

JUSTIFICATION AND SOCIAL MORALITY

by

Chad Van Schoelandt

---

Copyright © Chad Van Schoelandt 2015

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PHILOSOPHY

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2015



THE UNIVERSITY OF ARIZONA  
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Chad Van Schoelandt, titled Justification and Social Morality and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

\_\_\_\_\_  
Gerald Gaus Date: 4/30/2015

\_\_\_\_\_  
Thomas Christiano Date: 4/30/2015

\_\_\_\_\_  
Steven Wall Date: 4/30/2015

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

\_\_\_\_\_  
Dissertation Director: Gerald Gaus Date: 4/30/2015



## STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of the requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: Chad Van Schoelandt



## Acknowledgements

This research was made possible by generous support from the Earhart Foundation, the Institute for Humane Studies, the Mercatus Center of George Mason University and the University of Arizona Center for the Philosophy of Freedom. Work for this dissertation has been presented or work-shopped at Arizona State University, Tulane University, the University of Virginia, the 2013 Rocky Mountain Philosophy Conference, the 2015 American Philosophical Association Pacific Division Meeting, the International Consortium for Advanced Research in Social and Political Philosophy, the Institute for Humane Studies, the Mercatus Center, and the University of Arizona. Helpful feedback on published parts of the dissertation has come from the editors and reviewers of *The Philosophical Quarterly* and *Philosophical Studies*.

The list of people to acknowledge for comments and discussion over the years is vast, and the risk of accidentally leaving someone out is high. Since little in life is without risk, here goes: I wish to thank for comments and discussion Paul Dragos Aligica, J. David Alvis, Sameer Bajaj, Ben Bryan, Mark Budolfson, Allen Buchanan, Ronna Burger, Bryan Chambliss, Phoebe Chan, Edwin Curley, Fred D'Agostino, David Enoch, Kelly Gaus, Amanda Greene, Paul Gunn, Keith Hankins, Liz Harman, Cate Johnson, Rachana Kamtekar, Bill Kline, Brian Kogelmann, Andrew Lister, Loren Lomasky, Eric Mack, Meica Magnani, JP Messina, Ryan Muldoon, Jim Otteson, Carmen Pavel, Guido Pincione, Jon Quong, Kristin Reglitz, Jeremy Reid, Massimo Renzo, Tristan Rogers, David Schmidtz, Lucy Schwarz, Gordon Shannon, A. K. Shauku, Hillel Steiner, Stephen G. W. Stich, Virgil Storr, Bob Talisse, Bob Taylor, John Thrasher, John Tomasi, Peter Vanderschraaf, Justin Weinberg, Leif Wenar, and Bart Wilson.

I owe special thanks to a number of people. Bill Glod has provided me with feedback on

this and related projects, as well as both career advice and direction to scholarly opportunities. R.J. Leland has been a great friend since we were in the U.W. Milwaukee M.A. program together, and he is also my go-to person when I need to know what how a more “orthodox” Rawlsian would object to my work. I have learned much from discussing with him both his and my own work. I have been luck to share much time with Dan Shahar discussing ideas from their roughest to most polished state. I have benefitted from working with Tom Christiano, both regarding this dissertation and in his seminars. My views on responsibility have developed significantly in dialogue with Michael McKenna’s views, and conversation with him is always as insightful as it is lively. Dave Shoemaker, through both his publications and discussion, illuminates the complex and multifaceted nature of moral responsibility practices. Steve Wall provided important feedback and guidance for much of my work, particularly my “Markets, Community, and Pluralism.”

Dave Schmitz wrote that we come to deserve opportunities “by not wasting them—by giving them their due, as it were.” I have had tremendous opportunities thanks to Dave; giving them their due continues to be a tremendous, but uplifting, challenge. Though I entered the Ph.D. program the same year Kevin Vallier was graduating, he did more than anyone else to welcome me into the graduate community and help me see what being at Arizona is all about. Besides being a warm and welcoming person, Kevin is an insightful critic who has given me feedback on many drafts of this work. I have learned much from Kevin, particularly in our frequent exchanges about our disagreements about what may seem to be minor details, but which help cast light on the broader system of values and ideas we share.

I cannot say enough to thank Jerry Gaus. He has been extremely generous with his time and insight, reading and extensively commenting on many drafts of this and related material, as

well as discussing the ideas as they developed from their initial rough states to the present. Jerry's comments are always thorough, addressing the content, form and style of the work in ways that helped me not only in developing my ideas but also developing myself as a scholar. It has been a great honor to work with Jerry, and he has been a wonderful mentor and friend.

Lastly, I thank my parents for all of their encouragement and support. I dedicate this work to them.

## Table of Contents

<b>Abstract.....</b>	<b>12</b>
<b>Justification, Coercion, and the Place of Public Reason.....</b>	<b>14</b>
I.    Justification and the Dissenter.....	15
II.   The Limits of Idealization.....	19
III.  Let’s Be Reasonable.....	22
IV.   Relocating Public Reason.....	30
<b>Once More to the Limits of Evil.....</b>	<b>41</b>
I.    Addressing Evil.....	44
II.   Silencing the Reactive Attitudes.....	47
II.i.  To Whom It May Concern.....	48
II.ii. Demanding What Should Have Been.....	50
II.iii. Refusing to Listen.....	52
II.iv.  Necromancy.....	53
II.v.   Expressive Epiphenomena.....	54
III.  Presupposition and Resentment.....	58
III.i.  Representational Content.....	59
III.ii. Demands and Understanding.....	62
IV.   Return to the Limits of Evil.....	66
V.    Conclusion.....	72
<b>Rawlsian Functionalism and the Problem of Coordination:</b>	
<b>Rethinking the Priority of Principles to Institutions.....</b>	<b>74</b>
I.    Rawls’s Functionalism.....	75
II.   Coordination as a Function-based Constraint.....	79
III.  Institutionally Mediated Coordination.....	87
IV.   A Conception for the Basic Structure.....	95
V.    What is to be Done?.....	103
<b>Markets, Community, and Pluralism.....</b>	<b>107</b>
I.    The Community Principle.....	108
II.   Pluralism.....	111
III.  Repugnant Motives.....	115
<b>References.....</b>	<b>118</b>

## List of Figures

Figure 1. Asymmetric Payoff, Three Equilibria (\*) Stag Game (ordinal utilities).....p. 84

## **Abstract**

A common conceptual framework depicts morality as an alien force commanding us from on high; in contrast, this dissertation presents a picture of morality that is deeply social. It is not an abstract morality that commands us, but we who place demands on each other. On this picture, we are equal participants in morality, rather than mere subjects of morality. This participation has fundamentally important implications for the shape and structure of morality; or so this dissertation argues. By way of introducing the work as a whole, I will here note some of the key facets of the social nature of morality that the dissertation develops.

Our participation is primarily as enforcers, rather than followers, of morality. We hold people accountable to moral requirements through emotional responses like resentment, as well as actions and relations that follow from that attitude. As I argue, these emotions carry an important representational content, displaying the other person as having shown ill will. This ill will can be best understood as a disregard for relevant moral considerations that are available to the resented agent. Despite the negative tone of resentment, it is an aspect of being in community with each other. Someone who can be resented is a co-member of a community with us upon whom we can make demands and who can make demands upon us. We may not share community with some people regarding some issues, such as across religious divides, while still seeing them as people with whom we share at least some form of community, as within the system of basic liberal rights. There are people, as I discuss, who fail to be eligible for responsibility to even basic demands. With such people we have no community; they are to us like forces of nature, and the most dangerous of them are for us monsters.

Though many endorse conceptions of community focused on shared experiences or values, I argue that such a notion of community is not appropriate for modern, diverse societies. In modern, particularly liberal, societies, we cannot expect to share religion, occupation, views of the good life, or the like, so these cannot constitute community among the members of society. A shared moral framework, however, provides a promising conception of community for diverse societies like our own. Our shared morality may thus be among the most important forms of community we can have on the large scale of modern society. That same diversity, however, raises problems for a shared morality. As I argue, our interpersonal moral demands will have to be justified to each other, given our different perspectives, and such justification may be

difficult. I address both the nature of this interpersonal justification, as well as the difficulties of achieving it, within this dissertation.

This dissertation shows that morality is social in yet another way. Focusing on justice, as a central part of the morality, I argue that the content of the principles to which we hold each other accountable itself emerges from our social institutions as those develop over time through our interactions. The diverse members of society must be able to share an understanding of their mutual expectations, but such members tend to disagree about how to interpret and apply moral values and principles. Social institutions, such as legal systems with courts to interpret law, can provide a common interpretation of expectations. If the rules that emerge from these institutions are justified to the members, then those rules may constitute justice within that society.

This dissertation, then, presents a picture of morality that is social through and through. Morality is constructed within our social institutions, enforced interpersonally, restricted to what is mutually justified to society's members, and ultimately constitutes one of our primary forms of community.

## Justification, Coercion, and the Place of Public Reason<sup>1</sup>

“The answer is.... You shoot him.” – Burton Dreben (2003, p. 329)

Critics barrage public reason liberals with objections based on the threat of evil and illiberal dissenters. Burton Dreben (2003, p. 329) reported that people often asked him, as a public reason liberal, “What do you say to an Adolf Hitler?” Public reason accounts claim that political institutions must be justified by appeal to reasons accessible to the public – all citizens.<sup>2</sup> Someone like Hitler, however, rejects liberalism’s foundational values, so liberal requirements cannot pass a public reason test that includes Hitler. Since we know people like Hitler exist, the objection seems decisive: either public reason liberals fail to show that liberal institutions are legitimate, or else whatever legitimacy they possess does not extend to the terrible dissenter.<sup>3</sup> Yet it seems that public reason liberals must explain the legitimacy of liberal institutions over terrible dissenters, for as Joseph Raz (1998, p. 40) argues, principled illiberals, “such as Nazis and religious extremists, are an important part of the reason why the rest of us think that governments have authority. They are the people we need a government to control.”

The illiberal dissenter objection is serious because it strikes at the conception of justification at the heart of the public reason project.<sup>4</sup> This article argues that answering the

---

<sup>1</sup> A version of this paper is published as Van Schoelandt (2015).

<sup>2</sup> Paradigmatic public reason accounts are Gaus (1996, 2011a) and Rawls (2005).

<sup>3</sup> Cf. Gaus (1990, p. 336, 1996, p. 160).

<sup>4</sup> For objections of this form, see Abbey (2007), Mack (2011), Okin (2005), Quong (2012, 2014), Raz (1998), Sleat (2013), Taylor (2011, Chapter 8). Gaus (2014a) provides a recent reply.

objection requires fundamentally rethinking the public reason project. After highlighting the distinctive interpersonal conception of justification in public reason accounts, I will detail a core dissenter-based objection to public reason based on a worrisome example advanced by Jonathan Quong. The two prominent strategies for dealing with dissenters involve *idealization of reasoning* and *requiring liberal values*; I show that these strategies do not succeed in a way compatible with the public reason project. That is, the prominent strategies leave public reason theorists with a dilemma between denying the legitimacy of using coercion to protect core freedoms against deeply illiberal people or abandoning the fundamental public reason project. I conclude by proposing a different answer to public reason liberalism's fundamental question: what requires justification?

### **I. Justification and the Dissenter**

The orthodox public reason answer to the fundamental question is that laws backed by coercive political power require justification in accordance with public reason.<sup>5</sup> Although details differ, all public reason liberals committed to a mode of justification internal to, or starting from, the beliefs and values of citizens. Because the public reason liberal also holds that these beliefs and values differ among individuals, for one person to justify a principle or institution requires that she engage the perspectives of others. Such justification is directed to others; I shall call it *interpersonal justification*.<sup>6</sup> This is to be contrasted with *impersonal justification*, which focuses

---

<sup>5</sup> E.g., Eberle (2002, p. 54), Estlund (2008, p. 41), Larmore (2013, p. 293), Macedo (2000, pp. 17, 35), Quong (2010, pp. 233–50), Rawls (2001a, pp. 18–9, 2005, pp. 12, 86, 136), Vallier and D'Agostino (1996), and even (Neufeld & Van Schoelandt, 2014).

<sup>6</sup> Interpersonal justification may be called “justificatory internalism,” as by Gaus (1996, Chapter 3), which is related to reasons internalism (à la Bernard Williams) in regarding the agent's psychological states, but distinct

on reasoning directly about the reasons there are, rather than the beliefs, values and commitments that are somehow implicit in the diverse views of actual people.<sup>7</sup> Though impersonal justification is not focused on the self-exploration of the individual theorist's own commitments – the theorist is reasoning about what reasons there are -- such reasoning must always be undertaken from the first-personal perspective.<sup>8</sup> It is me, the theorist, reasoning about what reasons there are. Steven Wall (2010, p. 136ff), a defender of such impersonalist justification, argues that public reason accounts fail to adequately address this crucial first-personal aspect of moral deliberations. To employ an example of Wall's: when Jill is seeking to decide whether or not euthenasia is morally wrong, her deliberations are about what *she*, from her perspective, should conclude about the reasons that there are regarding euthenasia. According to Wall (2010, p. 139), she should not be swayed in her belief, or her willingness to coercively enforce her belief, should she discover that Jack “cannot rationally accept it.” Impersonal justification appeals to the true or genuine reasons that apply to the subjects (or what is thought true). There is no direct concern for what the subject's own actual views are, particularly as real people may be confused or corrupt.<sup>9</sup>

In contrast, interpersonal justification is justification *to* someone. According to Quong,

---

from Quong's (2010, pp. 5–7, and ch. 5) conception of justification internal to liberalism.

