and the empirical, and he rejects any sharp distinction between the two.[65] For Gaus, "the

[64] Larmore, C., "The Moral Basis of Political Liberalism", p. 607.
[65] See, e.g., Gaus, G., "On Theorizing About Public Reason", p. 67.

beginning point of understanding 'true morality' is 'actual morality'"; a moral theory that is insensitive to "the facts of our social life" and to the role morality plays in "mak[ing] a cooperative social life possible among human beings" would, he says, be "academic in the most pejorative sense".[66] Gaus rejects the picture of political philosophy in which, as he colourfully puts it, "tablets inscribed with the truth…are sent down from on high to the rest to guide them".[67] So perhaps Gaus would say that, once it has been shown that public justification cannot be renounced, there is nothing more to show. Perhaps, in other words, the reactive attitudes argument is as good a justification as one can and should expect — and so to seek anything more, to ask whether the reactive attitudes are "truly" normative, is to beg the question against Gaus's meta-philosophy by presupposing some deep empirical/normative distinction.

But questioning whether the reactive attitudes argument is genuinely justificatory begs no questions against Gaus. For while he views some empirical conditions as necessary for morality, he does not take these to be *sufficient*: "'true morality' must also be a 'positive' (existing) morality, *although not all positive moralities qualify as true moralities*".[68] So at the same time as highlighting "the error of dismissing positive morality", Gaus also warns against "the opposite error": namely, that of adopting the "quasi-anthropological" view that "moral philosophy is essentially a sort of systematization of moral psychology and positive morality".[69] Gaus thus accepts the idea that, in moral theorizing, we can "stand back from our social institutions and take the perspective of what, we might say, 'morality itself tells us'".[70] Indeed, that Gaus would consider the question posed in this section to be legitimate and meaningful is clear from the fact that he himself expresses, and endeavours to answer, a concern along these lines. I find his answer, however, to be quite unclear, and so it is worth quoting Gaus in full here:

> Still, the worry gnaws: have we simply landed in a confused practice that we cannot reason ourselves out of? That may be a recipe for despair rather than contented resignation, much less justification. We not only

[66] Gaus, G., *The Order of Public Reason*, p. 102-3, 176.
[67] Gaus, G., "The Good, the Bad, and the Ugly: Three Agent-Type Challenges to *The Order of Public Reason*", p. 577.
[68] Gaus, G., *The Order of Public Reason*, p. 182 (emphasis added).
[69] Ibid., p. 173, 176, 177.
[70] Ibid., p. 180.

> wish to know where we have landed but to have some reassurance that it is a destination worth arriving at, and not the result of being marooned with no hope of rescue. Our current moral practice is made intelligible and sensible once we understand that human society depends on a social morality based on Rule-following Punishers. For such a system to have arisen, we must care about the moral actions of others, care about making demands on them, and hold that we have standing to make these demands. When Strawson says that the "existence of the general framework of attitudes is something we are given with the fact of human society," he presents us with a deeper truth than even he realizes: human society would not even be possible without this framework.[71]

Gaus's statement of the gnawing worry is admirably clear. But it is difficult to interpret his response or to see its relevance foregoing worry. The first oddity is that as Gaus begins his answer ("our current moral practice…"), he falls back into non-normative talk: he says that the reactive attitudes argument makes our current moral practice "intelligible and sensible". But even if the reactive attitudes argument makes public justification *intelligible*, in the sense of explaining how the practice of public justification arose, that would fall short of showing us that we *should* practice public justification; the whole point of the gnawing worry is that explaining a practice is not the same as justifying that practice. We want to know, especially in light of the fairly libertarian implications of public justification, whether our path-dependent evolutionary history has landed us in a just or unjust practice.[72] (Perhaps this is why Gaus adds the term "sensible" — a term that can be read both normatively and non-normatively.)

In any case, let us suppose that Gaus means "justified" where he says "intelligible and sensible". Still, the main problem with the response is that it simply restates the reactive attitudes argument. The heart of Gaus's reply here seems to be that public justification is justified by the fact that "human society would not even be possible without" public justification (because human society would not be possible without the reactive attitudes, and because the reactive attitudes presuppose public justification). As Gaus elsewhere puts it:

---

[71] Ibid., p. 192-3.

[72] See also Vallier, K., "Gaus, Hayek, and the Place of Civil Religion in a Free Society", *The Review of Austrian Economics* 30 (2017), pp. 339-45.

"absent endorsability the good simply cannot be achieved".[73] Now, as I have argued in Sections 2 and 3, and as suggested by the fact that most actual societies today have been able to arise and exist despite not being "orders of public reason", I am doubtful of the truth of the possibility claim here: I am doubtful both that human society requires the reactive attitudes (Section 3) and that the reactive attitudes presuppose public justification (Section 2). But, even granting that human society would not be possible without public justification, we are still left with the gnawing worry: for we still want to know whether the non-renounceability of public justification is a cause for joy or lamentation. In other words, in response to the gnawing worry ("the reactive attitudes argument shows that we are stuck with public justification: but the argument is not justificatory, because it does not show whether this is for good or for bad"), Gaus seems to simply repeat his prior claim: "we are stuck with public justification".

Consider a comparison. Some people believe that aggression is inherent to human beings — that it is so deep a part of our constitution that it is a fixed and unavoidable feature of human life. Others deny this, of course. But suppose ex hypothesi that, as an empirical claim, it is true: we are stuck with aggression. Gaus's position seems to be that in such circumstances it would make no sense to ask, or there would be no point in asking, whether aggression is a *desirable* trait, whether we *ought* to act aggressively and violently in response to provocation. It seems that, for Gaus, if no one can avoid aggression, then questions about its justifiability simply do not arise. But this does not seem right. Surely the thing to say here — at least if we wish to avoid, in Gaus's own words, the erroneous "quasi-anthropological" view whereby "moral philosophy is essentially a sort of systematization of moral psychology and positive morality" — is that, while we tend to act aggressively, we ought not do so. If we can, we should *resist* aggression; and if we cannot, we should *regret* our aggressive and destructive responses to provocation. That it is difficult or even impossible to act in a non-aggressive and friendly way does not undermine friendliness as a normative ideal any more than the difficulty or impossibility of creating an error-free criminal justice system undermines the normative ideal that there should be no convictions of the innocent and no acquittals of the guilty. In other words, the normative question of justification is still on the

[73] Gaus, G., "On Theorizing About Public Reason", p. 73.

table, even after it has been assumed that aggression is a fundamental feature of human life. The same could be said of a range of other distinctively human yet undesirable traits, such as jealousy, vanity and domination.

Of course, these are familiar, even banal, observations. But they bear mentioning here because they are, I think, precisely the kind of observations that illustrate the inadequacy of Gaus's response to the worry about the justification of the reactive attitudes. They demonstrate that even if the soundness of the reactive attitudes argument is granted, we are still left with what Gaus acknowledges is a "gnaw[ing]" question yet which he does not properly answer, namely: But is any of this *justificatory*?

Does Social Trust Justify the Public Justification Principle?

According to public reason liberalism, the state must abide by a principle of public justification. This principle holds that the laws and institutions of society must be in some sense justifiable to, or acceptable to, all reasonable citizens. Many contemporary liberal political philosophers are attracted to this approach insofar as it "focuses like a laser on the core aim of the liberal tradition: justifying coercion to all".[1] Different public reason liberals theorize the public justification principle in different ways and take it to have different political implications. For Rawls and for Quong, the public justification principle primarily implies that important political decisions must be justified in terms of shared public reasons such as fairness and equality, and not by appeal to controversial perfectionist doctrines.[2] For Gaus, the implications of the public justification principle are more radical and far-reaching in that it calls into question the legitimacy of much egalitarian liberalism and tilts towards the minimal state.[3] But, despite these differences, the idea of "justifiability to all" remains distinctive and central to all varieties of public reason liberalism.

However, the public justification principle is not self-evident. Many remain unconvinced that public justification is a genuine constraint on government action. At most, say the critics, acceptability-to-all is a *pro tanto* value: it would be nice, of course, if all citizens could accept all laws, but this kind of acceptability is just another item on the long and dense list of political desiderata, and not a particularly weighty item at that.[4] So public reason liberals need to provide a justification (or grounding, or moral foundation) for public justification. They need, that is, to answer the question: Why reason publicly? Why care about justifiability to reasonable citizens? Why not simply implement the political rules and institutions that are *correct* — or at least that

---

[1] Vallier, K., *Liberal Politics and Public Faith: Beyond Separation* (New York: Routledge, 2014), p. 261.

[2] Rawls, J., *Political Liberalism* (New York: Columbia University Press, 1996); Quong, J., *Liberalism Without Perfection* (Oxford: Oxford University Press, 2011).

[3] Gaus, G., "Coercion, Ownership, and the Redistributive State: Justificatory Liberalism's Classical Tilt", *Social Philosophy & Policy* 27 (2010), pp. 233-75; Gaus, G., *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World* (New York: Cambridge University Press, 2011).

[4] See, e.g., Enoch, D., "Against Public Reason", in D. Sobel, P. Vallentyne and S. Wall (eds), *Oxford Studies in Political Philosophy*, vol. 1 (Oxford: Oxford University Press, 2015), pp. 112-44, esp. at pp. 138-40.

are most likely to be correct — regardless of whether those rules and institutions could be reasonably rejected by some citizens?

Yet, surprisingly, many prominent public reason liberals have said little about this foundational question, and what they have said is often brief and lacking in detail. For instance, while the public justification principle (or the "liberal principle of legitimacy", as he calls it[5]) is the centrepiece of Rawls's *Political Liberalism*, he says little in favour of the principle. There are hints in Rawls that public justification is grounded in respect for persons. He says that only when citizens deliberate in terms of public reasons do "their judgements converge sufficiently so that political cooperation on the basis of mutual respect can be maintained".[6] But while some failures to deliberate and to make decisions in terms of public reasons (such as theocracy or dictatorship) certainly do undermine mutual respect, it is unclear why *all* alternative models of democratic discourse and decision-making short of public reason (e.g. deliberative democracy) are incompatible with mutual respect. Rawls never considers in detail a range of alternatives, explaining why *only* the liberal principle of legitimacy would treat all citizens with respect.

It is natural to speculate about the cause of this surprising neglect of foundations. Perhaps public reason liberals have said so little about the foundations of public justification because they have followed Rawls's encouragement that liberalism "stays on the surface, philosophically speaking".[7] In basing liberalism on values that all can accept, rather than on some comprehensive doctrine such as Millian individuality, "we try," says Rawls, "to bypass religion and philosophy's profoundest controversies so as to have some hope of uncovering a basis of a stable overlapping consensus".[8] In any case, one exciting current development in the public reason literature is a willingness to dive deeper into the moral and philosophical foundations of public justification. In recent years, that is, a number of detailed and sophisticated arguments for public justification have been proposed, including arguments from stability,[9] from civic friendship,[10] and from the reactive attitudes.[11]

---

[5] Rawls, J., *Political Liberalism*, p. 137.
[6] Ibid., p. 156.
[7] Rawls, J., "Justice as Fairness: Political not Metaphysical", *Philosophy & Public Affairs* 14 (1985), p. 230.
[8] Rawls, J., *Political Liberalism*, p. 152.
[9] Weithman, P., *Why Political Liberalism?* (Oxford: Oxford University Press, 2010).

One of the most interesting and empirically informed of these more recent defences of public justification is Kevin Vallier's argument from social trust.[12] This argument has two main premises. The first premise is that the creation and maintenance of a system of social trust requires the practice of public justification, and the second is that a system of social trust has very great teleological and deontological value. Together, these premises generate a strong case for the public justification principle.

In what follows, I argue that Vallier's argument from social trust to public justification is unsuccessful. In broad outline, my main objection is that Vallier's depiction of the moral and political landscape is too stark: on his view, either we engage in public justification or (as he often puts it) "politics is war".[13] Thus: "If we impose our sectarian conception of justice on others [instead of justifying it *to* them], there is an important sense in which we are at war with them".[14] But this language is overstated. As with the argument from respect: even if *some* alternatives to public justification (again, such as theocracy and dictatorship) are aptly described as forms of war, as sectarian impositions, and as vitiative of social trust, this is not true of all alternative models of political discourse and decision-making. In this paper, then, I first spell out in more detail, in Section 1, Vallier's social trust argument. I then critique, in Section 2, Vallier's claim that the creation and maintenance of a system of social trust requires the practice of public justification by pointing to an alternative conception of democratic discourse and decision-making that also seems able to create and maintain social trust despite lacking a public justification principle. I then consider and reject two responses that Vallier might make: that alternatives to public justification would produce *less* social trust (Section 3), and that such alternatives would not produce social trust *in the right way* (Section 4).

1 Vallier's social trust argument for public justification

[10] Lister, A., *Public Reason and Political Community* (London: Bloomsbury, 2013).
[11] Gaus, G., *The Order of Public Reason*.
[12] Vallier, K., *Must Politics Be War? Restoring Our Trust in the Open Society* (Oxford: Oxford University Press, 2019). As Vallier notes (p. 7), there are brief earlier suggestions that public justification might be grounded on social trust in Gaus (*The Order of Public Reason*, p. 315) and in Rawls (*Political Liberalism*, p. 86, 163, 168), though neither author really develops this idea.
[13] See ibid., p. 18 *et passim.*
[14] Ibid., p. 12.

In this section I outline Vallier's argument from social trust to public justification. I first specify, in Section 1.1, the specific form of the public justification principle that Vallier affirms and that he seeks to justify. I then spell out, in Sections 1.2 and 1.3, the two premises of Vallier's argument for that principle.

1.1 Vallier's version of the public justification principle

Vallier affirms the following public justification principle:

> A moral rule is publicly justified only if each member of the public has sufficient intelligible reason to comply with and internalize the rule.[15]

Three features of this account call for further elaboration and comment. First, a distinctive feature of Vallier's account is his understanding of the *scope* of public justification. Traditionally, public reason liberals argued that it is only "constitutional essentials and matters of basic justice" that must be publicly justified.[16] Since then, other public reason liberals have defended the broader view according to which it is all political decisions that must be publicly justified.[17] Vallier's public justification principle, however, is even broader in scope. On his view, it is not just legal-political rules but also the "moral rules" that are invoked in ordinary interpersonal life that are the proper object of public justification "because they are the central unit is sustaining and justifying social trust".[18]

Second, in order to understand and evaluate this public justification principle, we need to know what it means to say that each member of the public "has" sufficient intelligible reason to accept rules. As Gaus has argued, the notion of "having a reason" is ambiguous between at least two interpretations — internalist and externalist — and Vallier follows the Gausian internalist understanding of what it means to have a reason.[19] On the externalist view of

[15] Ibid., p. 116.
[16] Rawls, J., *Political Liberalism*, p. 230 *et passim*.
[17] Quong, J., *Liberalism Without Perfection*, pp. 273-81.
[18] Vallier, K., *Must Politics Be War?*, p. 81.
[19] Ibid., p. 92.

having a reason, an agent has a reason whenever there is a reason that applies to her. Gaus rejects the externalist picture because it "misconstrues the relation between having a reason and being a rational agent".[20] The externalist view implies, for instance, that Aristotle had a reason to accept particle physics. But, says Gaus, Aristotle had no such reason; indeed, only by ignoring his evidence and by being irrational could he have come to endorse such a reason. When we move from theoretical reasoning to practical reasoning, externalism might seem more plausible. For surely everyone has a reason to avoid genocide? But consider the case of a benighted tribal society that lacks the concept of an ethnic group. Of this case, Gaus says: "to attribute a reason *to* these peoples, to say that they *possessed* a reason not to eliminate other groups, while acknowledging that any justificatory force of this reason is inaccessible to them as reasoning beings is, I think, not only a misuse of language, but undermines the point of discourse about reasons and rationality".[21] It is one thing to say that *there is* a reason to avoid genocide that applies to these tribal societies. But to say that this reason — this "incomprehensible justifying consideration" — is a reason they *have* is, thinks Gaus, "close to bizarre".[22]

Instead, Gaus adopts an internalist account of having a reason. This view, of course, need not say that the reasons we have are those that we would immediately affirm at any particular moment. Some idealization of the reasoner is permitted. But, to avoid the inaccessibility problems of the external account, this idealization must be *moderate*. So the idea is that there must be "some not-too-difficult bridge to cross connecting what reasons a person can be said to possess and what the exercise of her rational faculties can lead her to".[23] Thus:

> 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).[24]

[20] Gaus, G., *The Order of Public Reason*, p. 233.
[21] Ibid., p. 235.
[22] Ibid.
[23] Ibid., p. 35.
[24] Ibid., p. 250.

Vallier appeals to his account of the foundations of public justification to justify his rejection of externalist and highly idealized accounts in favour of the internalist Gausian account of what it means to have a reason. For Vallier, as for Gaus, the reasons that a person is said to have must genuinely depend to a significant extent on the values and beliefs of the person in question — they must, that is, be genuinely *hers*. "If she cannot see the reasons as hers according to her moral standards," Vallier explains, "then her reasons cannot drive socially trustworthy behaviour, since trustworthy behaviour requires an appreciation of moral reasons".[25] Only reasons that are "psychologically accessible and morally motivating on reflection" can sustain a system of social trust.[26]

The third feature of Vallier's public justification principle worth clarifying is his reference to "intelligible" reasons. Many public reason liberals, including Rawls and Quong, require that public justification must proceed in terms of reasons that are shared or shareable. The idea here is that, given the fact of evaluative diversity, we exclude controversial reasons (e.g. reasons based on Catholicism) from politics and instead we deliberate and justify political decisions in terms of reasons that all citizens share or at least could in some sense share — reasons such as freedom, equality, fairness, and so on. By contrast to these "consensus" models of public justification, Vallier adopts a "convergence" model in which the notion of shareability is replaced by the more minimal notion of intelligibility. On this view, the reasons that figure in a public justification need not be shared: a political decision can be justified to Adam in terms of his Catholicism, to Betty in terms of her Marxism, and so on. The only constraint is that these reasons must be in some sense intelligible to others *as reasons* for those who rely on them, where this is to be contrasted with expressions of emotion, appeals to power, or irrelevant utterances.

We now have a sense of the public justification principle that Vallier affirms and that he seeks to justify through the social trust argument. Of course, there are various ways in which one might challenge this principle. An obvious question to ask is that of whether Vallier's public justification principle would guarantee sufficiently liberal outcomes. In this connection, it is worth noting

[25] Vallier, K., *Must Politics Be War?*, p. 79.
[26] Ibid., p. 89.

that Vallier follows Gaus in adopting a *procedural* (rather than substantive) account of what it means to have a reason. But, as Quong has pointed out,[27] and as Vallier[28] and Gaus[29] both concede, this means that there is no principled moral constraint on the reasons that an agent can end up having. What an agent "has a reason" to do or believe, on Vallier's and Gaus's accounts, is not a matter of the objective or substantive reasons that apply to her: for this is the externalist account of having a reason, which severs the connection between having a reason and being rational and which would risk the collapse of public justification into ordinary rational justification. Rather, what an agent has a reason to do or believe is a matter of "what the exercise of her rational faculties can lead her to" after a respectable amount of time.[30] So Vallier's and Gaus's accounts of having a reason cannot foreclose the possibility of an individual such as Quong's Carl, who has sufficient reason to kill infidels — in the sense that, even after a respectable amount of time spent thinking about the matter, he believes he is required to kill infidels.[31] Carl recognizes, we can assume, that there is *some* reason not to harm others, but he attaches tremendous weight to the elimination of infidels, and so his reason to kill infidels is undefeated.