<sup>7</sup> Steven Wall (2002, p. 386) refers to this as “correctness-based justification.”

<sup>8</sup> David Enoch, in defending impersonalist justification against public reason theorists, emphasizes this point, though he admits that it is “hard to distinguish, from the first-person perspective, between *that-p* and *that-I-believe-p*.” According to Enoch (2013), when you coerce based on impersonalist justification, “[y]our reason for action... is *that-p*, not *that-you-believe-that-p*.” Against this view, see Locke (2010, para. 1), Hobbes (1994, Chapter V.3), Rawls (2005, p. 61), and Gaus and Van Schoelandt (forthcoming)

<sup>9</sup> To be sure, externalists might discuss collective deliberation as a discovery procedure to help in arriving at the independently right answer. In addition, externalists are apt to pragmatically consider the actual views of people.

“whether or not some principle can be justified to a person, Peter, depends on what else Peter is justified in believing.” (2010, p. 141; cf. Rawls, 1999a, p. 508) This leaves the possibility that some proposition, A, “might well be true, but the grounds of A's truth are not accessible to Peter...” and so cannot be justified *to him* (2010, p. 141).<sup>10</sup> Though Quong only mentions the relevance of beliefs, justifying to someone also depends upon her values, or what Bernard Williams (1981, p. 105) calls her subjective motivational set, constituted by desires, as well as “dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects... embodying commitments of the agent.”<sup>11</sup> The agent’s values help ground what reasons she has in this internal sense. For interpersonal justification, we say that an agent has a reason to endorse a law when such endorsement relevantly follows from her beliefs and values. An institution cannot be justified to someone if the justification depends upon values she finds alien. Suppose Abraham tries to justify legally requiring traditional dress, citing that modesty requires compliance with such a code. Elvira may agree that modesty favors the law, yet as someone who does not value modesty the imposition is not justified to her. As she sees it, the law would be imposing Abraham’s values on her. To show that a law is justified to an agent, we must show that her own beliefs and values should lead her to endorse the law.

Impersonalist critics of public reason argue that absurd results follow from a requirement that all citizens can access considerations to endorse our laws. These critics insist that it is not

---

<sup>10</sup> Quong (2010, p. 142) elaborates: “The mere fact that there exists a valid justification based on true premises for some proposition Q does not mean that we have justified Q to Peter,” for the latter justification depends “on Peter's wider epistemic situation.”

<sup>11</sup> On the relevance of each subject’s point of view and interests, see Nagel (1991, p. 35). Rawls (2005, p. 390), tellingly, refers to a subject’s “deepest religious and philosophical commitments,” rather than beliefs.

required that citizens be able to reason themselves into endorsing the law, but only that there *be* sufficient reason to do so, or that the law is justified by the truth or rightness of the values supporting it. Though we can discuss the problem of dissenters abstractly, an example will help. Consider Quong's case of Carl who, in terms of interpersonal justification, has sufficient reason to believe he is obligated to kill infidels (2014).<sup>12</sup> Carl does not have isolated murderous commitments, but instead “adheres to a religiously-inspired form of fascism which denies that infidels are free and equal, and maybe Carl is justified in adhering to this doctrine given his unusually intolerant upbringing and social environment, and the large epistemic costs involved in reconsidering many of his most strongly held beliefs.” (Quong, 2014; cf. Taylor, 2011, p. 249) These beliefs and deeply held religious commitments imply that he lacks access to sufficient reasons to endorse even basic liberal laws, like restrictions on his religious murders.

Public reason theorists have a serious problem if they conclude that true liberals cannot prevent Carl from killing infidels. Carl lacks sufficient reason to endorse the laws that would stop him. It thus seems, critics charge, that public reason requires allowing his murder and mayhem (Mack, 2011; Raz, 1998). Surely that would be a *reductio ad absurdum* of the public reason project. After all, impersonalists like Raz find it easy to explain the legitimacy of stopping Carl – simply appeal to the reasons that, in fact, apply to Carl, whether he knows it or not.<sup>13</sup> The relevant reasons may concern flourishing, but they may also concern facts about justice, Lockean rights, or whatever the real reasons may be. Before declaring the battle over, we must first consider two important replies from the public reason theorists.

---

<sup>12</sup> Cf. Quong's (2012, p. 55) related, though less pernicious, case of Anna.

<sup>13</sup> Raz (1999, p. 17) writes: “reasons are used to guide behaviour, and people are to be guided by what is the case, not by what they believe to be the case.”

## II. The Limits of Idealization

Public reason liberals focus on what citizens have reason to endorse, not what they in fact endorse. Sometimes agents fail to actually endorse what they have reason to either because of errors in reasoning, biases or poor information. Interpersonal justification does not require actual endorsement, but merely that such endorsement follows from the beliefs and commitments of the subject (Eberle, 2002, pp. 198–201; Gaus, 1996, pp. 130–1). This allows public reason liberals to idealize the subject, or model what she would endorse after reasoning well about the matter. What an idealized version of the agent recognizes establishes what reasons she has. Consequently, liberal protections are justified to Carl if his suitably idealized counterpart would endorse them. Suppose Carl was not thinking clearly about the protections the law provides him, or perhaps closer consideration of the purportedly infidel-killing passage would lead him to a different interpretation. We should therefore consider idealization’s potential to bring Carl into the liberal fold.

An agent can be either moderately or radically idealized. I will begin by considering the moderate idealization to which Gerald Gaus appeals in addressing some illiberal dissenters. He defends a form of moderate idealization in which we should consider the reasons the agent would recognize after a “respectable amount” of reasoning. Such reasoning “must be accessible” to the agent in question, qua a rational agent living in a world “in which cognitive activity has significant costs.” (Gaus, 2011, p. 253) In this case, rationality does not require the agent to “keep on with the quest to discover less and less accessible reasons.” (Gaus, 2011, p. 253) Interpersonal justification, on Gaus’s view, is achieved when agents “have accessible undefeated reasons to affirm” the law in question (2011a, p. 255).

Idealization should help answer some dissenters, as we can see in Gaus's (1990, pp. 292–

3) application of idealization to Nazis. Nazis made grossly inconsistent attributions to Jews, holding them morally accountable for supposed crimes against Germany while denouncing them as sub-human. They also justified their racist beliefs by appeal to manifestly erroneous science. Once we moderately idealize their reasoning to see what they have reason to endorse, we may conclude that they in fact have reason to endorse liberal protections for all people, including the people they want to kill when not idealized. Even Nazis may recognize on reflection that being open to moral condemnation requires a human standing, so the idealized members may recognize that Jews cannot be treated as subhuman, and from here that their basic rights must be respected. Yet for this response to succeed, Nazis must be rationally unable to render their illiberal commitments consistent in light of the new considerations. It is implausible that the moderate idealization favored by Gaus will always succeed in showing that each Nazi has reason to endorse liberalism. Moderate additions of better scientific data or further moral reasoning may result in minor changes essentially conserving the original view. The full range of corrections necessary to bring someone from a complex ideology like Nazism to liberalism likely requires idealization beyond that allowed by the moderate view.

To consider the more general case, while moderately idealizing may help with some dissenters, we cannot be confident that such idealization will neutralize every troubling case. Dissenters often have a variety of intellectual resources available to interpret new information compatible with their views, and may correct many errors of reasoning in ways that preserve reprehensible commitments, say by giving up the various beliefs that conflict with those commitments, and the illiberal conclusions may be over-determined, arising from diverse elements of their views. Moreover, much idealizing will still leave the agent under burdens of

judgment, weighing diverse values and interpreting in light of their complete life experience.<sup>14</sup> The dissenter can find many ways to respond to challenges to his atrocious commitments, even if it involves a few theoretic epicycles, or even a few contradictions.<sup>15</sup> At this point, the obvious thing to do is idealize further.

Radical idealization, according to Gaus (2011a, p. 236; note excluded), considers what the agent would endorse as a fully rational agent that “follows impeccable epistemic norms... changes his beliefs by making all the inferences from his current set of fully affirmed beliefs, and who employs the fullest possible information set.”<sup>16</sup> Since we are justifying *to* Carl, we begin with his actual beliefs and values, and then consider how those would change through reasoning and investigation. The difference from moderate idealization is that radical idealization goes beyond a “respectable amount of reasoning,” even beyond the reasoning any human can be expected to carry out. This reasoning process could, eventually, lead Carl to fundamentally change his world-view. A fully rational Carl may not only have a different interpretation of his religion, but an entirely different religious view, perhaps convinced by an exhaustive consideration of historic evidence and philosophic arguments about the existence and nature of God. It is in no way obvious, however, why we should be confident that fully rational Carl would settle on liberal views. As confident as we liberals may be in our liberal conclusions, we cannot know from our very non-ideal reasoning not only that those conclusions would remain at the end of our fully rational deliberation, but also that they would be the result of such deliberation from

---

<sup>14</sup> On the burdens of judgment, see Rawls (2005, p. 56).

<sup>15</sup> Agents may also maintain their commitments in light of objections they cannot answer if they have reason to believe that relevant authorities in their tradition could answer the objection.

<sup>16</sup> Gaus rejects such idealizations as ultimately incoherent.

any starting set of commitments. Where one ends in deliberations is likely influenced by where one begins (in Bayesian terms, by one's priors), as well as influenced by the order in which diverse considerations and information are processed (Gaus, 2011a, pp. 239–43).

To be sure, there is no demonstration that radical idealization would not liberalize everyone. Perhaps a Kantian transcendental deduction would show that everyone is irrevocably committed to liberalism merely by seeing themselves as agents or some similar argument (e.g., Gewirth, 1980, Chapters 2, 4). Without such a proof on hand, however, we should remain skeptical that radical idealization would vindicate liberalism for the dissenters, or necessarily any substantive political conclusion. In any event, public reason theorists should reject radical idealization. Such idealization would produce, unsurprisingly, a perspective diverging radically from the agent's actual perspective, perhaps rejecting her actual deepest commitments.<sup>17</sup> Justifications from this perspective would be alien to the actual agent, and it is strange to claim a law is justified *to* an agent when the justification appeals not only to evidence and values she lacks, but to reasoning she, with her human cognitive limitations, could never complete.<sup>18</sup>

### **III. Let's Be Reasonable**

Quong (2014) believes moderate idealization accounts, such as Gaus's, fail because their premises are too normatively sparse or modest to yield robust normative principles showing the error of the dissenter. He thus insists that justification must begin with a commitment to certain values ensuring that it will produce sufficiently liberal principles. Quong's public reason account

---

<sup>17</sup> This starkly contrasts with Rawls (2005, p. 390) aim for justifying to citizens without "criticizing their deepest religious and philosophic commitments."

<sup>18</sup> Weithman connects "intellectually alien" justification to subjection (2002, pp. 201–2). Vallier (2014, Chapter 5) and Wolterstorff (2007, p. 152f.) raise related worries in terms of violation of integrity.

thus limits the justificatory public to “reasonable” people, *defined* as people both accepting and giving deliberative priority to the liberal values of freedom, equality, and fairness (2010, p. 291, 2012, p. 54). Call these the “liberal values.” Only people endorsing, with deliberative priority, liberal values are owed a justification, and it is not hard to justify liberal laws to them. Quong thus avoids the problem of illiberal dissenters undermining the legitimacy of liberal laws, for Quong does not think that justification is owed to such people. Or, more precisely, justification does not need to be provided to unreasonable dissenters based on their own beliefs and values, but can instead be given in terms of the liberal values even if the dissenter neither endorses those values nor would endorse them after idealized reasoning (Quong, 2010, p. 313).<sup>19</sup>

It is essential to Quong’s account that all reasonable people not only endorse the liberal values, but also give those values deliberative priority over all other values. The contrast here is with what Quong calls an “All Things Considered Reasoner”: someone endorsing the liberal values, but weighing them against other values. Such reasoners may accept a *pro tanto* argument from the liberal values to liberal principles, but think those principles can be defeated in light of their comprehensive values.<sup>20</sup> As Quong argues, inclusion of All Things Considered Reasoners makes a public reason account vulnerable to illiberal dissenters. To see this, consider Carl\* who accepts the liberal values, but balances them against his religious values. On balance, Carl\* still

---

<sup>19</sup> Quong’s argument shows why Rawls’s own account cannot answer the dissenter-based objection, and Rawls never explicitly addresses this form of argument, so I do not discuss Rawls’s own account in detail.

<sup>20</sup> Rawls (2005, p. 386) holds that the liberal values ground a “freestanding argument” that is only “*pro tanto*, it may be overridden by citizens’ comprehensive doctrines once all values are tallied up.” This overriding is possible because, on Rawls’s (2005, p. 386) account, “it is left to each citizen, individually or in association with others, to say how the claims of political justice are to be ordered, or weighed, against nonpolitical values.” Vallier (forthcoming, sec. V) also defends permitting such balancing against the public values.

thinks he should kill infidels, or otherwise violate basic liberal protections.<sup>21</sup> Put generally, we cannot include in the justificatory public those who consider their other values in the justificatory public and “also guarantee that the content of the theory will remain suitably liberal.” (Quong, 2012, p. 55; emphasis removed)<sup>22</sup> Quong (2005, p. 304, 2010, p. 37 cf. 146) excludes such people, for inclusion makes “justice hostage to... unjust views” – or, put another way, hostage to the views of the “many ignorant, immoral, self-obsessed, or otherwise troublesome people...”<sup>23</sup> Real people, after all, “might hold *mistaken beliefs about... normative matters.*” (Quong, 2010, p. 146) Only by restricting justification to those giving priority to the public values can one guarantee justification for the sorts of liberal laws those values support.<sup>24</sup> Since an appeal to reasonableness can thus answer dissenter-based objections only if the public values are given priority, Quong requires that priority (2012, p. 52).