I am sympathetic to the worry that Vallier's public justification principle does not guarantee liberal outcomes because of the possibility of individuals with illiberal beliefs exercising a veto. But here I object to Vallier's theory from the other direction, as it were. In this paper, I am interested less in his move from public justification to liberal outcomes than in his move from social trust to public justification. Another way of putting this is that I intend to focus not on Vallier's answer to the question of what would pass the public justification test but rather on his answer to the more foundational question of why we should even care about that test in the first place.

1.2 A system of social trust requires public justification

Vallier argues that adherence to the practice of public justification would create and maintain a system of social trust between citizens — where a system of

[27] Quong, J., "What Is the Point of Public Reason?", *Philosophical Studies* 170 (2014), pp. 546-7.
[28] Vallier, K., *Must Politics Be War?*, pp. 94-7.
[29] Gaus, G., "The Good, the Bad, and the Ugly: Three Agent-Type Challenges to *The Order of Public Reason*", *Philosophical Studies* 170 (2014), pp. 570-1.
[30] Gaus, G., *The Order of Public Reason*, p. 35.
[31] See Quong, J., "What Is the Point of Public Reason?"

social trust is understood as "the general expectation that all members of society will comply with moral rules" and is measured by international survey questions such as "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?"[32] When a moral or legal rule is publicly justified, each individual sees herself as having reason to comply with that rule. This means that "persons see themselves as having reason to be trustworthy by complying with the rule" and so "social trust forms through the perception of this properly incentivized trustworthiness".[33] As Vallier also explains, public justification facilitates the emergence of social trust within highly diverse societies — societies that share no bonds of blood and that are not united by a common religion or worldview, and thus societies within which high social trust might at first seem impossible or improbable — because members of a society regulated by public reason can obey, and can expect others to obey, the laws and institutions of society without acting against their deepest ideals and convictions.[34]

A crucial claim that Vallier makes, however, is that public justification is not just *sufficient* for the creation and maintenance of a system of social trust but also *necessary* for such a system. Alternatives to public justification, in other words, would be incapable of sustaining such a system. As Vallier puts it: "maintaining a system of trust *requires* that our moral rules be publicly justified".[35] "Moderately idealized intelligible reasons appropriately morally incentivize trustworthiness *and are the only reasons that do so*."[36] And: "publicly justified moral rules have the *unique power to sustain a system of trust.*"[37] This necessity claim — that social trust requires public justification — is crucial because, without it, Vallier would not have answered the sceptics who question why public justification matters so much. After all, if other models of deliberative-democratic discourse and decision-making short of public justification were *also* able to secure social trust — if, that is, public justification were sufficient but not necessary for social trust — then considerations of social trust would not be able to answer the question of why

[32] Vallier, K., *Must Politics Be War?*, p. 36, 50.
[33] Ibid., p. 79.
[34] Ibid., p. 118.
[35] Ibid., p. 127 (emphasis added).
[36] Ibid., p. 118 (emphasis added).
[37] Ibid. (emphasis in original)

we should select public justification over familiar and more traditional liberal-democratic theories.

Yet Vallier does not argue for the necessity claim in the way one might expect. He does not, that is, argue for the necessity claim by way of a detailed comparison between the public justification model and a variety of alternative models, showing how none of these alternatives would facilitate social trust (or perhaps: none would facilitate social trust to a sufficient degree, or in the appropriate way).

Admittedly, there is *some* such comparative analysis. For instance, Vallier imagines a "small band of Marxists" who "ostracize local business owners for not handing their capital over to their employees".[38] Even if this Marxist reasoning is objectively correct, it cannot be justified *to* the business owners, and so, says Vallier, it is "normatively identical to insisting that American business owners fork over their capital because the Soviet Union has passed a law requiring capital redistribution".[39] The Marxist claim "has no rational uptake" for the business owners and so it "tend[s] to undermine, rather than support," social trust.[40] Vallier also considers, as another example of failure to engage in public justification, the enactment of a law that simply lacks moral justification altogether. "Unjustified laws" of this kind, he says, "cannot drive genuine trustworthiness" because even morally motivated individuals subject to such laws will defect and free-ride, thus undermining trust, whenever they believe they can get away with doing so.[41] Third, Vallier compares public reason liberalism with a *modus vivendi* regime — a regime whose "stability and persistence is based on mere power relations" and in which citizens are strategically, rather than morally, committed to political institutions.[42] Vallier explains that "a *modus vivendi* order cannot stably establish trust because it is grounded in non-moral reasons to act".[43] A fourth example of non-public reason comes in an interpersonal analogy. Vallier imagines a dinner party in which the host, John, orders Reba, the vegan guest, to eat meat. The fact that John "issue[s] meat-eating orders" despite "fully realizing that the order is

[38] Ibid., p. 138.
[39] Ibid.
[40] Ibid.
[41] Ibid.
[42] Ibid., p. 107.
[43] Ibid., p. 108.

incompatible with Reba's deep commitments", says Vallier, means that John "flagrantly disregards Reba's perspective" and undermines the system of trust between himself and his guest.[44]

But Vallier's argument for the claim that social trust requires public justification is not primarily by way of systematic analysis of the alternatives to public justification. Except for these scattered examples, Vallier largely relies on appeal to the metaphor of war and peace. He says that if exercises of political power are not acceptable from the point of view of moderately idealized citizens, then "these demands look like small-scale acts of war".[45] In such a case, the political order "becomes a mere contest for control".[46] And again: "public justification can sustain relations of social trust between real, diverse persons and so show that…liberal politics can be more than institutionalized aggression".[47] The flip side of Vallier's claim that a society that is not regulated by public justification is a society at war is his claim that a society regulated by public justification is a society at peace. Thus, for Vallier, "publicly justified moral rules [are] the building blocks of a morally peaceful society".[48] "The institutions that can sustain moral peace," he says, "are those that can be jointly endorsed".[49]

1.3 A system of social trust has very great value

Suppose we accept that a system of social trust requires public justification: anyone committed to social trust is thereby committed to public justification. Still, for this to be a compelling argument for public justification — an argument that explains why public justification is not merely a pro tanto value but a weighty constraint on the behaviour of states — we need to know more about the social trust commitment. In particular, we need to know why the creation and maintenance of a system of social trust matters so much. This, then, is the second premise of Vallier's argument for public justification.

---

[44] Ibid., p. 83.
[45] Ibid., p. 4.
[46] Ibid.
[47] Ibid., p. 7.
[48] Ibid., p. 109.
[49] Ibid., p. 2.

For Vallier, a system of social trust realises two kinds of value: teleological and deontological. Roughly, the "teleological" value of a system of social trust refers to the good consequences produced by social trust. Vallier draws extensively on empirical studies to demonstrate that societies with a high level of social trust exhibit a range of good outcomes: economic growth, educational attainment, economic equality, cooperation between strangers, and so on. Importantly, a high level of social trust allows people to trust those outside of their immediate in-group and to form valuable relationships such as friendship and romantic love with people of diverse backgrounds. "A society without social trust," Vallier states, again with reference to the idea of war, "is one that is likely to be in a kind of cold (or hot) civil war, with persons constantly pitted against one another and living in persistent fear that conflict can break out at any time. Justified social trust allows people to rationally let their guard down in order to form loving relations with others".[50]

However, as Vallier explains, a system of social trust cannot be justified merely on the basis of its teleological value.[51] After all, if social trust is justified merely by reference to its good consequences, we would have no reason to be trustworthy if better consequences could be achieved through free riding, opportunism, and other small-scale acts of untrustworthy behaviour. A public justification principle grounded in the teleological value of a system of social trust would fail to explain why the public justification principle cannot generally be overridden by competing moral values. Vallier thus argues that social trust also realises the deontological value of respect for persons, and this explains why we should adhere to the public justification principle even when its contravention might produce better consequences.[52]

Vallier connects respect to a system of social trust by means of two arguments. The first is the "Argument from Trustworthiness".[53] According to this argument, "within a system of social trust, respect for others requires that we be trustworthy to live up to the trust that others have placed in us".[54] "If we fail to return trustworthy behaviour in response to the trust that others have placed

[50] Ibid., p. 58.
[51] Ibid., pp. 58-63.
[52] Ibid., pp. 66-78.
[53] Ibid., p. 66-9.
[54] Ibid., pp. 66-7.

in us," explains Vallier, then "we are manipulative and deceptive and show a general indifference to the worth of others".[55] The second argument for the deontological value of a system of social trust is the "Argument from Accountability".[56] According to this argument, within a system of trust "we will hold other moral agents accountable for wrongdoing only in cases where they see themselves as having sufficient moral reason of their own to comply with the moral rule they have violated". [57] By "confining our practice of accountability towards those we trust in this way, we express due regard for them".[58]

In short, then, a system of social trust has both teleological value (the system produces good outcomes such as economic growth, economic equality, relationships within and between social groups) and deontological value (a system of social trust in which we are trustworthy, and in which we hold others accountable for wrongdoing only when they see themselves as having reasons to comply with the violated rule, realizes respect for persons).

2 Getting the comparison right: why social trust does not require public justification

In what follows I shall grant that the creation and maintenance of a system of social trust between highly diverse individuals is of very great value, and that if social trust *did* require public justification then this would constitute a powerful case for the public justification principle. Instead, what I shall challenge is Vallier's claim that social trust requires public justification.

Of course, Vallier's claim looks plausible if we take his examples — which involve Marxism, a modus vivendi based on mere power relations, morally unjustified policies, and aggressive carnivorism — as paradigms of non-public reason. Vallier's imagined episode in which a small band of Marxists tries to ostracize business owners and to force them, in ways reminiscent of the Soviet Union, into handing their capital over to their employees would indeed plausibly undermine social trust. But appeals to cases such as these are

[55] Ibid., p. 8.
[56] Ibid., p. 69-73.
[57] Ibid., p. 70.
[58] Ibid., p. 71.

unfortunate and tendentious because it is unclear whether our judgement that social trust would be undermined is elicited by the Marxists' failure to publicly justify their behaviour per se or instead by a range of other factors present in the description of the case that are strictly distinct from, but easily conflatable with, such failure: for instance, that the Marxist reasoning is morally unsound, that the Marxists show utter disregard for the business owners' perspective,[59] and that the Marxists pursue their ends in belligerent, antagonistic, Soviet-style manner. Can we really be confident that our judgement that the Marxists would undermine social trust is a product of their failure to reason publicly, and not a product of these other contingently associated factors?

Another way of putting this point is that to reject public justification is not to say that "politics is war", that "anything goes", or that there are no constraints at all on political discourse and decision-making. Those who reject public justification do not in general claim that, therefore, citizens may engage in mindless sectarianism, may flagrantly disregard other citizens' perspectives, and may browbeat and ostracize others on the basis of discredited, morally unjustified ideologies. So once we consider more plausible and attractive alternatives to public justification, it becomes less clear why all alternative models and ideals short of public justification must vitiate social trust.

I should say that Vallier is not alone in setting up dichotomous comparisons that stack the deck in favour of public reason liberalism. The strategy of appealing to tendentious alternatives to public justification as a way of motivating the public justification principle is ubiquitous in the relevant literature.[60] Public reason liberals going back to Rawls often get undue rhetorical and argumentative mileage out of the suggestion that, if one rejects public justification, then *of course* one must be some kind of theocrat or imperialist. Thus Rawls defends his view, in his *Commonweal* interview, by suggesting that we have to choose between either public reason liberalism or the wars of religion:

---

[59] This is distinct from failure to provide public justification because it is possible to show regard for someone's perspective in ways that are different from public reason liberals' way of showing such regard. See Section 2.1.

[60] For an excellent discussion of this point, see Eberle, C., *Religious Conviction in Liberal Politics* (Cambridge: Cambridge University Press, 2002), p. 109-15.

> How many religions are there in the United States? How are they going to get on together? One way, which has been the usual way historically, is to fight it out, as in France in the sixteenth century. That's a possibility. But how do you avoid that? See, what I should do is to turn around and say, What's the better suggestion, what's your solution to it? And I can't see any other solution....Again, what's the alternative?[61]

Similarly, Dworkin seeks to bolster the plausibility of political liberalism by contrasting it to the practice of inflammatory conservative evangelists such as Jerry Falwell, as if to reject the former is to accept the latter:

> Government must be neutral in matters of personal morality, [and] must leave people free to live as they think best so long as they do not harm others. But the Reverend Jerry Falwell, and other politicians who claim to speak for some "moral majority", want to enforce their own morality with the steel of the criminal law. They know what kind of sex is bad, which books are fit for public libraries, what place religion should have in education and family life, when human life begins, that contraception is sin, and that abortion is capital sin. They think the rest of us should be forced to practice what they preach.[62]

More recently, Eberls-Duggan similarly motivates the public justification principle by unduly narrowing the moral landscape; invoking, like Vallier, the metaphor of war, she contrasts public reason liberalism to the view that other "citizens are, to the extent that they disagree with us, understood as enemy combatants in a struggle to impose our views".[63] And Gaus trades on a characterisation of violators of the public justification principle that tendentiously casts them in an unappealing light: anyone who denies the public justification principle must, suggests Gaus, see herself as belonging to a dictatorial and superior priestly class. "A morality for our democratic age," Gaus says, "takes seriously that people of good will in a diverse society have differing views of what constitutes an acceptable way of living together, and there is no

---

[61] Rawls, J., *Collected Papers*, ed. S. Freeman (Cambridge, MA: Harvard University Press, 1999), p. 620.
[62] Dworkin, R., "Neutrality, Equality, and Liberalism" in D. MacLean and C. Mills (eds), *Liberalism Reconsidered* (Totowa, NJ: Rowman and Allanheld, 1983), p. 1.
[63] Ebels-Duggan, K., "The Beginning of Community: Politics in the Face of Disagreement", *Philosophical Quarterly* 60 (2010), p. 56.

priestly class who can dictate to the rest how we are to live together."[64] We should reject the picture of politics, he says, on which "tablets inscribed with the truth…are sent down from on high to the rest to guide them…and *hoi polloi* who do reason the matter through, yet cannot grasp the truth of the message, are told that they really do have reason to do as they are told".[65]

In response to this tendency, I spell out a more sympathetic alternative model — namely, Eberle's "ideal of conscientious engagement" — which, I contend, would create and maintain a system of social trust. To be clear about my strategy in this section: I do not intend to endorse Eberle's account. Perhaps this account has *other* important defects.[66] My claim is just that Eberle's account is on a par with public justification so far as social trust is concerned; thus Vallier's appeal to social trust does not give us any reason to prefer public reason liberalism to alternative liberal-democratic theories such as Eberle's that lack a public justification requirement. Another way of putting this is that, even if considerations of social trust do generate important constraints on political discourse and decision-making, such considerations *underdetermine* the nature of those constraints: social trust does not tell us why we should opt for the public justification constraint rather than some other constraint.

2.1 Eberle's ideal of conscientious engagement

One of most attractive and well-developed alternative conceptions of deliberative-democratic discourse and decision-making is Eberle's ideal of conscientious engagement.[67] Indeed, Eberle develops his ideal of conscientious engagement explicitly as part of a critique of public reason liberals' argument

[64] Gaus, G., "On Theorizing About Public Reason", *European Journal of Analytic Philosophy* 9 (2013), p. 84.
[65] Gaus, G., "The Good, the Bad, and the Ugly", p. 577.
[66] For instance, in Section 3 of Chapter 4 I argue against this account on other grounds.
[67] Recently, Pallikkathayil has developed another attractive ideal of political discourse and decision-making that has much in common with Eberle's ideal of conscientious engagement: namely, the "duty of responsiveness". See Pallikkathayil, J., "Disagreement and the Duties of Citizenship", *American Philosophical Quarterly* 56 (2019), 71-82. "The duty of responsiveness," explains Pallikkathayil, "directs citizens to be responsive to the presence of disagreement in their political advocacy, but in a very different way than the duty of restraint would recommend" (p. 71). Another account that looks likely to secure social trust despite lacking a public justification principle is Christiano's "wide view of deliberative democracy": see Christiano, T., "Must Democracy Be Reasonable?", *Canadian Journal of Philosophy* 39 (2009), pp. 1-34; Christiano, T., *The Constitution of Equality: Democratic Equality and Its Limits* (Oxford: Oxford University Press, 2008).

from respect to public justification: respect does not require public justification, Eberle argues, because citizens who adhere to the ideal of conscientious engagement also express respect for their fellow citizens.[68]

Eberle's ideal of conscientious engagement is constituted by a variety of moral constraints. One of these constraints is that a citizen should only advance laws and policies that enjoy a "high degree of rational justification" (understood in terms of considering the relevant evidence, discharging various epistemic obligations governing belief formation, and so on).[69] Another is that a citizen should not support a law or policies that "denies the dignity of her compatriots".[70] A third constraint is that, during political deliberation, a citizen should "listen to her compatriots…with the intention of learning from them".[71] Fourth, the ideal of conscientious engagement even requires a citizen to "*pursue* public justification for her favoured coercive policies".[72] But crucially — and this is what distinguishes the ideal of conscientious engagement from public reason liberalism — citizens need not withhold their support for policies if, after looking for reasons that are acceptable to other citizens, no such reasons are ultimately forthcoming. "So long as a…citizen sincerely and conscientiously *attempts* to articulate a rationale for his favoured coercive policies that will be convincing to his compatriots, then he has thereby discharged his obligation to respect them — even if his attempt ends in failure".[73]

With the ideal of conscientious engagement in view, we can more accurately test whether the failure to abide by public justification would undermine social trust. In characterizing those who fail to abide by public justification as supporting policies that we have independent reasons to think are wrong (such as the overthrow of capitalism or the imposition of carnivorism) or as exhibiting unattractive character traits (such as arrogance, dogmatism, and a

[68] Billingham has since appealed to the ideal of conscientious engagement to show that *civic friendship* does not require public justification. See Billingham, P., "Does Political Community Require Public Reason? On Lister's Defence of Political Liberalism", *Politics, Philosophy & Economics* 15 (2016), pp. 20-41. Here I continue the generalization of Eberle's argumentative strategy by applying it to the value of social trust.
[69] Eberle, C., *Religious Conviction in Liberal Politics*, p. 104.
[70] Ibid., p. 105.
[71] Ibid.
[72] Ibid.
[73] Ibid., p. 82.

willingness to ostracize others), Vallier introduces rhetorical distractions. To properly test the claim that social trust requires public justification, we should screen off these extraneous factors and should imagine instead someone who, while refusing to abide by public justification, is sympathetic and admirable in other respects — someone, in other words, who abides by the ideal of conscientious engagement.[74]

Consider, then, the following portrait:

> *Edwin.* Ordinarily, the policies that Edwin supports are not just rationally justified (in the sense that they are well supported by the available evidence) but also interpersonally or publicly justified (in the sense that they are acceptable to other citizens). But this isn't always the case, especially when it comes to environmental issues. For Edwin subscribes to a fairly unorthodox ecological philosophy according to which the Earth is regarded "not as an inanimate globe of rock, liquid and gas driven by geological processes, but as a sort of biological superorganism, a single life-form, a living planetary body that adjusts and regulates the conditions in its surroundings".[75] Deeply concerned about the human impact on the environment, Edwin supports and campaigns for far-reaching measures to address the various forms of environmental degradation. He is aware that some reasonable citizens reject this philosophy and the environmental policies that flow from it. He is aware, that is, that some citizens see no sufficient reason, from within their system of beliefs and values, for such radical environmental policies — even after a respectable amount of reflection on the matter. Edwin has engaged such citizens in many conversations with the intention of learning from them; and he has tried his very best, in the course of these many civil and polite deliberations, to articulate a rationale (in terms of familiar liberal values such as fairness to future generations, equality, freedom, and so on) for his favoured policies that will be persuasive to his compatriots. But ultimately he is unsuccessful and it turns out (ex hypothesi) that the radical environmental policies that he favours depend *ineliminably* on acceptance of this controversial

[74] See ibid., p. 112.
[75] Huggett, R., *Physical Geography: The Key Concepts* (Oxford: Routledge, 2010), p. 84. See also Lovelock, J., *Gaia: A New Look at Life on Earth* (Oxford: Oxford University Press, 1979).