Without the need to vindicate laws in the light of diverse commitments, justification for Quong not only begins with the liberal values, but ends there too. Notice, however, that doing so effectively adopts an impersonalist position *vis-à-vis* those deemed unreasonable. Let me

---

<sup>21</sup> Quong’s gives the example of Anna, who balances the public values against religious concerns, and on balance she rejects laws prohibiting employment discrimination (2010, p. 169, 2012, p. 55; cf. related cases at 58). Okin (2005, p. 242) argues against including in the justificatory public members of faiths that exclude women from the priesthood, undermining the opportunities and self-respect of women. Quong (2005, pp. 312–3), however, writes that it would be reasonable to *permit or prohibit* gender discrimination in the Catholic priesthood.

<sup>22</sup> Gaus (2003, p. 138) raises doubts about tailoring a view of justification to secure particular outcomes.

<sup>23</sup> Quong (2010, pp. 162, 166, 190, 216 and 233) makes the same claim in terms “unjust” people, sometimes with accomplices.

<sup>24</sup> The public values must be seen as wholly sufficient for reasoning about political justice. If reasonable people disagree about what *other* values are relevant, they would still require Rawls's overlapping consensus test.

explain. Quong (2012, p. 53) makes it clear that the liberal values are “not themselves subject to any test of public [or interpersonal] justification, nor do we check to see if these values can be the subject of an overlapping consensus amongst *real* citizens here and now.”<sup>25</sup> So Quong is not, for instance, appealing to these values as justificatory for us because these values are implicit in our liberal-democratic culture or the like.<sup>26</sup> Since the liberal values provide the basis for justification independent of the views of other people, they must be discovered through some form of first-personal reasoning, as the “correct” values.<sup>27</sup> That is, they are discovered and verified like any other impersonal justification. We must coerce dissenters in accordance with our own commitments, according to Quong (2010, p. 314), “because to do otherwise would be to ignore what we... believe political justice demands.” Coercion, then, is justified when it is in accord with the proper values (correct beliefs about “normative matters”), like on any impersonalist account, and not justified to the subjects as in a public reason account.<sup>28</sup>

It may seem that this must be wrong. Quong (2010, p. 313) appears committed to

---

<sup>25</sup> Quong's test begins “with certain fairly substantive commitments” instead of “a commitment to public justification...” (2012, p. 56).

<sup>26</sup> Rawls grounds the values in democratic culture (2001a, p. 15). Taylor criticizes this view for the implication that liberalism is only contingently justified and may not be justified at all for those in undemocratic cultures (2011, pp. 238, 298).

<sup>27</sup> Quong (2010, p. 2) claims we “correctly think of ourselves as free and equal from the moral point of view.” The first-personal nature of the reasoning can also be seen in Quong's discussion of reflective equilibrium (2010, pp. 155–6). In related vein, Blain Neufeld argues that political liberalism's foundation principle “should be affirmed as the 'true' or 'correct' one...” (2005, p. 287).

<sup>28</sup> Quong (2010, p. 313) claims state action is justified, regardless of subjects' views, purely “by appealing [to] the values of freedom, equality, or fairness.”

interpersonal justification and justification *to* dissenters when he writes: “*All* persons are offered a justification for the exercise of political power that they could endorse *in their capacity* as free and equal citizens committed to the idea of society as a fair system of social cooperation.”<sup>29</sup> The proper interpretation of this critical claim is not clear, however. It is difficult to see how Quong can be interpreted as claiming that the justification to “all persons” could mean all actual persons, for not all actual persons are in fact committed to that liberal idea. And, in any event, he stresses that “the justification of liberal principles at no point depends on the beliefs of real people.” (Quong, 2010, p. 144, cf. p. 149 and, 2012, p. 53) The best interpretation, I believe, is that justification is addressed to all persons committed to the liberal idea in the sense that justification is provided to each *insofar as she fills the role of citizen* – a role partially constituted by liberal commitments.<sup>30</sup> In seeing oneself as a free and equal citizen, one sees oneself as committed to fair cooperation. This interpretation, though, still leaves us with two problems. First, though *some* justification is indeed provided to each via appeal to her role as citizen, the justification may not be decisive in light of her full variety of roles and commitments. Second, appeal to the commitments of citizenship only provide justification to a subject actually accepting or taking up that role (at least under moderate idealization). Given their interpersonal understanding of justification, only if Carl *sees himself* as a liberal citizen can liberals appeal to the intrinsic commitments of citizenship when justifying to him. Even if there are impersonal

---

<sup>29</sup> Charles Larmore (1990, p. 352) similarly claims that we can justify to dissenters on the counterfactual supposition that they share certain liberal values. Wall (2010, p. 134) proposes a related move converting any impersonal justification into supposed interpersonal justification by attributing to the subject “[a]wareness and appreciation of the full range of evaluative considerations that apply to the situation at hand.”

<sup>30</sup> I thank Jon Quong and Jerry Gaus for very helpful discussions regarding this point.

reasons for Carl to take up the role, we cannot suppose it is a role that he takes up. Conscripting him into a role would not provide interpersonal justification. This may constitute some form of justification, but it is not justification *to* the subjects in the sense central to public reason liberalism.

Claiming legitimacy to coerce a dissenter like Carl based on values he rejects, and may be unable to reason himself to endorsing, makes sense only if those values justify coercion independent of the views of the subject. For a perfectionist like Raz (1988, p. 53), the relevant justificatory values are the “reasons which apply to” the subjects, whether or not those subjects are in a position to know those reasons.<sup>31</sup> The state is justified to, say, prohibit drugs not because all of the subjects can appreciate reasons, from their own perspectives, to abstain. Instead, the laws are justified by independent facts about the harms of drug use. Alternatively, Lockeans like Eric Mack or Robert Nozick hold that states are only justified in protecting natural rights, including rights of entitlement to property. What Raz and the Lockeans share is the view that justification comes from facts independent of the views of the subjects about those facts, be they facts about the nature of good lives, rights, or whatever. Quong’s disagreement with Raz and other perfectionists comes strictly from Quong’s insistence that some reasons that there are, which might otherwise be relevant to political justification (such as those based on human flourishing) should be set aside.<sup>32</sup> On Quong's view, political justification can only appeal to the

---

<sup>31</sup> Raz (2006, pp. 1025–6) adds a “knowability condition” for authority, but this remains in stark contrast to the access requirement of interpersonal justification.

<sup>32</sup> Perhaps Quong’s political liberalism is distinguished from perfectionism for holding only that the liberal values are relevant for political justification. Note, however, that Lockean and some other externalist accounts also have a purely political scope.

sub-set of the reasons concerning duties of justice (2010, pp. 118–20).<sup>33</sup> These justifying reasons are specifically the *actual* duties of justice, presumably *known* by state officials or other elites, not what the subjects think their duties are or can themselves into thinking, so dissent does not undermine the legitimacy of an imposition (Quong, 2010, pp. 128, 313). All the justificatory work is “done by our views regarding who holds which rights and who is under which duties: the allocation of legitimate authority is simply parasitic on our beliefs about the distribution of rights and duties.” (Quong, 2010, p. 116)<sup>34</sup> In the end, these impositions are based not on what can be interpersonally justified, but instead simply on personal or sectarian commitment.<sup>35</sup>

We are forced to conclude that the apparatus of public reason is superfluous in Quong's account. Though he discusses agreement between reasonable people, the justificatory work is in fact done by the conditions of reasonableness. Recall that Quong holds that only the views of reasonable people count, and reasonable people all, by definition, give deliberative priority to the very same liberal values. When considering what institutions are legitimate, we thus can forgo investigating what can be agreed to and, instead, simply appeal to the required liberal values. On this account, a reasonable person need never verify the results of her personal deliberation from

---

<sup>33</sup> According to Quong (2010, p. 135, cf. pp. 128, 130, 313): “Since we cannot rightfully refuse to do what justice requires, others can justify their legitimate authority over us by demonstrating they are merely enforcing the claims of justice that others have against us.” This distinction from the perfectionist vanishes if, as Wall (1998, p. 12) argues, perfectionist reasons are not additional to justice-based reasons, but instead are part of what determines the duties of justice.

<sup>34</sup> Quong (2010, pp. 115, 110) also holds that the duties, including to promote justice, exist prior to the authority.

<sup>35</sup> Quong (2010, p. 313) claims that “[i]f you believe freedom, equality, and fairness are the fundamental political values from which reasoning about justice must proceed, then you cannot also believe that other people are exempt from the normative requirements implied by these values.”

anyone else's perspective.<sup>36</sup> Given the definition of reasonable people, however, we can simplify further to say that a law is legitimate when justified by the liberal values. Thus, agreement really is superfluous, with the liberal values really providing the justification.<sup>37</sup>

The superfluity of agreement is further demonstrated by Quong's denial that agreement on any values or principles outside the liberal set is politically justificatory. In particular, he argues that policies to promote perfectionist values cannot be justified even if all of the actual reasonable people in a society rationally agree on them. Suppose that, literally, everyone in a society was committed to opposing use of methamphetamine on flourishing grounds and, derivatively, to supporting its prohibition. Quong (2005, p. 310, cf. 2010, p. 150ff) claims that such laws would remain unjustified because laws must remain neutral with regard to conceptions of the good, and seemingly any values outside the liberal values. For Quong (2005, p. 310), justification through actual agreement cannot vindicate without subjecting the results to "to any independent standard of justification."<sup>38</sup> The aim of his account, then, is not to justify *to* people with their own views (including moderately idealized versions of them), but to justify in terms of

---

<sup>36</sup> In contrast, Rawls's original position can involve the reasoning of a single agent, but his overlapping consensus stage necessarily involves multiple people with diverse perspectives.

<sup>37</sup> Quong (2014, sec. III) claims we "have the authority to coerce people like Carl because we can justify our decisions by appeal to fundamental values or ideals that everyone (including Carl) ought to endorse and whose implications they ought to accord deliberative priority. If Carl cannot see the force of these reasons, that is not a defect in our justifications, but rather a defect in Carl." James Bohman and Henry S. Richardson (2009, pp. 257–60) similarly argue that normative constraints dispossess agreement. Readers may wish to compare my argument to Quong's own "spare-wheel objection" to the "external conception" of political liberalism (2010, pp. 146–7).

<sup>38</sup> Quong (2014, sec. II) endorses justification through "substantive" claims. Note that on Gaus's account it is only consent of the moderately idealized agent, and not actual consent, that is justificatory.

the correct reasons, reasons that provide a standard independent of everyone's views. So for Quong, the correct reasons can justify even against universal dissent, and no other reasons justify even with actual consensus. Insofar as the public reason project is a distinctive project at all, such sectarianism must be avoided.<sup>39</sup>

Thinking back to our initial problem, critics charge that public reason objectionably lets misguided, ignorant and wicked people undermine the legitimacy of good laws. Those critics call for abandoning public reason's interpersonal form of justification for impersonal justification, or justification according to the reasons that there are independent of subjects' views. Though Quong claims to defend a public reason account, he abandons interpersonal justification for the impersonalism of critics. As he says, on his view, saying that principles could be endorsed really means "that those principles can be validly constructed from a normative ideal" he specifies (Quong, 2010, p. 144). The general lesson is that requirements of liberal values, or "reasonableness," address dissenters at the cost of introducing impersonal justification, and abandoning the justification definitive of the public reason project.

#### **IV. Relocating Public Reason**

The problem of the dastardly dissenter shows that interpersonal justification cannot be strictly required for permissible coercion. In light of this objection, let us consider the prospects for reforming public reason liberalism. The most obvious response to our problem is to deny that interpersonal justification is necessary for justifying political coercion, though interpersonal

---

<sup>39</sup> Contrast with Quong (2012, p. 58). Though David Estlund (2008, pp. 48–9) also appeals to a restricted, possible public, he defends acceptability to this "qualified" public as a necessary, but not sufficient, condition for legitimacy. The acceptance (or moderately idealized acceptance) of actual "disqualified" citizens may also be required. Cf. discussion of membership in the public in Gaus and Van Schoelandt (forthcoming).

justification may be one of a plurality of justifying considerations for coercive laws. While the fact that a law is not interpersonally justified would count against the law, it might still be justified by the weight of other considerations.<sup>40</sup> The practical importance of public reason would depend upon its relative weight and how it may influence more marginal cases. Alternatively, one could think that successful public justification is never necessary, though publicly pursuing such justification promotes respect, social trust or values.<sup>41</sup> Public reason, on this account, does not determine which laws ought to be implemented and enforced, but instead guides public discourse.

It is clear that both the pluralist and the pure pursuit variations “rescue” public reason only by downgrading its importance. On these accounts, public reasons is relevant to justifying political coercion, determining when political coercion is permissible, but is never necessary. They are hybrid views relying on appeals to impersonal as well as interpersonal justification. Nevertheless, they do provide ways to appeal to public reason when justifying coercion while avoiding the devastating implications of dastardly illiberal dissenters. In thinking through their view, public reason theorists should first reflect on the type of goods that cannot be secured

---

<sup>40</sup> We can understand Raz (1998, p. 51) to be proposing a view like this, for he considers it a cost to force people to act against their consciences; presumably it is easier to justify laws not requiring such contra-conscientious force. Related views are defended by Enoch (2013) and Wall (2013, pp. 488, 498n13). That this pluralism is proposed by their staunchest critics, should lead public reason theorists to think, in the words of Admiral Ackbar, “It’s a trap!”