> holistic ecological doctrine. Edwin regards this situation as tragic — he wishes he did not have to impose his views on citizens who from their point of view lack sufficient reasons to accept the policy. Yet, given the significance of these issues, he cannot, in conscience, do otherwise than to support and campaign for this unpopular environmental policy. Edwin thus violates (Vallier's version of) the public justification principle because (to borrow Vallier's wording) it is not the case that "each member of the public has sufficient intelligible reason to comply with and internalize" the policies he advances.[76]

The question we have to ask is: could a political community populated by citizens like Edwin achieve social trust? It is difficult to see why not. Edwin is conscientious and engaged. The policies he supports are not mere hunches; they are not based on whims or on guesswork. He spends long hours thinking deeply about environmental issues, and only supports policies for which he has a high degree of rational justification. He ensures that the policies he supports do not deny the dignity of his compatriots. He is cooperative, deliberative, civil, and polite. He takes no joy in imposition and tries his best to inhabit the distinctive perspectives of his fellow citizens and to find justifications that they will find convincing.

Edwin's behaviour is thus nothing like various examples that Vallier distractingly uses to portray the violator of public justification. Unlike *modus vivendi*, Edwin is morally and not merely strategically committed to his society's main social and political institutions. Unlike the "band of Marxists", Edwin is not hostile and confrontational; he is not in the grip of morally discredited ideologies; he is sensitive to the disagreements of his fellow citizens and listens to them with the intention of learning from them; and he has done what he can pursue a public justification for his favoured policy. As Eberle would say of Edwin, "his support for [far-reaching environmental policies] is motivated, not by gleeful imposition of power and not by indifference to his compatriots' cares and concerns, but by his commitment to act in accord with moral convictions for which he has, by hypothesis, a high degree of rational justification".[77] More generally, Edwin is not plausibly described as being "at war" with his fellow

[76] The case of Edwin is based on Eberle's case of Elijah. See Eberle, C., *Religious Conviction in Liberal Politics*, pp. 112-5.
[77] Ibid., p. 113.

citizens; his support for environmental policies, despite not being acceptable at the tribunal of each reasonable citizen's judgement, is not plausibly described as "institutionalized aggression" or as a "small-scale act of war" against those who disagree. Nor, it is worth adding, does he resemble Rawls's, Dworkin's, Gaus's or Ebels-Duggan's respective portrayals of violators of public justification: Edwin is not infused with apocalyptic religious fervour to "fight it out"; he does not follow either the substantive policies or the fulminatory style of sectarian evangelists such as Jerry Falwell; he does not regard himself as belonging to any superior "priestly class"; and he does not view other citizens "enemy combatants" to be defeated in a struggle for power.

So once we test Vallier's claim against the proper alternative — the comparison with all the morally relevant variables appropriately controlled — it becomes unclear why social trust requires public justification. Even if the particular alternative practices to which Vallier appeals would fail to sustain social trust, it is unclear why general adherence to the ideal of conscientious engagement could not sustain social trust by "ameliorat[ing] the tensions and frustrations inevitably generated by our disagreement" regarding "the application of power in a pluralistic society":

> If our concerns are simply discounted, if we aren't consulted, if our compatriots impose their will on us in a whimsical fashion, if our compatriots don't even attempt to discern some rationale that might convince us that some coercive law is morally appropriate, then we'll find it much more difficult to acquiesce [and to trust others] than if our compatriots hear our concerns, consider our objections on their merits, make a good faith effort to explain why that coercive law is in fact morally defensible, and, throughout the process, make clear that they are willing to abide by such constraints out of respect for our dignity as persons.[78]

Why, then, must the presence of people like Edwin within society undermine the formation and maintenance of a system of social trust? Is there anything Vallier can say at this point except to observe that social trust would not form amongst those citizens who reasonably reject Edwin's favoured policies — an

[78] Ibid., pp. 106, 107-8.

observation that would amount to a bare assertion of the indispensability of public justification to social trust and that would thus approach circularity?

3 How much social trust?

In this section and the next, I defend the argument of Section 2 against two objections. The first objection is that, even if deliberative democracy would create *some degree* of social trust, it would create *less* social trust than public justification creates. Even if failure to reason publicly doesn't totally destroy social trust, mightn't it diminish social trust? So can we be confident that Eberle's account would sustain as *high* a level of social trust as that sustained under public reason liberalism?

Since no actual observable society perfectly corresponds either to deliberative democracy or to public reason liberalism, it is difficult to make precise and reliable empirical statements about the relative levels of social trust under various political theories. We are in the realm of empirical conjecture. Nonetheless, my own sense is that, if anything, the level of social trust might actually be *higher* under Eberle's ideal than it would be under the public justification principle. Deliberative democracy, that is, does a better job at realizing social trust than does public reason liberalism. There are two main reasons for this, having to do with *integrity* and *sincerity*.

First, then, many critics of public reason liberalism have argued that the public justification principle undermines the wholeness or integrity of religious citizens by requiring them, when engaged in politics, to put their most deeply held yet non-public religious commitments to sleep. While this objection applies most naturally to "consensus" models of public reason, it also applies to Vallier's preferred "convergence" model of public reason insofar as citizens are required to withhold support for policies which they privately endorse on religious grounds but which they cannot publicly justify to all other citizens. By requiring religious citizens to "split" their public and private identities in this way, the public justification principle, it is claimed, prevents citizens of faith from living a religiously integrated life. Perry puts this point forcefully:

> One's basic moral/religious convictions are (partly) self-constitutive and are therefore a principal ground — indeed, the principal ground — of political deliberation and choice. To "bracket" such convictions is therefore to bracket — to annihilate — essential aspects of one's very self. To participate in politics and law — in particular to make law, to break law, or to interpret law — with such convictions bracketed is not to participate as the self one is but as some one — or rather some thing — else.[79]

By contrast, Eberle's ideal of conscientious engagement does not impel religious citizens to "translate" their religious reasons into secular reasons when entering the political sphere and thus does not threaten their wholeness or integrity. Rather than taking religious perspectives off the table, politically speaking, the ideal of conscientious engagement encourages citizens to actively and seriously engage with such perspectives. What Rawls calls the "strains of commitment" are thus lower for religious citizens under these forms of deliberative democracy than they are under public reason liberalism.[80] So, insofar as one is less likely to be a committed and responsible member of a system of social trust the ground rules of which threaten one's integrity, it is plausible to suppose that deliberative democracy would be better able to foster social trust (at least in one respect, namely across the religious/secular divide) than would public reason liberalism.

The second (and related) reason why deliberative democracy would plausibly realize more social trust than public reason liberalism concerns openness and sincerity. One common objection to public reason liberalism is that it permits, encourages, and perhaps even requires citizens to be insincere in political deliberations. Consider a pro-life activist whose opposition to abortion stems from disputed religious beliefs. If this activist wants to engage in political debate consistent with public reason liberalism, it appears that she will need to *dissimulate*: she will need, that is, to find suitably public arguments against abortion (e.g. the slippery slope from abortion to infanticide) even though

---

[79] Perry, M., *Morality, Politics and Law* (Oxford: Oxford University Press, 1988), pp. 181-2.
[80] Rawls, J., *Political Liberalism*, p. 17.

these are not the arguments that *really* move her and are merely "puppets manipulated from behind the scenes" by her religious reasons.[81]

By contrast, Eberle's ideal of conscientious engagement does not encourage any such insincerity or dissimulation. On these views, citizens may appeal to the whole truth as they see it and are invited to deliberate openly and frankly (yet respectfully), offering in good faith what they take to be the best reasons that apply in a given case. So, assuming that openness and sincerity contribute to social trust, the deliberative-democratic approach to political discourse and decision-making appears likely to produce more, rather than less, social trust than public reason liberalism.

Let me close this section by stressing its tentative character. It is difficult to make reliable conjectures about the level of social trust that would be sustained under one or another political theory. I am doubtful that there will be much to choose between the level of social trust sustained under Vallier's version of public reason liberalism and Eberle's version of deliberative democracy. The point is just that, *if anything*, deliberative democracy looks the more likely candidate to generate high social trust because it is less restrictive and exclusionary (whilst still requiring citizens to avoid sectarianism and to be attuned to the reasonable pluralism within society).

4 Social trust for the right reasons?

At this point, a defender of Vallier's social trust argument might respond as follows:

> It is true that familiar deliberative-democratic alternatives to the public justification principle could sustain a system of social trust (Section 2) and it is even true that they could sustain a *high degree* of social trust (Section 3). But they would not create high social trust *in the right way and for the right reasons*. We're not just trying to establish and maintain a system of social trust any which way: social trust could, after all, be

[81] Rawls, J., *Political Liberalism*, p. 454. For an insincerity objection directed at Vallier's preferred convergence form of political liberalism, see Quong, J., *Liberalism Without Perfection*, pp. 265-73.

> established and maintained by brainwashing.[82] Instead, we're trying to establish and maintain a system of social trust for the morally appropriate kind of reasons. And the morally appropriate kind of reasons are reasons that are recognized as valid by each member of a system of social trust. Thus Vallier says: "only publicly justified moral rules can sustain social trust *in the right way*".[83] And: "moderately idealized intelligible reasons *appropriately* morally incentivize trustworthiness and are the only reasons that do so".[84] And: "the foundation for public justification lies in its capacity to promote social trust *in the right way*".[85] In short, then, only public justification could establish and maintain social trust for the right reasons.