<sup>41</sup> Such a position is defended by Eberle (2002, pp. 189–91). Note that Eberle’s requirement still regards interpersonal justification, unlike William Galston’s (1991, p. 109) view that “we show respect when we offer [dissenters], as explanations, what we take to be our true and best reasons for acting as we do.” I thank Jerry Gaus for discussion of these views.

without public justification. We now know that public reason is not necessary for coercion to be permissible, and we are unlikely to find public justification necessary for any other permission. I propose, then, that we investigate whether there are goods for which public reason is a constitutively necessary element.

Rawls (1999c, p. 59, 1999e, p. 209) holds that mutually acceptable principles make possible “true community between persons in their common practices; otherwise their relations will appear to them as founded to some degree on force and circumstance.”<sup>42</sup> This is certainly not a claim that it would be *impermissible* to have true community in a society based on any other principles. Instead, true community is not possible outside of interpersonally justified institutions. True community represents an ideal of consensually relating to each other, and those relations require, as a constitutive element, that the terms of the social practice be acceptable to the member. While it is possible, and may well be permissible, to force Carl to comply with our commands, we then found our relations with him on force. In such a practice, Carl is alienated from and subjugated by us, not in community with us.<sup>43</sup>

What sort of community requires public reason?<sup>44</sup> The driving insight of public reason tradition has been the claim that recognizing others’ status as free and equal and the requirement

---

<sup>42</sup> Cf. Paul Weithman’s (2010, p. 44) discussion distinguishing “inherent” from “imposed” stability.

<sup>43</sup> R. J. Leland (2014) and Andrew Lister (2013a, Chapter 5) defend accounts of public reason based on community differing from my own not only in their conceptions of community, but also in that they each understand public reason in terms of restrictions on the reasons for policies. I have benefitted greatly from discussing these issues at length with R. J. and Andrew.

<sup>44</sup> Some conceptions of community must be rejected as incompatible with social diversity (Benn, 1988, Chapter 12; Rawls, 2005, p. 37ff; Van Schoelandt, 2014). For an interesting discussion relevant to liberal community, see Bernard Bosanquet (2001a) -- I think Loren Lomasky for directing me to this underappreciated thinker.

of interpersonal justification: that status implies the requirement. We have already investigated and rejected the traditional view that recognizing another as free and equal gives her a standing that requires that any coercion against her be justified to them. We have seen that public reason cannot supply necessary condition for permissible coercion. But there is a more fundamental matter at stake. In recognizing another as free and equal, we recognize her as a co-member of the moral community. She is someone not only who can issue moral demands to us and to whom we can address moral demands, but also someone we can hold morally responsible for moral failures. It is inconsistent with this general standing to address to her moral commands that she cannot reason her way to endorsing – one is then claiming an authority to command the will of a free and equal other. In contrast, we respect the subjects of the moral demands as free and equal when we only issue demands that they could be expected to endorse from their own perspective. Doing so would hold them only to their own commitments, rather than having to ground the authority of the command on our superiority, including our superior insight into impersonal reasons. In this way, interpersonal justification is essential for our free and equal status and grounds these practices of social authority and responsibility.<sup>45</sup> Thus understood, public reason liberalism presents a structure of social morality in which questions of coercion arise, but coercion is not the direct or primary object of interpersonal justification.<sup>46</sup> The motivating

---

<sup>45</sup> Gaus's account directly regards the authority of moral demands (2011, p. 19).

<sup>46</sup> This is one place that my account importantly diverges from that of Gaus. Though he agrees that interpersonal justification is necessary for authoritative moral demands, Gaus (2011a, sec. 17.3) also maintains that a general principle requiring the public justification of coercive acts and laws can itself be publicly justified within communities composed of what he calls "self-directing" moral persons. I hold that, instead, we have patchworks of entrenched particular rights and norms restricting interference without justification defined in terms other than those of interpersonal justification. Of course, our coercively backed laws are often moralized, requiring

concerns and justificatory apparatus of public reasons are inappropriate as the sole consideration in justifying political coercion, but are inescapable to the demands and patterns of responsibility in a moral community.

We hold our fellow members of the moral community responsible through moral emotions like indignation and resentment.<sup>47</sup> Examining these emotions illuminates further the necessity of public reason. In an insightful passage, Adam Smith considers what we should think about someone who, like Carl, murders out of a sincere, but misguided, sense of religious obligation. According to Smith, “the defense of society requires” legal punishment, no matter the perspective of the culprit. We should, however, punish such crimes “with reluctance, when they evidently proceed from false notions of religious duty.”<sup>48</sup> Beyond, and in significant ways independent of, questions of coercion are questions about how we should *feel* about, and relate to, the culprit. This is not a matter of legal responsibility, but a distinct and separable matter of moral responsibility, which is associated with our moral emotions.<sup>49</sup> Smith’s position is that,

---

interpersonal justification to vindicate their implicit claims to authority. Such justification is not directly required for non-moralized laws. Though Vallier otherwise treats all coercive laws as requiring interpersonal justification, he accepts that this is properly only so for laws claiming moral authority (2014, p. 41 note 61). I thank an anonymous reviewer for pressing me for clarification.

<sup>47</sup> On social-moral practice as a system of demands with inter-personal accountability through the reactive attitudes, see Darwall (2006, pt. I), Gaus (2011a, Chapters 11–12), and Shoemaker (2015). Though Southwood (2010, p. 87) does not commit to the necessity of the moral reactive attitudes, he does hold that community is at least partially constituted by social authority based in internalized norms.

<sup>48</sup> Smith (2009, p. 176).

<sup>49</sup> Shoemaker (2011b) defends separating the conditions for criminal responsibility from moral responsibility. Wall (2010, p. 144) separates issues of coercion, criticism, and blame, and insightfully indicates that seeing that

recognizing the misguided views in this case, we should feel reluctance and regret in relying on force, and must withhold the indignation appropriate toward other criminals. The other's perspective is essential in determining the kind of stance we should take to her, and the kinds of responsibility we should attribute to her.<sup>50</sup>

Emotions have associated cognitive content representing the object of the emotion a certain way.<sup>51</sup> Fear, for instance, represents the feared object as dangerous. Peter Strawson (2003, pp. 75–6) and others argue that “moral reactive attitudes,” like resentment and indignation, are responses to the perceived quality of an agent's will in terms of due regard for others.<sup>52</sup> Concern for the quality of will creates many exculpating conditions, including that the agent is exempt from moral demands because of an incapacity, or the agent's circumstances give her an excuse. So, when Abraham is indignant at Elvira for some perceived offense, his indignation should subside if he learns that she acted in the midst of a psychotic episode, she was

---

relevant impersonal justifications are inaccessible to an agent “should temper both our judgment of the person and how we respond to him.”

<sup>50</sup> Forms of responsibility are importantly distinguished by Gary Watson (1996), and Shoemaker (2011a, 2013b, 2015).

<sup>51</sup> See Gaus (1990, Chapter 2), Hurley and Macnamara (2011), Nussbaum (2001, Chapter 1), Rawls (1999a, sec. 73, 1999j, p. 107), Smith (2009, sec. I.i.3.5–10), Solomon (1973). The cognitive content of emotions may be a non-belief-based way of “seeing as,” as when in the grip of a phobia you do not *believe* that the bunny is dangerous, but you see it as dangerous nonetheless (Calhoun, 2003).

<sup>52</sup> Michael McKenna (2011, p. 59) holds that the quality of an agent's will consists in her regard for other moral agents, as well as a broader range of moral reasons. Shoemaker (2013b, 2015) persuasively argues that we should in fact understand the variety of responsibility responses to regard different qualities of will. Acceptance of a quality of will thesis conflicts with Rawls's (1999a, pp. 421–2, 1999j, pp. 105–6) understanding of the moral emotions as responses to compliance with or violation of moral principles.

non-culpably ignorant of the consequences of her action, or an emergency necessitated her action. Any of these conditions could undercut the appearance that Elvira acted maliciously. Now when Abraham acknowledges that he is appealing to a rule that is not interpersonally justified to Elvira – that is, which he holds that she cannot reason herself into accepting – he has reason to excuse her from certain reactive attitudes. Elvira is not showing objectionable disregard for him when she disobeys the rule, for she cannot (even when moderately idealized) reason her way to endorsing it. That being so, she cannot see how obeying the rule is required by a due concern for Abraham, particularly in light of other values she believes to be at stake. I take this to be Smith’s point about the misguided religious murderer who, in doing his best to fulfill what he conscientiously thought was his religious obligation, showed no ill will. The murderer’s misguided view earns our pity and excuses him from indignation, though our own safety requires that we restrain him.<sup>53</sup>

Given these presuppositions embedded within the reactive attitudes, we cannot rationally sustain them unless our demands are interpersonally justified. Here, again, public reason constitutes part of a good, rather than serving as a condition on the permissibility of a behavior. We can, of course, abandon this way of relating to some people, taking up what Strawson (2003, p. 82ff) called the “objective view” of people like Carl. We can see Carl as a patient, or a

---

<sup>53</sup> A number of theorists conceive the moral emotions specifically as forms of nascent forms of communication. On such accounts, indignation is inappropriate when directed at those relevantly incapable of uptake or understanding. I believe these accounts too can be shown to have an important place for public reason as explicating one of the conditions for the capability to take up the message in an instance of indignation. For impressive discussion of the communication account, see Coleen Macnamara (2013), as well as McKenna (1998, p. 127, 2011, Chapter 4), Shoemaker (2007), and Watson (1987, pp. 264–67).

potential strategic partner and opponent, but not someone we can relate to in moralized terms. Sometimes we must relate to people strategically or through force, but public reason makes possible the mutual recognition of co-members in a moral community.

It may seem that in declaring some outside the moral community we declare an open season to destroy them as dangerous tigers (Benn, 1988, p. 102; Locke, 1681, para. 11, 16). Fortunately, we have justified norms among ourselves restricting treatment of outsiders. Even literal tigers are protected by norms against certain kinds of abuse. These norms regulating treatment of tigers are not justified to, or obligating for, tigers themselves. Nor do the tigers have standing to hold us to the norms. These are our norms, regarding tigers, to which we can hold each other.<sup>54</sup> More extensive norms regulate treatment of fellow humans. We, for instance, should respect the outsider's right to free speech and not use her for medical experiments, even though she is not in a position to demand it of us. The liberal tradition presents wide-ranging reasons for extending the protection of liberal rights in this way, and the robustness of these rights leaves me confident that significant restrictions on harsh treatment are publicly justified.<sup>55</sup> We are also restricted, as individuals or members of smaller groups, by our own private morality and other ideals (Strawson, 1961), and by stable conventions for peaceful coexistence (Moehler,

---

<sup>54</sup> Cf. Darwall (2006, p. 28) and Gaus (1990, pp. 369, 374).

<sup>55</sup> Rawls (1999a, sec. 35) argues that though the intolerant do not have standing to demand to be tolerated, we owe it to our constitution (and I believe ultimately each other) to tolerate even the intolerant. Benn (1988, p. 102) argues that "the category of psychopath is very uncertainly defined," so "the status of moral persons is better safeguarded" by extending rights protection to psychopaths. Peter Vanderschraaf (2011) relatedly argues for extending protections to people unable to productively cooperate. Loren Lomasky (1987, p. 40) defends derivative rights for some through social relations to primary rights holders. Cf. Korsgaard (2004), Gaus (1990, p. 371, 1996, p. 166), Kavka (1986, p. 242), Kukathas (2003, Chapter 4), and Quong (2010, sec. 10.1–2).

2014, sec. 5).

At this point, some people are likely to dismiss public reason as unimportant, thinking they can forgo social-moral authority and simply rule through the barrel of a gun. To do so, however, would come with very high costs. The objective view surrenders not only the possibility of moral community, but also relations infused with morality. Without the support of the moral emotions there cannot be such central human relations as friendship, love, or mutual trust (Gaus, 1990, pp. 290–2; Rawls, 1999a, pp. 427–8, 1999j, pp. 109–12; Shabo, 2012; Shoemaker, 2015, Chapter 3; Strawson, 2003, p. 75). For this reason, Rawls (1999a, p. 427) argues that abandoning the moral emotions would disfigure ourselves. Of course, no matter how basic these relations are, we do not need to extend the possibility of friendship to everyone. But the importance of these relations pushes us to form and maintain a mutually justified moral practice with many people.

There is another reason to not be satisfied with mere coercion and relations within the objective view. Peaceful cooperation and prosperity depend upon moral relations and resultant trust. Consider even cooperation within markets, often thought to be domains of amoral strategic interaction. Well-functioning markets involve extensive exchanges among strangers possible only where honesty, trust, reliability and fair dealing are ubiquitous and supported by the moral emotions (Rose, 2011, Chapter 2; Schwab & Ostrom, 2008; Stringham, 2011). Though cooperation among purely strategic individuals develops in some circumstances, such cooperation is constrained and fragile (Bowles & Gintis, 2011, Chapter 5; Gaus, 2011a, Chapter II.6).<sup>56</sup> Strategic cooperation faces not only losses from opportunist behavior and the costs of

---

<sup>56</sup> Though see the impressive studies by Leeson (2007, 2009) and Skarbek (2010, 2011, 2014) of governance and emergent order among some of the individuals least known for moral virtue.

monitoring and enforcement, but the complete absence of certain opportunism-vulnerable investments and cooperative activities. Without mutual trust, resources and entrepreneurial energy away from productive, and into unproductive or even destructive, uses (Baumol, 1990; Boettke & Coyne, 2003; Hobbes, 1994, Chapter XIII.9). Prosperity can be secured only by an extended moral community, and not through mere force.

Even an effective coercive institution requires far reaching social authority. As Gauthier (1969, p. 168; cf. Schwab & Ostrom, 2008, p. 223) writes, cooperation of enforcers “cannot rest entirely on the fear of some further power. At some point, barring a Kafkaesque regress without limit, some degree of voluntary co-operation on the part of some persons is the condition of the continued working of all social and political bodies.”<sup>57</sup> Effective coercive systems depend upon cooperation and trust supported by moral relations at least among the coercing agents. The individual initially happy to simply coerce without social-moral authority soon finds that she cannot be truly effective in her coercive plans without that authority. We may not need to have everyone within our moral community, but our own ends are often best served within an expansive and inclusive moral community.

There are two ways to develop community where it is lacking. Holding our practice fixed, we can always attempt to win the hearts and minds of dissenters, encouraging them to change their perspectives and convert into community members. If, however, the practice depends on many deeply contentious matters, such as following a particular religion or sectarian political ideology, then the conversion strategy is hopeless. As Rawls (2005, pp. 36–7) argued, pluralism about religious, philosophic and moral issues is the natural result of the free exercise of reason. Converting people to share a single doctrine would thus itself require the oppressive use of force

---

<sup>57</sup> Cf. Bannan et al. (2010, p. 9).

that is itself antithetical to true community. The only hope for expansive community in any modern society is to shape the social so that diverse members have reason to endorse it from their separate perspectives.<sup>58</sup> It is the place of public reason to illuminate the possibilities for social morality in a diverse and bounded world.

---

<sup>58</sup> Rawls (1999h, p. 431) writes that “there is no practicable alternative superior to the stable political unity secured by an overlapping consensus on a reasonable political conception of justice.” Lomasky (1987, pp. 81–2) argues that the broad acceptability of the moral community requires limiting the demands placed on individuals, and particularly entrenching individual liberties. If Lomasky is right, then inclusive communities will have what Gaus (2010) calls a “classical liberal tilt.”

## Once More to the Limits of Evil

“I’m as mad as hell and I’m not going to take this any more!”

Howard Beale, *Network* (1976)

Peter Strawson provides a fundamental insight that we can understand moral responsibility in terms of the emotional responses to others. In particular, emotions constitute one mode through which we hold each other responsible; examining such emotions reveals the corresponding form of responsibility.<sup>59</sup> Central to this approach, both for Strawson and most philosophers following him, is resentment, along with the closely related emotions of indignation and guilt.<sup>60</sup> Through examining these emotions and the conditions of their fittingness, we can discover the conditions of moral responsibility.<sup>61</sup> For instance, Strawson examines the conditions

---

<sup>59</sup> Strawson (2003).

<sup>60</sup> As is typically done in the Strawsonian literature, I will treat resentment, indignation and guilt as differing only in how the experiencer of the emotion relates to the supposed wrong-doer (or person otherwise worthy of being the target of these emotions). That is, while a victim experiences resentment toward a wrong doer, indignation is experienced by a third party (non-victim) toward a wrong-doer, and guilt is experienced by a wrong-doer directed to herself. While I focus on resentment, it may be that the more basic emotion of anger is more apt, as argued by David Shoemaker, (2015, Chapter 3). Cf. Aristotle (2004, bk. II.2) I thank Dave Shoemaker and Ronna Burger for helpful discussion of this point.

<sup>61</sup> More precisely, we learn the conditions for one form of responsibility. A complete account of responsibility would take account of a broader range of emotions, particularly other negative emotions like shame and regret, as well as positive emotions like pride and gratitude. Such a complete account is likely to find that responsibility is multifaceted instead of unified, with diverse means of holding responsible associated with different standards

that would exculpate someone from liability to fitting resentment to clarify what sorts of freedom are or are not needed for responsibility. He argues that we withhold resentment of an apparent wrong-doer when we discover that she was coerced, but we would not withhold resentment of her simply on the discovery that the universe is deterministic. Strawson thus attempts to show that, within our emotion-constituted practice, though responsibility requires some form of freedom, it is not a form of freedom that is affected by determinism.

A number of theorists have taken up this Strawsonian method of understanding responsibility through resentment. This paper takes up two strands in the literature originating with Gary Watson's highly influential "Responsibility and the Limits of Evil." According to Watson's interpretation of Strawson, the reactive attitudes are communicative entities or forms of moral address. More recent versions of this "expressive view" further specify resentment as containing a directed message or a form of address in an ongoing conversation.<sup>62</sup> Were the expressive view correct, it would clarify the way we hold people responsible, showing that we are in a central way attempting to communicate with them. Such an understanding would further illuminate the conditions of responsibility, such as a capacity for understanding that makes our communicative attempts intelligible.

Much as Strawson examined common exculpating conditions in order to clarify the conditions of responsibility, Watson draws attention to a case at "the limits of evil" where we are to consider a cruel, cold-hearted murderer and how we would normally respond to him.<sup>63</sup> Though

---

to which people can be held, or we might say different qualities of will. For an exploration of that sort, see Shoemaker (2013b).

<sup>62</sup> See, for instance, Macnamara (2013, 2015) McKenna (2011); Shoemaker (2007).

<sup>63</sup> Cf. McKenna (2008) and Watson (2008). Outside the limits of evil, a number of diverse "marginal" cases are

interacting with such people is not a daily occurrence, heartless killers are no mere philosopher's fancy; they are all too real. A theory of responsibility should adequately capture our responses to such cases. Watson contends that at least some versions of the expressive view mistakenly find the murderer not responsible, as if evil were its own excuse, and Michael McKenna defends one way for an expressive view to handle the case. I will address the case and their analyses further below.

This paper shall argue that the expressive view is mistaken and should be replaced with an alternative "presupposition view." To make this case, the paper has four parts. The first elaborates the expressive view, particularly how it interprets certain Strawsonian concepts like the basic demand for good will, and considering some of the developments of the view since Watson's seminal work. The second part shows why the expressive view must be rejected. I will argue that expressive views give a confused timeline for holding responsible, because it places the demand for regard too late for the demand to be complied with. Furthermore, I will argue that there are many people who we properly resent despite not being appropriate targets of communicative efforts. The expressive view cannot appropriately capture these cases without abandoning the focus on expression and communication.

Criticism of the expressive view is not merely destructive, for considering those objections will both clear the ground for an alternative account and clarify the form an alternative must take. So, in light of the criticisms discussed in part two, part three develops an alternative presupposition view of the emotions through which we hold people responsible. On this view, the reactive attitudes include a cognitive content representing the resented agent as having had a certain quality of will. This representation, rather than conditions of address,

---

usefully explored by Shoemaker (2007, 2015).

ground the appropriateness of resentment. These appropriateness conditions, because they are not based in communication, avoid the difficulties plaguing the expressive view and more clearly explain what we are doing when we hold people responsible.

With the presupposition view in hand, we can conclude in part four by reconsidering limits of evil. The expressive view points in the wrong direction, raising red herrings like concern about whether the murderer is open to dialogue or refuses to listen to us. A re-examination of the case in terms of the presupposition view will both clarify the view and illustrate how to think through cases of evil that are often more complex than they may first seem.

## **I. Addressing Evil**

Watson reconstructs Strawson's account as an "expressive theory of responsibility" in which the "reactive attitudes are incipiently forms of communication, which make sense only on the assumption that the other can comprehend the message."<sup>64</sup> Many people follow Watson in this regard and understand the reactive attitudes in terms of messages or forms of address.<sup>65</sup> Since theorists differ in the details and develop the view in diverse ways, I will focus on the broad framework of the expressive view and discuss some of its key variants as needed.

---

<sup>64</sup> Watson (2008, pp. 119, 122) Cf. McKenna (McKenna, 2008).

<sup>65</sup> E.g., Macnamara (2013, 2015) and Shoemaker (2007). Though Phillip Pettit understands responsibility via resentment as related to "discursive control" and the agent's standing within a discourse, he does not hold that resentment expresses a demand. Instead, he holds that the discursive features are necessary for an agent "being fit to be held responsible ... not in their actually being held responsible." (2001, p. 26) I thus understand Pettit's account as being a version of the presupposition account I defend below.

Strawson holds that the reactive attitudes “are essentially reactions to the quality of others' wills toward us as manifested in their behaviour: to their good or ill will or indifference or lack of concern.”<sup>66</sup> We resent not those whose actions set back our interests in some way, but those agents that express contempt, disregard, malice – that is, ill-will.<sup>67</sup> To those showing us good will, we respond with gratitude. On the expressive view, reactive attitudes like resentment communicate, or express, the demand for good will. We can paraphrase Strawson to get a statement of this demand:

*The Basic Demand:* the normative expectation of, and demand for, the manifestation of a certain degree of good will or regard, or at least lack of ill-will, toward ourselves and beings of moral significance.<sup>68</sup>

Typical members of a moral community, on this account, mutually address each other with the basic demand and in turn understand the demand from others. According to the expressive view, cases of exculpation from responsibility can be understood in terms of failures of address. Macnamara, for instance, writes that “[s]ince a blamee is an addressee [of resentment understood as a communicative entity], eligibility for the role of blamee requires eligibility for the role of addressee.”<sup>69</sup> Such “eligibility requires that one has the capacities to give uptake to the address in

---

<sup>66</sup> Strawson (2003, p. 83)

<sup>67</sup> Adam Smith (2009, sec. II.iii.I.5) writes: “When our enemy appears to have done us no injury, when we are sensible that he acted quite properly, that, in his situation, we should have done the same thing, and that we deserved from him all the mischief we met with; in that case, if we have the least spark either of candour or justice, we can entertain no sort of resentment.”

<sup>68</sup> Strawson (2008, p. 29).

<sup>69</sup> Macnamara (2015, sec. 2). Some theorists would distinguish “blame” from responsibility, but I take Macnamara here to be using “blame” to mean hold responsible via resentment (or like emotions), with a “blamee” being someone at whom resentment is directed.

question.”<sup>70</sup> So, according to the expressive view, children must be exempt since with their inadequate understanding of moral communication, as Watson puts it, “the reactive attitudes lose their point as forms of moral address.”<sup>71</sup>

Another key aspect of Watson’s article is that he raises a troubling difficulty for the expressive view. On that view, trees, children and psychotics can all be understood to be altogether exempt from responsibility because they are unable to relevantly understand moral address and thus leave no point in trying to so address them. Normally functioning adults, on the other hand, are responsible to each other because they form a sort of community within which attempts at address are mutually understood. The exempt are outsiders of that community. Watson thinks this concern for shared community as a basis of mutual address “suggests shared ends, at some level, or a shared framework for practical reasoning,” or as he also puts it, “a shared framework of values.”<sup>72</sup> Watson, however, objects that such an account has the implausible implication that extreme evil is its own exempting condition.<sup>73</sup>

Consider Watson's example of the murderer Robert Harris.<sup>74</sup> I take the key facts to be that Harris, besides numerous other crimes, shot and killed two teenage boys while stealing their car for a bank robbery. The killing was in cold blood, not at all necessary for getting the car. Harris expressed, instead of remorse, amusement at the incident. His disregard for civility and basic

---

<sup>70</sup> Macnamara (2015, sec. 2).

<sup>71</sup> Watson (2008, p. 123).

<sup>72</sup> Watson (2008, p. 125)

<sup>73</sup> Watson (2008, p. 125) writes: “If holding responsible requires the intelligibility of moral address, and if a condition of such address is that the other be seen as a potential moral interlocutor, then the paradox results that extreme evil disqualifies one for blame.” Put another way, “some forms of evil will be exempting conditions.”

<sup>74</sup> Watson (2008, pp. 125–8).

moral relations made even his fellow Death Row inmates despise him and celebrate his execution. He is certainly a prime candidate for blame. Harris rejects our moral community's shared framework of values, and so, Watson alleges, Harris must fail to be a member of the moral community.<sup>75</sup> Thinking directly about moral address, there is no hope for moral dialogue with someone like Harris – “he is hardly a potential moral interlocutor....”<sup>76</sup>

On the expressive view, however, since Harris is outside the moral community and moral address would be pointless, Harris must be exempt from moral responsibility. This cannot be right, so Watson proposes amending the account of responsibility. Specifically, Watson argues that neither “comembership in the moral community” nor “the significant possibility of dialogue” can be necessary for moral responsibility.<sup>77</sup> Resenting Harris, according to Watson, should not be understood as a robust form of moral address, but instead merely “a denunciation forlorn of the hope of an adequate reply.”<sup>78</sup>

## II. Silencing the Reactive Attitudes

Though many Strawsonians claim the reactive attitudes *express* the basic demand, Strawson indicated two distinct relations between these attitudes and the demand. On the one hand, Strawson says the reactive attitudes “involve, or express,” the basic demand.<sup>79</sup> On the

---

<sup>75</sup> Repudiating the moral community makes Harris a “moral outlaw” (Watson, 2008, p. 128).

<sup>76</sup> Watson, (2008, p. 125).

<sup>77</sup> Watson, (2008, p. 125).

<sup>78</sup> Watson, (2008, p. 125).