I believe that this move to social trust for the right reasons fails to rescue the argument for public justification. To see why, it is helpful to first consider a parallel move made by those who seek to argue from the value of stability to public justification.

One prominent answer to the question of why we should reason publicly involves grounding public justification in the value of *stability*. This grounding was notably hinted at in the introduction to Rawls's *Political Liberalism* in which he discusses the European Wars of Religion. Rawls explains that this period saw a "clash between salvationist, creedal, and expansionist religions" — a "mortal conflict moderated only be circumstance and exhaustion".[86] "Political liberalism starts by taking to heart the absolute depth of that irreconcilable latent conflict."[87] The public justification approach — in which we forego the "zeal to embody the whole truth in politics" and seek to "bypass religion and philosophy's profoundest controversies so as to have some hope of uncovering a basis of a stable overlapping consensus" — can thus be understood for Rawls to be grounded, at least in part, in considerations of stability.[88]

---

[82] See Vallier, K., *Must Politics Be War?*, p. 75.
[83] Ibid., p. 77.
[84] Ibid., p. 118.
[85] Ibid., p. 6.
[86] Rawls, J., *Political Liberalism*, p. xxv-xxvi.
[87] Ibid., p. xxvi.
[88] Ibid., p. 152, 442.

Of course, the obvious problem with this argument is that, as an empirical matter, it is highly implausible that social stability causally depends on public justification. The failure to adhere to public justification constraints does not cause war and bloodshed — as evidenced by the fact that most modern liberal democracies are not "orders of public reason" and yet are not marred by mortal conflict. It is true that *some* alternatives to the public justification principle — such as the aggressively salvationist religious principle of the seventeenth century — lead to instability. But that is not the relevant and appropriate comparison. As stated in Section 2, the proper comparison is between public reason liberalism and some attractive deliberative democratic theory that lacks a public justification principle.

At this point, the relevant public reason liberals counter by clarifying that it is not *empirical* stability that depends on public justification but rather some kind of *normative* stability — that is, stability for the right reasons. Unfortunately, these public reason liberals are never particularly clear on what exactly the "right" reasons for stability are. But, roughly speaking, a society is "stable for the right reasons" when each new generation of citizens freely endorses and affirms its political institutions in wide and general reflective equilibrium. As Rawls, for instance, puts it, "the kind of stability required of justice as fairness is based, then, on its being a liberal political view, one that aims at being acceptable to citizens as reasonable and rational, as well as free and equal, and so as addressed to their public reason".[89]

The problem with this move from empirical stability to normative stability, however, is that makes the stability story into another way of telling the public justification story. If a society is stable for the right reasons when its laws and institutions are "acceptable to citizens as reasonable and rational", then stability for the right reasons merely *restates*, rather than justifies, the public justification principle. If, that is, a political order's being stable for the right reasons *just is* its being acceptable to all reasonable citizens, then stability for the right reasons cannot explain why we ought to care about public justification. Of course, it may be that there are deeper reasons for caring about stability for the right reasons. Perhaps, for instance, a society that achieves stability for the right reasons realizes a valuable kind of political autonomy. But

[89] Ibid., p. 143.

then it will be *those* autonomy-based reasons, rather than normative stability, that do the moral work in justifying public reason. In the absence of deeper reasons of this kind, the appeal to normative stability will not persuade anyone who is not already persuaded by the public justification principle.[90]

Arguments from stability to public justification thus face a dilemma: either they are couched in terms of empirical stability, in which case they rely on the factually false premise that without public justification (e.g. under a system of deliberative democracy) society would fall into instability, bloodshed and civil war; or such arguments are couched in terms of normative stability, in which case they merely re-label or presuppose the public justification principle rather than providing compelling independent justification for it.

A parallel difficulty besets the move to social trust for the right reasons that is suggested at the beginning of this section. For the obvious question to ask is: What would it take to sustain social trust "in the right way"? What are the right reasons for social trust? Vallier is clear on this point: "*moderately idealized intelligible reasons* appropriately morally incentivize trustworthiness and are the only reasons that do so".[91] As he similarly puts this point:

> The person who has the reasons must be able to see, on reflection, that *the reasons are hers* and motivate her to act morally in accord with either her personal moral standard or the public moral standard prevalent in her social order. If she cannot see the reasons as hers according to her moral standards, then her reasons cannot drive socially trustworthy behavior, since trustworthy behavior requires an appreciation of moral reasons…Inaccessible, nonmotivating reasons cannot drive trustworthiness.[92]

The problem with this argument becomes clear when we recall the dialectical situation. The critic wants to know why we ought to abide by the public justification principle — why, to use Vallier's formulation of the public justification principle, we need to ensure that "each member of the public has sufficient intelligible reason to comply with and internalize" political laws and

[90] I thank Paul Billingham and Anthony Taylor for discussion of this point.
[91] Vallier, K., *Must Politics Be War?*, p. 118.
[92] Ibid., p. 79 (emphasis added).

institutions. In response, Vallier appeals to the idea of social trust: only publicly justified laws and policies would establish and maintain social trust. When the critic counters that alternative democratic theories such as Eberle's ideal of conscientious engagement seem perfectly capable of securing social trust, the response is that Eberle's ideal of conscientious engagement does not secure the right *kind* of social trust — does not, that is, secure social trust on the basis of "intelligible reasons".

The problem, then, is that, as with the appeal to stability for the right reasons, this begs the crucial question. We want to know *why* it matters so much that each citizen has intelligible reasons to accept laws and policies; pointing to the fact that, by ensuring that each citizen has intelligible reasons to accept laws and policies, we would create a system of social trust on the basis of intelligible reasons (as opposed to a system of social trust based on commitment to an ideal of conscientious engagement) is no argument at all. Someone who is not antecedently persuaded of the public-reason-giving requirement would fail to be persuaded by the tautological claim that only public reason would bring about social trust on the basis of public reasons. The foundational question is just pushed back. Instead of asking why we should care about ensuring that all citizens have intelligible reasons to comply with laws and policies (as opposed to ensuring that an ideal of conscientious engagement is satisfied), we will now just want to ask: *Why should we care about establishing social trust driven by intelligible reasons, as opposed to establishing social trust driven by joint commitment to an ideal of conscientious engagement?*

It is true that Vallier does discuss the value a system of social trust — a discussion that I summarize in Section 1.3. But that discussion is more geared towards the general and non-comparative question of *why a system of social trust matters* rather than the specific and comparative question of *why a system of social trust driven by intelligible reasons, as opposed to a system of social trust driven by joint commitment to deliberative democracy, matters*. Vallier's account of the teleological and deontological value of a system of social trust cannot answer this fine-grained and comparative question because much if not all of this value could also be realized by a system of social trust driven by joint commitment to deliberative democracy.

Consider first Vallier's claim that a system of social trust has *teleological* value: a system of social trust, Vallier says, produces good outcomes such as economic growth, educational attainment, and the possibility of interpersonal relationships between persons of diverse backgrounds. It is highly implausible that these outcomes are tied to something so specific as a system of trust driven by intelligible reason and that they would not also be produced by a system of trust driven by joint commitment to deliberative democracy. Is it really plausible to suppose that, in a society governed by Eberle's ideal of conscientious engagement, citizens would have lower educational attainment or that individuals from diverse backgrounds would not be able to form romantic relationships? Certainly Vallier's empirical evidence does not support this: he appeals to empirical studies showing the role of social trust in producing good outcomes, but the liberal democracies from which these studies are drawn are not "orders of public reason" or well-ordered Rawlsian societies, and so these studies do not support (and may indeed refute) the claim that good outcomes such as economic growth and interpersonal relationships are tied exclusively to a system of social trust driven by intelligible reasons in particular.[93]

Consider next Vallier's claim that a system of social trust realizes the *deontological* value of respect for persons. Again, it is unclear why only a system of social trust driven by intelligible reasons — and not a system of social trust driven by joint commitment to Eberle's ideal of conscientious engagement — could realise respect for persons. The argument here is similar to the argument in Section 2, but this time in relation to respect. Consider again Edwin, a citizen who adheres to the ideal of conscientious engagement. He thinks hard about issues and only supports policies for which he enjoys a high degree of rational justification. He is cooperative, deliberative and polite. He is

[93] Perhaps *judges* in liberal democracies such as the UK and USA are a particular group whose behaviour more closely approximates the ideal of public reason than the ideal of conscientious engagement. (I thank David Miller for this point.) But I see no reason to think (and Vallier provides no evidence for thinking) that the *general* practice of moral and political discourse and decision-making in liberal democratic societies such as the UK and USA more closely approximates Vallier's public justification principle than Eberle's ideal of conscientious engagement (and thus that the teleological benefits of social trust in liberal democratic societies can be chalked up exclusively to public reason liberalism) — this especially so in the USA, where appeals to religion in public life can be quite common. Indeed, after his famous reference to the US Supreme Court as "the exemplar of public reason", Rawls adds that "it is the *only* branch of government that is visibly on its face the creature of that reason and that reason alone" (Rawls, J., *Political Liberalism*, 231, 235, emphasis added).

not "manipulative and deceptive" and does not "show a general indifference to the worth of others", in Vallier's terms.[94] Edwin engages his fellow citizens in political debate and listens to them with the intention of learning from them. He tries in good faith to find rationales and arguments that will be persuasive to those citizens who disagree with him; but, sometimes, he is ultimately unsuccessful in doing so. In light of all this, it is difficult to see how Edwin disrespects for his fellow citizens just because he does not abide by the public justification principle. Indeed, Edwin's behaviour strikes me as not just not disrespectful but in fact morally exemplary. What basis, then, is there for saying that a system of trust populated by conscientious and engaged citizens like Edwin would fail to realize the deontological value of respect for persons? Why couldn't general adherence to an ideal of conscientious engagement sustain a system of social trust in a respectful manner?