<sup>79</sup> Strawson (2008, p. 30).

other, he writes that the reactive attitudes “rest on, or reflect” the basic demand.<sup>80</sup> These statements present different views about the place of demands as they relate to the reactive attitudes, for expression places the demand interior to the emotion while reflection implies that the demand is in some way exterior to and independent of the emotion. This section will argue that we must reject the former characterization, as found in the expressive view. Such criticism will shed light for developing, in the next section, an alternative view that understands the relation of resentment and the basic demand more in terms of reflection.

### **II.i. To Whom It May Concern**

The first problem with the expression view comes from that fact that the reactive attitudes have a definite directionality. Resentment is felt toward some particular individual and we need to know the conditions for resentment to be appropriate or inappropriate toward that very individual. If the expression view is to appropriately ground the conditions for responsibility and exculpation therefrom, the view must explain why the characteristics of the resented agent, and not other agents, are relevant. That is, we must know why the conditions for appropriate resentment concern the ability of the resented agent to understand the address, rather than the ability of other agents to understand the address. Macnamara tries to address this problem by arguing that resentment has a function such that it must carry a directed message.<sup>81</sup> That is, on Macnamara’s account, resentment developed specifically for sending a message to, and getting uptake of that message by, resented agents.

The problem is that, as Macnamara herself recognizes, the reactive attitudes may “have

---

<sup>80</sup> Strawson (2008, p. 29).

<sup>81</sup> Macnamara (2015, sec. 3).

many functions[,]” much as the “tongue is both for eating and talking...”<sup>82</sup> When an item has multiple functions, then the mere fact that one function is not being fulfilled in some use does not seem to undermine the legitimacy of that use. So, even if we grant that one function of resentment is to gain a corresponding uptake by the resented, it may well have other functions. Resentment may also serve to communicate to other members of the community, or act as a means of self-affirmation for the resenting agent, or even serve non-communicative functions like motivating retribution.<sup>83</sup> There could then be cases for which resentment is appropriate, in that the resentment fulfills one of its other functions, even when the resented agent cannot have uptake.

Unless it can be shown that all functions of resentment require uptake by the resented agent, then the conditions of appropriate resentment may not adequately depend upon fact about the resented agent. This problem is particularly pressing since Strawsonian accounts understand an agent to be responsible when it is appropriate to hold the agent responsible. If resentment serves a communicative function with regard to agents other than the resented, then it can turn out that some agents are responsible because of the potential that other agents have for uptake. Even many paradigmatically exempted agents may be found responsible because others in the community can understand the message. For instance, if other community members understand the message when I resent the behavior of an infant, or if I can understand the message of my

---

<sup>82</sup> Macnamara (2015, sec. 3).

<sup>83</sup> On the evolution of anger to motivate retribution, see Boyd, Gintis, Bowles & Richerson (2003); Gerald Gaus (2011a, sec. 7); Robert L. Trivers (1971, p. 49). Communication to oneself and to members of the community besides the wrongdoer is plausible if resentment involves a kind of protest, as moral blame involves in the account of Angela M. Smith (2012).

own resentment of the weather, then the view would seem to declare that the infant and weather are each morally responsible beings. It is thus essential that the expressive view explain why the communication is strictly targeted to the resented agent, and that there be no other functions of resentment, such as communicating to others, if it is to have the conditions of responsibility appropriately connected to features of the responsible agent herself.

I should highlight an implication of the present objection. One may be tempted to replace moral address with some other function of the reactive attitudes, either in terms of some other communication-related function or in terms of some non-communicative function. From this newly identified function, the theorist would then attempt to derive the appropriateness conditions for the reactive attitudes. Such an account will face the same problem of ruling out other functions that are insufficiently related to the capacities of the target. We may, of course, also note that any function-based account, even if we grant that it has properly identified the natural function of a trait, would have to explain the normative relevance of that function. For my own part, I am deeply skeptical of inferences from natural functions to practical requirements, but we should now bring our focus back to the moral address view.

## **II.ii. Demanding What Should Have Been**

The expression view is puzzling when we consider the fact that resentment is, at least typically, backward looking. We resent people for what they did in the past. A demand included in resentment could only be a demand regarding future action, but resentment holds people accountable for past actions. It would be better, as Nelkin suggests, to understand resentment as “a kind of holding to the demands to which the agent was subject at the time of action but did not

fulfill.”<sup>84</sup> Though resentment does not express *that* demand, the resentment in a way depends upon it. The expressive view thus gets the timing of the demand off and misunderstands the relation of resentment to the demand.

It seems there may be current demands implicit in resentment, including demands that relate to the basic demand. For instance, resentment could include the demand that the target empathize, right now, with the person she wronged.<sup>85</sup> Though there could be such a demand implicit in resentment, that is not the basic demand for good will at the time of action. In fact, to the demand for empathy seems to presuppose a separate and prior basic demand insofar as the this demand for empathy focuses on the way the target has wronged the resentful agent. Whatever other demands may be implicit in the resentment, the basic demand must be before an episode of the resentment. We may say that those attitudes presuppose, rather than express, the basic demand.

A closely related problem is that the expressive account, by placing the demand in the episode of resentment, will tend to ask about the capacities for comprehension of the agent at the wrong time. An account of responsibility should track an agent’s capacities at the time of action, not the time of resentment. The expressive account, however, in considering the ability of the

---

<sup>84</sup> Nelkin (2015, n. 10). Nelkin here is concerned about “blaming,” which she seems to intend as inclusive of resentment. She writes that, “insofar as one is truly accountable for a past action, it is because one was *at the time of the action* subject to a demand—a demand to act differently at the time, whether that demand was actually expressed or not.” (2015, sec. 2)

<sup>85</sup> Shoemaker (2007, p. 93) writes: “It looks as if what is being demanded by the emotional aspect of moral address is empathy by the target of the reactive attitudes, imaginatively stepping into the shoes of the other person in order to feel what one has put him or her through.” Likewise, he writes that resentment is an “invitation for the offender to take a moment to appreciate what he has done from within the other’s perspective.” (2007, p. 100)

agent to understand the resentment may find an agent responsible for actions performed prior to the development of that ability. For instance, suppose Alex comes out of a psychotic episode during which he damaged Beverly's car. Now that the psychosis has passed, Alex can understand resentment directed toward him, but the appropriateness of the resentment depends upon the abilities at the time of the action, during the psychotic episode, rather than on his current abilities. We see again, then, that the expressive view places the demand, and the abilities to understand the demand, too late.

### **II.iii. Refusing to Listen**

This subsection will address a problem of "refusal" and the following subsection will introduce a related problem of "necromancy." They are related problems, in that each considers an agent with whom moral dialogue is infeasible and each of these problems has inspired modifications of the expressive view. I will discuss each of these problems before discussing the modifications that avoid these particular problems but raise their own problems in turn.

The problem of refusal is simply that moral address is futile with someone who refuses to listen. Such futility would seem to require the expressive view to exempt from responsibility people who are uninterested in moral dialogue. For instance, Watson describes Harris as "hardly a potential moral interlocutor..." since any "invitation to dialogue would be met with icy silence... or murderous contempt."<sup>86</sup> Attempts to address Harris would certainly be futile, but this does not seem to undermine the appropriateness of our resentment. To be clear, the futility is not merely in the fact that Harris will not change his behavior. Watson writes that the "point" of moral address is not the consequences, but rather "to address the other with a complaint and a

---

<sup>86</sup> Watson, (2008, p. 125).

demand.”<sup>87</sup> It is not clear what point there is, however, in addressing a complaint to someone who will not even hear it, so the expressive view seems to make sufficient obstinacy an exempting condition.

#### **II.iv. Necromancy**

Worse than those who refuse to listen are those who cannot listen. The expressive view must somehow explain how resentment can be appropriate, as it seems to be, even when the target is not available for address because shielded by distance or death. Corpses are particularly inapt targets of demands; they cannot understand the message and are rather poor conversation partners. We cannot address our complaints to the dead, so a straight-forward moral address account will thus find that it is inappropriate to resent the dead. Our common practice, however, seems to include resentment for people even after they are dead. Resentment for the dead cannot effectively address the dead, so the appropriateness of post-mortem resentment suggests that resentment is not best understood in terms of address or expression. Of course, even if one wanted to argue that resentment for the dead is misplaced, the same line of argument also applies to individuals we simply cannot address, such as criminals in hiding, the physically distant, or people who we cannot fully identify (e.g., whoever committed the unsolved murder). These provide a wide range of cases in which resentment seems to remain appropriate despite the impossibility of addressing the target.

---

<sup>87</sup> Watson (2008, p. 123).

## II.v. Expressive Epiphenomena

Watson and McKenna propose modifications of the expressive view that avoid the problems of refusal and necromancy, but these modifications amount to rejecting the expressive view. That is, in order to better capture key cases, Watson and McKenna modify the expressive view in ways that prevent us from using the conditions of appropriate address or effective communication to understand the conditions or nature of responsibility. Examining this shift will point the way to a better understanding of the nature of our reactive attitudes.

In light of the problem of refusal, Watson maintains that resentment should be understood as a communicative entity, but rejects the idea that we specifically need dialogue. On Watson's proposal, we should understand resentment of Harris not as a robust form of moral address, but instead merely "a denunciation forlorn of the hope of an adequate reply."<sup>88</sup> McKenna, on the other hand, proposes that what matters is whether or not the target has the "capacity" to enter the moral dialogue, not whether she actually would choose to do so.<sup>89</sup> It is thus appropriate, on McKenna's account, to address Harris so long as he can understand the demand and the relevant values the demand draws upon, even if Harris in fact rejects those values or obstinately refuses to listen or understand us. Regarding the dead, McKenna proposes that resentment is appropriate when it could play the relevant conversational role under counter-factual conditions including that the target was alive.<sup>90</sup> Though the target may in fact be dead, it may be true that moral address would have been effective were the target alive.

---

<sup>88</sup> Watson (2008, p. 125).

<sup>89</sup> McKenna (2008, p. 142 cf. p. 130).

<sup>90</sup> McKenna (2011, p. 177).

The most basic version of the expressive view tells us that resentment is a form of address and we can understand its appropriateness in terms of the appropriateness of addressing the target, as when we consider whether or not the target can understand the address. Watson and McKenna, however, propose that we move away from these considerations. On Watson's view, we express denunciations "forlorn of the hope of an adequate reply," and the appropriateness of such expressions is not grounded in the availability of the target for address. Watson's proposal, however, leaves us with no clear condition for the appropriateness of resentment in terms of communication at all. We can, after all, express our hopeless denunciations just as well to infants or the weather as to Harris.

McKenna's view seems to even more clearly move away from concern for the conditions of address and communication. On his account, there is an explicit appeal to availability for address, but for many unavailable people McKenna appeals to the fact that they would be available for address under certain counterfactual conditions. The counterfactual considerations show that it is not really conditions of communication that matter. Instead, what matters must be whatever is true about the target that the counterfactuals are chosen to bring out. Working through a pair of examples may help us clarify the concern.

Consider what McKenna's account would say about resentment for Harris now that he has been executed. We know that Harris is not available for address (barring necromancy), and that he would not listen even if he were. McKenna, however, says that resentment will be appropriate if Harris had a capacity to understand our address through understanding our values (whether or not he endorsed those values) were he alive and were we to express our resentment to him. The dead Harris is an appropriate object of resentment, then, when certain things are true of him, such as that he had the relevant understanding. The counterfactual availability for address

is simply picking up on other features of Harris, particularly what he knew and understood at the time of his crimes, so it seems that the facts about address are merely epiphenomenal to the real grounding of responsibility.

Considering a non-responsible agent will further illuminate the case. Suppose that Harriet committed acts much like those of Harris, but she did so after a psychotic break brought on by exposure to a powerful psychoactive chemical for which she is not at all responsible. She was at the time of her violence, and remains still, out of touch with reality, including being unable to understand our attempts to address her. Were she not psychotic, she would be perfectly able to understand our address. This counterfactual availability for address, however, would not lead McKenna to say that we may appropriately resent Harriet. The difference between the cases does not seem to be that Harriet's case is too remote of a possibility, for "were she sane" does not seem to be a more distant possibility than "were he alive." The difference must arise from certain features that are held fixed or adjusted when we consider the counterfactuals. On McKenna's account, we must be asking something like: holding fixed the intentions, beliefs, and understanding of values that the person had at the time of action, but leaving other features open to counterfactual specification, could she possibly understand us? The relevant difference between Harris and Harriet must be found in their intentions, beliefs, understandings, or whatever else is held fixed. Once this is seen, however, it is unclear why we would need to appeal to the counterfactuals at all. We could, instead, merely focus on the relevant features that we hold fixed. Those features are what really matter.

McKenna may hold that consideration of counterfactual address serves an important heuristic purpose and that such consideration is useful for thinking through cases. This would fit particularly well with the fact that he proposes a "model" for holding responsible in terms of

conversation, and models may be valuable without being perfect. The appeal to address, then, may be thought to be informative. We see a very similar strategy deployed with hypothetical contracting in moral and political philosophy. It might be charged that a hypothetical contract is not worth the paper it is not written on, for the force of a contract comes only from actual agreement to it.<sup>91</sup> The moral significance of the hypothetical contract, however, is not said to come from the contracting. Instead, consideration of the hypothetical contract is meant to bring out or clarify what is already true of an agent, as when the fact that the agent would contract to form some institution in question evidences that that institution is beneficial for the agent if complied with by all. The latter fact is the ground of moral significance, and the hypothetical contracting is merely a sign or way of thinking through cases.<sup>92</sup> It seems that McKenna's appeal to hypothetical availability for address is meant to serve a heuristic role in much the same way

---

<sup>91</sup> Responding to John Rawls, Ronald Dworkin (1976, p. 18) writes: "A hypothetical contract is not simply a pale form of an actual contract; it is no contract at all." In similar vein, A. John Simmons (1999, p. 744) says of Robert Nozick's account that, "[s]howing that it is possible for a (certain kind of) state to arise and function without immorality and that having such a state would be a good thing – that the state is justified, on Nozick's model – is obviously not the same thing as showing that a particular actual state (even of that kind) did in fact arise and does in fact function in morally acceptable ways." (Cf. Schmidtz, 1996, p. 92) Some theorists seem to dissent from this view, including Thomas Nagel who holds that to "show that they all have sufficient reason to accept [a system of compulsory institutions] is as close as we can come to making this involuntary condition voluntary." (1991, p. 36)

<sup>92</sup> See, for instance, Dworkin (1976, p. 25); Gaus (2012a); Rawls (1999a, sec. 4). For Lockeans like Simmons (1999, p. 746), a hypothetical contract may show that certain kinds of states are prudential to form or good things to have, while only actual consent confers authority on an actual state. Oliver Sensen (forthcoming) presents an interpretation of Kant according to which the hypothetical contract illuminates outputs of practical reason, but specifically independent of our interests or desires.

that hypothetical consent plays that role for some contract-based theories.

Hypothetical availability for address may well serve that heuristic role, but we must emphasize that the account does not then understand resentment to be a form of address whose appropriateness can be understood in terms of the present appropriateness of address. The comparison to the use of a hypothetical contract is informative here, specifically in that appeals to a hypothetical contract are very different from actual contracting or voluntarist views. If a theory justifies institutions through an actual contract, then the theory is elaborated primarily through considering issues like the normative significance of volition, the limits of self-binding, and the necessary conditions for morally relevant consent. A hypothetical consent account, on the other hand, is not advanced by further consideration of consent, but is instead elaborated through considering the principles or values that ground the description of the hypothetical contracting situation. In like fashion, it seems that a heuristic of hypothetical appropriate address is going to be elaborated through considering the values grounding the counterfactuals, the considerations that make some particular counterfactuals and not others necessary to consider. We would not, however, learn about the conditions for, or nature of, responsibility through focus on the conditions for availability to address, for whether or not the agent is so available is not really the point.

### **III. Presupposition and Resentment**

The critical fires that consumed the expressive view will also illuminate an alternative view we must now consider. Instead of focusing on the conditions of address, the alternative view must find the conditions of responsibility in the responsible agent and locate the basic demand prior to the resentment. This view will take up Strawson's suggestion that the reactive

attitudes rest on or reflect, rather than express, the basic demand. Let us put the alternative view thus:

*Presupposition View:* The reactive attitudes represent their target as having shown insufficient regard and presuppose the basic demand applied to the target at the time of the target's action.

This section will develop aspects of the presupposition view in more detail.

### **III.i. Representational Content**

The presupposition view agrees with the expression view that the reactive attitudes have propositional content, but while the expressive view understands this as the content of a message, the presupposition view understands it as representational. Emotions in general, not just resentment, have associated cognitive content representing the object of the emotion a certain way.<sup>93</sup> An emotion's representational content grounds the appropriateness of episodes of that emotion in terms of either whether the content accurately represents its object or whether the emotional-experiencing agent is justified in believing that such representation is accurate.<sup>94</sup> To consider a common example, fear represents its object as dangerous and can thus be seen to be fitting insofar as the object really is dangerous or the agent has sufficient reason to believe that the

---

<sup>93</sup> See Gaus (1990, Chapter 2); Hurley & Macnamara (2011); Nussbaum (2001, Chapter 1); Rawls (1999a, p. 73, 1999j, p. 107); Adam Smith (2009, sec. I.i.3.5–10); Solomon (1973); Wallace (1998, Chapter 2). The cognitive content of emotions may be a non-belief-based way of “seeing as,” as when in the grip of a phobia you do not *believe* that the bunny is dangerous, but you see it as dangerous nonetheless (Calhoun, 2003).

<sup>94</sup> My argument in this paper does not depend on whether we go for coherence with the agent's beliefs and evidence or, instead, consider independent facts about the target. It might be best to say that it is internally appropriate for an agent to resent someone only when her evidence suggests that the target displayed ill will, while the target is in fact worthy of resentment when he in fact displayed ill will.

object is dangerous. In the same way, resentment represents its object as having shown ill will or insufficient regard.

Expressive views see resentment as directed at the resented agent, specifically in terms of being a form of address to her or containing a message for her. Such views correctly spot that there is directionality, but they mistake its nature. Resentment is not directed *to* the resented, but is instead *about* the resented. This emotion is directed in the same way that fear is directed to, because about, the feared object. One is resentful *of*, not necessarily *to*, the resented. Because the representation is about the target, representing that target as having a certain sort of will, the presupposition view properly locates the conditions of responsibility on the target herself.<sup>95</sup> That is, unlike the expressive view, the presupposition view not only captures the fact that resentment is directed, but also readily captures the direction. The relevant question is whether the content is accurate, whether the target agent really had ill will, rather than who could or could not understand a message with that content.

By understanding the propositional content as representational rather than communicative, the presupposition view understands resentment as structurally like other emotions. A syndrome including paradigmatic representational content, motivation and physiological symptoms typically characterize an emotion. To carry on our example of fear, it (paradigmatically) represents the target as dangerous, motivates avoidance, and includes elevated heart rate, various changes in blood circulation, muscle tension and so on. I have not focused on the motivational and physiological aspects of resentment; those features are not in dispute between the expressive and presupposition views and we can remain ecumenical regarding how

---

<sup>95</sup> Cf. Rosen (2015).

a full theory may specify the details of those features.<sup>96</sup> These views differ regarding the place of propositional content, and thus on whether some additional communicative feature is necessary for resentment though not for other emotions. The expressive view must hold that resentment is structurally different from other emotions, for it is not plausible to hold that other emotions likewise contain messages for their targets. We do not address earthquakes with our fear, and our disgust does not contain a message for slime mold. Thus, the expressive view understands resentment as different from other emotions not only in the specifications of its representation, motivation and physiological character, but also in having an additional communicative aspect. The presupposition view parsimoniously conceives of resentment as, essentially, just another emotion.

The presupposition view answers the problems we raised for the expressive view. Resentment, in representing the target as having shown ill will presupposes that the target was under the demand at the time of action. The presupposition view thus places the demand at the right time. This further provides the answer to the problem of necromancy, and distance generally, since the resentment does not need to communicate anything to a deceased agent. The appropriateness of resenting the dead is grounded in accurately representing the target as having shown ill will when they were alive. The target's current non-living status, or any other reason she may not be available for address, does not undermine the representation of her. We can thus see that major problems are avoided by holding that resentment rests on, reflects, or presupposes, rather than expresses, the basic demand.

---

<sup>96</sup> One central point of dispute will regard whether resentment comes with a motivation to harm the target or merely to engage in other behaviors. That dispute may itself be cashed out in terms of the representational content as regarding whether or not resentment comes with seeing harm to target as deserved or otherwise good.

### III.ii. Demands and Understanding

One might wonder when and how the demand for good will is expressed such that the target could have been under the demand at the time of action. At this point we may clarify, or perhaps reinterpret, the basic demand. Not only is the demand not expressed in the reactive attitudes, it may not have been expressed at all. Resentment presupposes that the resented agent was, as Nelkin puts it, “subject to... a demand to act differently at the time, whether that demand was actually expressed or not.”<sup>97</sup> It might seem strange to think of a demand that has not been expressed, but we should note that Strawson himself claimed that the demands in question were themselves constituted by our dispositions to attitudes like resentment. According to Strawson, “[t]he making of the demand *is* the proneness to such attitudes.”<sup>98</sup> So the demand does not necessarily have to be expressed for the target agent to be under the demand.

The essential point is that through resentment and our disposition to it, we hold the target to certain expectations. These are not expectations specifically for behavior, but instead for regard. An agent’s quality of will is normally described in terms of one’s regard for other people, but we can follow McKenna in saying that the good or ill quality of one’s will is constituted by “the regard or concern one has toward others (or oneself), and toward the relevance of moral considerations, as manifested in one’s conduct.”<sup>99</sup> The “moral considerations” clause is meant to

---

<sup>97</sup> Nelkin (2015, sec. 2).

<sup>98</sup> Strawson (2008, p. 34). It can be noted that this claim also favors the presupposition view as an interpretation of Strawson, for here an episode of resentment is not expressing the demand, but the disposition, existing prior to the episode, constitutes the demand itself.

<sup>99</sup> McKenna (2011, p. 59). The regard “is revealed in her reasons for action, her immediate proximal intention in acting as she did, and the attitudes she has for relevant others, as well as the broader motives or plans into which

account for the regard that may be owed to non-persons, such as animals.<sup>100</sup> Once that condition is added, the regard or concern for oneself and others is redundant; regard for people, insofar as it is demanded, is part of the relevant moral considerations. So, in discussing an agent's quality of will we will be considering the regard that agent shows toward morally relevant considerations, or, to put it another way, moral reasons.

The basic demand, then, is a demand for good will understood in terms of the target agent giving adequate consideration to certain morally relevant features of her situation. Of course, if she is non-culpably ignorant, say, of the effects of her actions, then resentment is out of place. She does not show disregard for consideration about which she is unaware. This helps bring out that the demand is on the agent's reasoning given certain features of her perspective.<sup>101</sup> To see

---

specific reasons and intentions for action are nested.” (2011, p. 18) Furthermore, omissions can be just as revealing (2011, p. 59).

<sup>100</sup> Gaus argues that we must distinguish “between (a) to whom must we justify ourselves? And (b) who are the beneficiaries of our moral duties[.]” (1990, p. 369, cf. 1996, p. 166) See also McKenna's (2011, pp. 11–2) distinctions between morally responsible agents, moral agents, persons, and moral subjects. Producing similar conclusions, Darwall writes: “although I am bound to insist that moral obligation, like the concept of right, cannot be understood independently of authoritative demands, the thought that moral obligations can be owed to beings who lack second-personal competence might be able to be elaborated in terms of trustees'... authority to demand certain treatment on their behalf.” (2006, p. 29) Other theorists provide many diverse strategies for showing that the interests beings who are not themselves part of the community of mutually morally responsible agents are morally relevant (Benn, 1988, p. 102; Kavka, 1986, p. 242; Korsgaard, 2004; Lomasky, 1987, p. 40; Morris, 1991, pp. 89–90; Rawls, 1999a, sec. 35; Vanderschraaf, 2011).

<sup>101</sup> I here leave open precisely what features of an agent's perspective are relevant. I am not attempting to provide a complete account of responsibility, and theorists may fill in the details in diverse ways while remaining within the framework I defend.

the person as having violated the demand for good will, then, is to see her as having had access to certain relevant moral considerations and failed to show due regard for those considerations in her reasoning.

Minimally, the agent does not show relevant disregard for considerations about which she is non-culpably ignorant when those considerations are in the form of ordinary descriptive facts, such as the physical consequences of an action.<sup>102</sup> It is, however, also broadly accepted that some kind of moral understanding is necessary, such as an understanding of moral principles. For instance, in considering Harris's understanding of the relevant moral features, McKenna raises the possibility that responsibility may require not only "*knowledge that* moral demands are made," but also "*knowledge of how* to guide one's conduct in light of those demands, and *knowledge of why* one should do so."<sup>103</sup> McKenna neither endorses nor rejects this requirement, for he is assessing the responsibility of Harris who satisfied this requirement well enough. We, however, are exploring more of the limits of evil where it is relevant that McKenna finds it at least plausible that accountability requires a significant understanding of the values supporting our demands. Other theorists clearly endorse requirements of this sort. For instance, Wallace holds that the required understanding includes "knowing when [the principle's] demands might require adjustment in light of the claims of other moral principles."<sup>104</sup> Or consider Shoemaker's view that a responsible agent "must be able not only to understand the various concepts deployed

---

<sup>102</sup> Though I emphasize those who hold that moral understanding is necessary, this view is controversial. For important arguments on each side of the dispute about exculpation through moral ignorance, see Harman (2011) and Rosen (2002). I argue that responsibility depends upon an agent's values in Van Schoelandt (2015, sec. 4).

<sup>103</sup> McKenna (1998, p. 131).

<sup>104</sup> Wallace (1998, p. 157).

in the moral principle in question but also to appreciate the justifications for the principle....”<sup>105</sup> Whether we think in terms of principles and justifications, or frameworks of values, the overall pattern we see is that it is not enough for the agent to understand merely what action we are demanding of them right now.<sup>106</sup> Responsibility requires that the target be able to significantly understand the values or justifications for those demands.<sup>107</sup> The sort of disregard that warrants resentment, then, seems to involve a failure to adequately account for the moral considerations that the agent adequately understands, while an agent who non-culpably does not understand these considerations displays ignorance rather than ill will.<sup>108</sup>

We can see here that the expressive view rightly emphasizes the importance of the target agent’s understanding, though the expressive view mischaracterizes the source of that requirement. What matters is, as McKenna writes, that the target has “an understanding of the shared framework of values from whence such demands arise.”<sup>109</sup> Such an understanding matters

---

<sup>105</sup> Shoemaker (2007, p. 75). Unfortunately, it is unclear to me how great the gap is in Shoemaker’s account between being “able to understand” and actually understanding. I thank McKenna for pointing out the importance of this gap.