In short, then, Vallier gives us no reason to think that there is anything special, in terms of teleological value or deontological value, about a system of trust driven by intelligible reasons, as opposed to a system of trust driven by joint commitment to deliberative democracy. In light of this, the argument from social trust to public justification faces a dilemma. If, on the one hand, Vallier's claim is that social trust in its ordinary empirical sense depends on public justification, then this looks implausible once we consider appropriate and sympathetic comparisons: for, as I argue in Section 2 and 3, it is difficult to see why high social trust could not be achieved outside of public justification (e.g. by Eberle's ideal of conscientious engagement). Yet if, on the other hand, Vallier's claim is that a special moralized kind of social trust — social trust on the basis of intelligible reasons — depends on public justification, then, as I argue in this section, while tautologically true, this simply presupposes the public justification principle rather than providing compelling independent justification for it.

[94] Ibid., p. 8.

Bibliography

Ackerman, Bruce. 1990. *Social Justice and the Liberal State*. New Haven, CT: Yale University Press.

Ackerman, Bruce. 1990. Neutralities. Pp. 29-43 in R. Douglass, G. Mara and H. Richardson (eds), *Liberalism and the Good*. London: Routledge.

Aristotle. 1990. *The Politics*, trans. H. Rackham. Cambridge, MA: Harvard University Press.

Arneson, Richard. 2003. Liberal Neutrality on the Good: An Autopsy. Pp. 191-218 in S. Wall and G. Klosko (eds), *Perfectionism and Neutrality: Essays in Liberal Theory.* Oxford: Rowman & Littlefield.

Baccarini, Elvio. Having a Reason and Distributive Justice in *The Order of Public Reason*. *European Journal of Analytic Philosophy*, 9, 25-51.

Barry, Brian. 1995. *Justice as Impartiality.* Oxford: Clarendon Press.

Barry, Brian. 2001. *Culture and Equality.* Cambridge: Polity Press.

Billingham, Paul. 2016. Does Political Community Require Public Reason? On Lister's Defence of Political Liberalism. *Politics, Philosophy & Economics*, 15, 20-41.

Billingham, Paul. 2017. Liberal Perfectionism and Quong's Internal Conception of Political Liberalism. *Social Theory and Practice*, 43, 79-106.

Billingham, Paul and Taylor, Anthony. 2018. Liberal Perfectionism, Moral Integrity, and Self-Respect. *The American Journal of Jurisprudence*, 63, 63-79.

Boettcher, James. 2007. Respect, Recognition, and Public Reason. *Social Theory and Practice*, 33, 223-49.

Bradley, Ben. 2012. Doing Away with Harm. *Philosophy & Phenomenological Research*, 85, 390-412.

Brighouse, Harry. 1995. Neutrality, Publicity, and State Funding of the Arts. *Philosophy & Public Affairs,* 24, 35-63.

Caney, Simon. 1991. Consequentialist Defences of Liberal Neutrality. *The Philosophical Quarterly*, 41, 457-77.

Caney, Simon. 1996. Impartiality and Liberal Neutrality. *Utilitas*, 8, 273-93.

Caney, Simon. 1998. Liberal Legitimacy, Reasonable Disagreement, and Justice. *Critical Review of International Social and Political Philosophy*, 1, 19-36.

Chan, Joseph. 2000. Legitimacy, Unanimity, and Perfectionism. *Philosophy & Public Affairs*, 29, 5-42.

Christiano, Thomas. 2008. *The Constitution of Equality: Democratic Equality and Its Limits*. Oxford: Oxford University Press.

Christiano, Thomas. 2009. Must Democracy Be Reasonable? *Canadian Journal of Philosophy*, 39, 1-34.

Clarke, Simon. 2014. Consequential Neutrality Revivified. Pp. 109-21 in R. Merrill and D. Weinstock (eds), *Political Neutrality: A Re-evaluation*. New York: Palgrave Macmillan.

Colburn, Ben. 2010. *Autonomy and Liberalism.* Oxford: Routledge.

Cordelli, Chiara. 2017. Neutrality of What? *Critical Review of International Social and Political Philosophy,* 20, 36-48.

Couto, Alexandra. 2014. *Liberal Perfectionism*. Berlin: De Gruyter.

Dancy, Jonathan. 2000. *Practical Reality.* Oxford: Clarendon Press.

Dworkin, Ronald. 1978. Liberalism. Pp. 113-43 in S. Hampshire (ed.), *Public and Private Morality*. Cambridge: Cambridge University Press.

Dworkin, Ronald. 1983. Neutrality, Equality, and Liberalism. Pp. 1-11 in D. MacLean and C. Mills (eds), *Liberalism Reconsidered*. Totowa, NJ: Rowman and Allanheld.

Dworkin, Ronald. 1985. *A Matter of Principle.* Oxford: Clarendon Press.

Dworkin, Ronald. 1995. Foundations of Liberal Equality. Pp. 190-306 in S. Darwall (ed.), *Equal Freedom: Selected Tanner Lectures on Human Values.* Ann Arbor: University of Michigan Press.

Ebels-Duggan, Kyla. 2008. Against Beneficence: A Normative Account of Love. *Ethics*, 119, 142-70.

Ebels-Duggan, Kyla. 2010. The Beginning of Community: Politics in the Face of Disagreement. *The Philosophical Quarterly*, 60, 50-71.

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

Enoch, David. 2015. Against Public Reason. Pp. 112-44 in D. Sobel, P. Vallentyne and S. Wall (eds), *Oxford Studies in Political Philosophy*, vol. 1. Oxford: Oxford University Press.

Feit, Neil. Forthcoming. Harming by Failing to Benefit. *Ethical Theory and Moral Practice*.

Fischer, John and Ravizza, Mark. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.

FitzPatrick, William. 2008. Moral Responsibility and Normative Ignorance. Answering a New Skeptical Challenge. *Ethics*, 18, 589-613.

Frankfurt, Harry. 1971. Freedom of the Will and the Concept of a Person. *Journal of Philosophy*, 68, 5-20.

Freeman, Samuel. 2007. *Justice and the Social Contract*. Oxford: Oxford University Press.

Galston, William. 1991. *Liberal Purposes: Goods, Virtues, and Diversity in the Liberal State*. Cambridge: Cambridge University Press.

Gaus, Gerald. 1996. *Justificatory Liberalism: An Essay on Epistemology and Political Theory*. Oxford: Oxford University Press.

Gaus, Gerald. 1999. Reasonable Pluralism and the Domain of the Political: How the Weaknesses of John Rawls's Political Liberalism can be Overcome by a Justificatory Liberalism. *Inquiry*, 42, 259-84.

Gaus, Gerald. 2003. Liberal Neutrality: A Compelling and Radical Principle. Pp. 137-65 in S. Wall and G. Klosko (eds), *Perfectionism and Neutrality: Essays in Liberal Theory*. Oxford: Rowman & Littlefield.

Gaus, Gerald. 2010. Coercion, Ownership, and the Redistributive State: Justificatory Liberalism's Classical Tilt. *Social Philosophy & Policy*, 27, 233-75.

Gaus, Gerald. 2011. *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*. Cambridge: Cambridge University Press.

Gaus, Gerald. 2011. A Tale of Two Sets: Public Reason in Equilibrium. *Public Affairs Quarterly*, 25, 305-25.

Gaus, Gerald. 2013. On Theorizing About Public Reason. *European Journal of Analytic Philosophy*, 9, 64-85.

Gaus, Gerald. 2014. The Good, the Bad, and the Ugly: Three Agent-Type Challenges to *The Order of Public Reason*. *Philosophical Studies*, 170, 563-77.

Goodin, Robert and Reeve, Andrew. 1989. Do Neutral Institutions Add Up to a Neutral State? Pp. 193-210 in R. Goodin and A. Reeve (eds), *Liberal Neutrality*. London: Routledge.

Griffin, James. 2008. *On Human Rights*. Oxford: Oxford University Press.

Hanna, Jason. 2018. *In Our Best Interest: A Defense of Paternalism*. Oxford: Oxford University Press.

Hills, Alison. 2007. Intentions, Foreseen Consequences and the Doctrine of Double Effect. *Philosophical Studies*, 133, 257-83.

Holtug, Nils. 2002. The Harm Principle. *Ethical Theory and Moral Practice*, 5, 357-89.

Huggett, Richard. 2010. *Physical Geography: The Key Concepts*. Oxford: Routledge.

Hurka, Thomas. 1993. *Perfectionism*. Oxford: Oxford University Press.

Jiang, Quing. 2013. *A Confucian Constitutional Order*, eds D. Bell and R. Fan and trans. E. Ryden. Princeton: Princeton University Press.

Kagan, Shelly. 1998. The Additive Fallacy. *Ethics*, 99, 5-31.

Kamm, Frances M. 1996. *Morality, Mortality, Vol. II.* Oxford: Oxford University Press.

King, Martin Luther. 1986. Remaining Awake Through a Great Revolution. Pp. 268-78 in J. Washington (ed.), *A Testament of Hope: The Essential Writings and Speeches of Martin Luther King, Jr.* New York: Harper Collins.

Klosko, George. 2003. Reasonable Rejection and Neutrality of Justification. Pp. 167-89 in S. Wall and G. Klosko (eds), *Perfectionism and Neutrality: Essays in Liberal Theory*. Oxford: Rowman & Littlefield.

Klosko, George. 2004. An Empirical Approach to Political Liberalism. Pp. 129-48 in S. Young (ed.), *Political Liberalism: Variations on a Theme*. Albany, NY: University of New York Press.

Kramer, Matthew. 1998. Rights Without Trimmings. Pp. 7-111 in M. Kramer, N. Simmonds, and H. Steiner (eds), *A Debate Over Rights*. Oxford: Oxford University Press.

Kramer, Matthew. 2011. *The Ethics of Capital Punishment: A Philosophical Investigation of Evil and Its Consequences*. Oxford: Oxford University Press.

Kramer, Matthew. 2014. *Torture and Moral Integrity: A Philosophical Enquiry*. Oxford: Oxford University Press.

Kramer, Matthew. 2017. *Liberalism with Excellence*. Oxford: Oxford University Press.

Kymlicka, Will. 1989. Liberal Individualism and Liberal Neutrality. *Ethics*, 99, 883-905.

Kymlicka, Will. 2001. Western Political Theory and Ethnic Relations in Eastern Europe. Pp. 13-105 in W. Kymlicka and M. Opalski (eds), *Can Liberal Pluralism be Exported?* Oxford: Oxford University Press.

Laborde, Cécile. 2017. *Liberalism's Religion*. Cambridge: Harvard University Press.

Larmore, Charles. 1987. *Patterns of Moral Complexity*. Cambridge: Cambridge University Press.

Larmore, Charles. 1999. The Moral Basis of Political Liberalism. *The Journal of Philosophy*, 96, 599-625.

Leland, R. J. and van Wietmarschen, Han. 2017. Political Liberalism and Political Community. *Journal of Moral Philosophy*, 14, 151-75.

Lister, Andrew. 2013. *Public Reason and Political Community*. London: Bloomsbury.

Lott, Micah. 2006. Restraint on Reasons and Reasons for Restraint: A Problem for Rawls' Ideal of Public Reason. *Pacific Philosophical Quarterly*, 87, 75-95.

Lovelock, James. 1979. *Gaia: A New Look at Life on Earth*. Oxford: Oxford University Press.

MacCallum, Gerald. 1967. Negative and Positive Freedom. *The Philosophical Review*, 76, 312-334.

Macedo, Stephen. 1990. *Liberal Virtues: Citizenship, Virtue, and Community in Liberal Constitutionalism.* Oxford: Clarendon Press.

Mang, Franz. 2013. Liberal Neutrality and Moderate Perfectionism. *Res Publica*, 19, 297-315.

de Marneffe, Peter. 2006. The Slipperiness of Neutrality. *Social Theory and Practice,* 32, 17-34.

McKenna, Michael. Where Frankfurt and Strawson Meet. *Midwest Studies in Philosophy*, 29, 163-80.

Mele, Alfred. 2006. *Free Will and Luck*. Oxford: Oxford University Press.

Merrill, Roberto. 2014. Introduction. Pp. 1-21 in R. Merrill and D. Weinstock (eds), *Political Neutrality: A Re-evaluation*. New York: Palgrave Macmillan.

Metz, Thaddeus. 2001. Respect for Persons and Perfectionist Politics. *Philosophy & Public Affairs*, 30, 417-42.

Mill, John S. 1985. *On Liberty*, ed. G. Himmelfarb. London: Penguin Books.

Mill, John S. 1998. *Utilitarianism*, ed. R. Crisp. Oxford: Oxford University Press.

Mulhall, Stephen and Swift, Adam. 1996. *Liberals and Communitarians*. Oxford: Blackwell Publishing.

Nickel, James. 2007. *Making Sense of Human Rights*. Oxford: Blackwell.

Neufeld, Blain. Reciprocity and Liberal Legitimacy. *Journal of Ethics & Social Philosophy*, 4, 1-7.

Nussbaum, Martha. 2011. Perfectionist Liberalism and Political Liberalism. *Philosophy & Public Affairs*, 39, 3-45.

Pallikkathayil, Japa. 2016. Neither Perfectionism nor Political Liberalism. *Philosophy & Public Affairs*, 44, 171-96.

Pallikkathayil, Japa. 2019. Disagreement and the Duties of Citizenship. *American Philosophical Quarterly*, 56, 71-82.

Patten, Alan. 2012. Liberal Neutrality: A Reinterpretation and Defense. *The Journal of Political Philosophy*, 20, 249-72.

Patten, Alan. 2014. *Equal Recognition: The Moral Foundations of Minority Rights.* Princeton: Princeton University Press.

Patten, Alan. 2017. Equal Citizenship, Neutrality, and Democracy: A Reply to Critics of Equal Recognition. *Critical Review of International Social and Political Philosophy*, 20, 127-41.

Pereboom, Derk. 2009. Free Will, Love, and Anger. *Ideas y Valores*, 141, 5-25.

Pereboom, Derk. 2014. *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press.

Perry, Michael. 1988. *Morality, Politics and Law*. Oxford: Oxford University Press.

Quong, Jonathan. 2006. Cultural Exemptions, Expensive Tastes, and Equal Opportunities. *Journal of Applied Philosophy*, 23, 55-73.

Quong, Jonathan. 2011. *Liberalism without Perfection.* Oxford: Oxford University Press.

Quong, Jonathan. 2012. *Liberalism Without Perfection*: A Précis. *Philosophy and Public Issues*, 2, 1-6.

Quong, Jonathan. 2013. On the Idea of Public Reason. Pp. 265-80 in J. Mandle and D. Reidy (eds), *The Blackwell Companion to Rawls.* Oxford: Wiley Blackwell.

Quong, Jonathan. 2014. *Liberalism Without Perfection*: Replies to Lister, Kulenovic, Zoffoli, Zelic, and Baccarini. *Philosophy and Society*, 25, 96-122.

Quong, Jonathan. 2014. What is the Point of Public Reason? *Philosophical Studies*, 170, 545-553.

Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Rawls, John. 1999. *A Theory of Justice: Revised Edition*. Cambridge, MA: Harvard University Press.

Rawls, John. *Collected Papers*, ed. S. Freeman. Cambridge, MA: Harvard University Press.

Rawls, John. 1985. Justice as Fairness: Political not Metaphysical. *Philosophy & Public Affairs*, 14, 223-51.

Rawls, John. 1996. *Political Liberalism.* New York: Columbia University Press.

Raz, Joseph. 1986. *The Morality of Freedom.* Oxford: Oxford University Press.

Raz, Joseph. 1989. Liberalism, Skepticism, and Democracy. *Iowa Law Review*, 74, 761-86.

Scanlon, T. M. 2008. *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge, MA: Harvard University Press.

Schouten, Gina. 2019. *Liberalism, Neutrality, and the Gendered Division of Labor.* Oxford: Oxford University Press.

Schwarzenbach, Sibyl. 1996. On Civic Friendship. *Ethics*, 107, 97-128.

Sher, George. 1997. *Beyond Neutrality: Perfectionism and Politics.* Cambridge: Cambridge University Press.

Sher, George. 2009. *Who Knew? Responsibility Without Awareness*. Oxford: Oxford University Press.

Sher, George. 2018. Confessions of a Quidnunc. *The American Journal of Jurisprudence*, 63, 49-61.

Simmons, John. 1993. *On the Edge of Anarchy: Locke, Consent and the Limits of Society*. Princeton: Princeton University Press.

Sommers, Tamler. 2007. The Objective Attitude. *The Philosophical Quarterly*, 57, 321-41.

Strawson, Peter. 2003. Freedom and Resentment. Pp. 72-93 in G. Watson (ed.), *Free Will*. Oxford: Oxford University Press, 2003.

Swanton, Christine. 1980. The Concept of Interests. *Political Theory*, 8, 83-101.

Tadros, Victor. 2016. *Wrongs and Crimes*. Oxford: Oxford University Press.

Tasioulas, John. 2015. On the Foundations of Human Rights. Pp. 45-70 in R. Cruft, M. Liao and M. Renzo (eds), *Philosophical Foundations of Human Rights*. Oxford: Oxford University Press.

Taylor, Anthony. 2018. Public Justification and the Reactive Attitudes. *Politics, Philosophy & Economics*, 17, 97-112.

Thomson, Judith J. 1991. Self-Defense. *Philosophy & Public Affairs*, 20, 283-310.

Turner, Piers N. 2018. Social Morality in Mill. Pp. 375-400 in P. N. Turner and G. Gaus (eds), *Public Reason in Political Philosophy: Classic Sources and Contemporary Commentaries.* New York: Routledge.

Vallier, Kevin. 2014. *Liberal Politics and Public Faith: Beyond Separation.* Oxford: Routledge.

Vallier, Kevin. 2019. *Must Politics Be War? Restoring Our Trust in the Open Society.* Oxford: Oxford University Press.

Van Schoelandt, Chad. 2015. Justification, Coercion and the Place of Public Reason. *Philosophical Studies*, 172, 1031-50.

Waldron, Jeremy. 1987. Theoretical Foundations of Liberalism. *Philosophical Quarterly*, 37, 127-50.

Waldron, Jeremy. 1993. *Liberal Rights: Collected Papers 1981 – 1991.* Cambridge: Cambridge University Press.

Waldron, Jeremy. 1993. Special Ties and Natural Duties. *Philosophy & Public Affairs*, 22, 3-30.

Wall, Steven. 1998. *Liberalism, Perfectionism and Restraint*. Cambridge: Cambridge University Press.

Wall, Steven. 2009. Perfectionist Politics: A Defense. Pp. 99-117 in T. Christiano and J. Christman (eds), *Contemporary Debates in Political Philosophy*. Oxford: Blackwell.

Wall, Steven. 2013. Enforcing Morality. *Criminal Law and Philosophy* 7, 455-71.

Wall, Steven. 2013. Moral Environmentalism. Pp. 93-114 in C. Coons and M Weber (eds), *Paternalism: Theory and Practice*. Cambridge: Cambridge University Press.

Weithman, Paul. 2010. *Why Political Liberalism?* Oxford: Oxford University Press.

Wellman, Christopher. 1996. Liberalism, Samaritanism, and Political Legitimacy. *Philosophy & Public Affairs* 25, 211-37.

Wilkins, Burleigh. 1997. A Third Principle of Justice. *The Journal of Ethics,* 1, 355-74.

Yuracko, Kimberly. 2003. *Perfectionism and Contemporary Feminist Values*. Bloomington: Indiana University Press.