<sup>106</sup> Pettit argues that “in order to be fit to be held responsible, an agent must be in possession of concepts like those that bear on what ought to be done, what is justifiable or unjustifiable, what is worthy of blame or praise, and must be in a position to understand reactions like resentment and gratitude, indignation and approval.” (2001, p. 29)

<sup>107</sup> Susan Wolf seems to have a view of this sort, in that she sees sanity as necessary for being a responsible agent and requiring “that one’s *values* be controlled by processes that afford an accurate conception of the world.” (1987, p. 55 emphasis original)

<sup>108</sup> Responses other than resentment may well be appropriate. For instance, it may be fitting to experience contempt, even if not resentment, for someone who is ignorant of important moral principles or values.

<sup>109</sup> McKenna (2008, p. 130). McKenna (2011, p. 12) also writes: “A morally responsible agent must be able to

not specifically for the target to understand the address of resentment, but instead for the target to have access to the demands. The target can show neither regard nor disregard for a demand if she is simply unable to understand the demand or its legitimacy. A resented agent is thus seen as having an adequate understanding of morally relevant values and yet to have given them insufficient weight in her practical reasoning.<sup>110</sup>

#### **IV. Return to the Limits of Evil**

With this presupposition account in hand, we can now reconsider the limits of evil. Recall that Watson proposed using cases of evil, or more specifically of a cruel and remorseless murderer, to illuminate and test a philosophic account of responsibility. One thing an account of responsibility, at least in the Strawsonian tradition, should do is help explain our ongoing practice of holding people. That practice provides us with more or less fixed points for which we must account. This includes our withholding resentment in light of common exculpating conditions, as per Strawson's examples. The theorist can muster these fixed points in support of

---

understand morally salient considerations, and must be able to control her conduct in accord with those considerations.”

<sup>110</sup> Not all emotions require that the target have this sort of understanding, or even any understanding. We can be sensibly afraid of a killer robot though the robot does not understand fear or danger. Likewise, we can be disgusted by a slime mold that fairly clearly lacks the capacity for any conceptualization. I would emphasize, however, that on the presupposition view this requirement for understanding makes resentment different in content, but not structure, from other emotions. The expressive view takes the understanding to be necessary for the special communicative aspect of the emotion, while other emotions are not said to have any communicative aspect at all. The presupposition view places the requirement for understanding in the cognitive content of resentment, and emotions generally have some cognitive content or other.

a particular account or to show the inadequacy of an account which conflicts with the fixed points. Watson considers our resentment of a cruel murderer like Robert Harris as a fixed point for which we must account.

I have argued that the expressive view does not adequately explain resentment of a cruel and remorseless murderer like Harris. The expressive account framed the issues in a way that makes it natural to worry about the fact the Harris might refuse to acknowledge our resentment and fail to listen to our demands. Worse yet, after the murderer himself dies there is no possibility of communicating with him. Resentment, however, seems perfectly fitting for someone like Harris, both while he was alive and after his execution. The expressive view can only explain this fittingness by appeals to counterfactuals that prevent the conditions of address from driving the theory. The presupposition view, in contrast, says that resentment is appropriate if the cognitive or propositional content of that emotion accurately presents the resented agent. That is, resentment of the murderer is appropriate if the murderer showed ill will in the sense of disregard for relevant moral considerations available to her in the way the resenting agent sees her. The fact that Harris would not listen and now is dead does not change the facts about whether or not he displayed ill will when he committed his crimes.

We can elaborate this further, and in doing so we will see why some cases of evil may be more complex than they first appear. Let us begin by emphasizing that Harris and similar cruel murderers acted on their own intentions to knowingly bring about harm to others. Harris was aware of the outcome of his actions, rather than, say someone who caused another's death by mere accident. What needs to be established is that Harris also had the relevant understanding of morality such that he had available, and yet did not adequately regard, relevant moral considerations. If the moral principles against killing were too complex or obscure for him to

understand, apply, or see as justified, then it might be wrong to see him as having shown disregard. That is, if the moral considerations were not available to him, he would seem ignorant or confused, but not to have had ill will.

Harris shot two teenagers in the back in a remote field. The boys had never done anything to wrong Harris and they posed no apparent threat to him or anyone else. Killing them was not even a means to fulfilling his further plans, for he had already secured the car he was stealing from the boys for use in a bank robbery. The fact that there was not even a weak justification for his actions highlights Harris's cruelty. It is hard for us to imagine someone being unable to understand the wrongness of such senseless violence, and it is all the more difficult to imagine when we consider that all of the main views in moral philosophy find such wrongness easy to show. Any test of universalizability would prohibit such violence, and the pain involved seems on its face a bad thing. There are many difficult cases in morality, including those involving the permissibility and limits of defense and retribution, but this does not seem to be a difficult case. Moreover, it is not merely those who ponder moral philosophy who come this conclusion. Restrictions on violence are prevalent in ordinary life and understood even by young children.<sup>111</sup> There is even reason to believe that typical development produces in children psychological mechanisms to inhibit violence.<sup>112</sup>

Those various potential sources of the moral insight support a strong presumption that an adult human being has access to the moral considerations against senseless violence. It is thus

---

<sup>111</sup> Depending on the kind of understanding necessary, children typically understand the wrongness of unjustified violence by the age of eight or perhaps as early as four-years old. See, Nunner-Winkler & Sodian (1988).

<sup>112</sup> Blair (1995).

rational for us to see an agent who engages in such violence as having shown disregard.<sup>113</sup> This presumptive story, however, can be overcome by evidence about the particular person and their limitations. For instance, it seems that psychopaths have various emotional deficits preventing them from experiencing typical moral emotions, having lasting care for others (or even themselves), or developing the typical internalized inhibitions on violence.<sup>114</sup> Furthermore, there may be people who, while not pre-disposed to psychopathy, have an unfortunate upbringing that prevents them from developing the relevant moral insight. A cruel and harsh upbringing or early childhood trauma, such as the conditioning experienced by a child soldier, may have this effect, leaving the person essentially morally ignorant, confused, or even insane.<sup>115</sup> Even if the agent has some sort of intellectual understanding of the moral requirements, his trauma could result in being unable to adequately see that there are real normative boundaries (as opposed, for instance, to simply cheap talk about such boundaries).<sup>116</sup> Such an individual, if they really did not

---

<sup>113</sup> It is worth noting that Plato (2005, p. 733d–734c) endorsed the opposite presumption, treating wrong-doing as a sign of illness and holding that any instance of vice results “either because of ignorance or lack of self-control, or both....”

<sup>114</sup> Blair (1995); Shoemaker (2007, sec. II).

<sup>115</sup> For examples of the conditioning and long term trauma suffered by child soldiers, see Pietrasik (2014). Recent experimental work finds that cases of upbringing-induced moral blindness reduces without eliminating responsibility. See, Faraci & Shoemaker (2010, 2014) I share Faraci and Shoemaker’s (2014, p. 14) speculative explanation of that data that “while it is not *difficult* for [the deprived individual] to identify and do the right things because of that childhood, it is nevertheless not overly demanding to expect him to do so.”

<sup>116</sup> In a related discussion, Laurence Thomas (1998, p. 158) writes that with some forms of abuse “there is a gulf—a moral gulf, in fact—between the proper normative story that the abused child learns and his corresponding normative experiences.” Without the normative experience, Thomas (1998, p. 161) argues, the moral belief may not have adequate “normative pull,” and the child may not be able to “fully grasp the importance of” the moral

understand the normative boundaries, would not expect or demand others to not harm him either. He might experience his interactions with others as like boxers in the ring, seeing it as bad for one to be hit, yet each being fully permitted to hit the other.

As Watson emphasized, Harris suffered a rather abusive childhood. Consider the following newspaper summary of his early life:

“He was born several months prematurely after his father kicked his mother with heavy construction boots. Because his mother was an alcoholic, he was born with fetal alcohol syndrome. Over the years, Mr. Harris was subjected to constant beatings, often to the point of unconsciousness. He traveled with his mother and siblings, living in tents and the family car, until he was finally abandoned by his mother at the age of 14.”<sup>117</sup>

Let us be clear, however, that this past is not relevant to holding Harris responsible for his actions because Harris’s past evokes pity. Pity alone is insufficient to undermine our resentment, for we are able to experience multiple emotions simultaneously and to see someone in multiple lights.<sup>118</sup> Consider the way that we may both pity and fear someone with a highly contagious disease, seeing them as in an unfortunate position and personally dangerous; we have reason to assist and reason to avoid her. If Harris’s past is relevant, it is as evidence that our presuppositions might not be met; the childhood trauma gives us some reason to think that Harris *might* not have been able to develop the moral understanding necessary for responsibility. Given his upbringing, he *might* have understood himself to be living in a world of violence where each is charged only to look after himself. The past, then, serves as evidence for how he might have actually been at the time of his crimes, particularly whether or not he had the kind of will we see

---

values.

<sup>117</sup> Gross (1990). Cf. Morain (1992). One can compare the traumatic childhood of violence and sexual abuse of serial rapist and murderer Carl Panzram, as discussed in his prison journals (2002).

<sup>118</sup> I thank Santi Sanchez for helpful discussion of this point.

him as having when we resent him.

If the cruel murderer's past did not prevent him from having the moral insight we presume, then the presupposition view readily justifies our resentment. Through resentment, we see him as having shown disregard for the moral consideration that was available to him, and he in fact did just that. If, however, he really lacked that understanding, the presupposition view holds that resentment is not fitting. This may at first seem to be a problem for the presupposition view, for resentment of the cruel murderer is a fixed case for which we want to account. What I want to suggest here is that the fixed case for us really involves cruel murderers for whom we maintain the presumption that they have the insight. Part of their evil is that they inflict senseless harm even in full understanding of the wrongness of that action. If we truly see someone as non-culpably lacking that moral understanding, we start to see him in a different light.

Consider the way we feel toward at least the intelligent monsters of fiction. Such monsters, like the Gorgon of Greek myth or predatory extraterrestrials of science fiction, are literally alien and inhuman. When monsters attack people, they may seem cruel, we may feel fear, dread, or hatred of them, but we do not typically feel resentment. We see monsters as forces of nature without any relevant quality of will, rather than as members of our moral community to be held accountable.<sup>119</sup> Those truly lacking moral insight appear to us as monsters, rather than

---

<sup>119</sup> Of course, even if monsters cannot be held morally accountable, they may still be controlled, driven away, kept out, confined, or slain. As Adam Smith (2009, p. 176) describes a case of a (religiously) misguided murderer, "the defense of society requires" legal punishment, but we should punish such crimes "with reluctance, when they evidently proceed from false notions of religious duty." That is, we may still have to coerce the individual, even when resentment would not be fitting. Cf. Shoemaker's elaboration of the separation of criminal from moral responsibility in (2011b, and 2013a). In a related vein, Plato justifies legal coercion as a means to control and limit harm from people presumed to be morally ignorant or ill. For Plato, as Mary Mackenzie (1985, p. 202)

morally accountable fellows, and the presupposition view helps us discern monsters from the responsibly evil. The expressive view leaves us stumbling in the dark. The presupposition view helps us navigate the complex terrain at the limits of evil.

## **V. Conclusion**

We often express our resentment, indignation, and even guilt. Such expressions can be found as much in eloquent, yet impassioned, speeches as in wild and violent outbursts. I will close this paper by turning attention to the iconic expression of anger, and I think indignation, from the film *Network*.<sup>120</sup> In the well-known scene that provides the epigraph of this paper, news anchor Howard Beale implores his television audience to yell from their windows that they are “as mad as hell and ... not going to take this any more!” What I think we can see, even in this scene, is that even when resentment is expressed, it is not itself fundamentally expressive. Resentment, instead, fundamentally involves a way of seeing the world and how others are relating to us. Inclusive of that is a certain way of seeing ourselves. Beale was shaking his audience from the complacency in which “we sit watching our TVs while some local newscaster tells us that today we had fifteen homicides and sixty-three violent crimes, as if that's the way it's supposed to be!” It is this way of seeing the world that Beale rebels against, trying to get them to see the world a different way, specifically seeing what is presupposed within our resentment. Experiencing resentment is one way to see, as Beale put it, “I’m a human being, goddammit, my

---

describes, “considerations of injury are divorced from those of injustice.” Some key nuances of Plato’s view are addressed in Malcolm Schofield (2012). I thank Rachana Kamtekar, Kristin Reglitz and Jeremy Reid for helpful discussion of Plato’s *Laws*.

<sup>120</sup> Sidney Lumet (1976).

life has value!”

## **Rawlsian Functionalism and the Problem of Coordination: Rethinking the Priority of Principles to Institutions**

“It is not the consciousness of men that determines their existence, but, on the contrary, their social existence determines their consciousness.”

Karl Marx, *A Contribution to the Critique of Political Economy*<sup>121</sup>

One broad tradition of political philosophy sees justice as a social technology whose point is to facilitate cooperation.<sup>122</sup> On such a functionalist view, justice is not so much discovered as invented. As a social tool, an acceptable conception of justice must be appropriate given the constraints we, the potential users of the tool, face. Though his functionalism is often underappreciated, John Rawls is a prominent figure in this tradition. For Rawls, we construct a conception of justice to serve as a basic social charter providing a public perspective and means to adjudicate competing claims on society’s basic institutional structure. There are many alternative ways to specify a functionalist understanding of justice, but for the sake of this discussion I take for granted some of the basic elements of Rawls’s account, such as the

---

<sup>121</sup> Marx (1979, pp. 11–12). Cf. Marx’s (1977, p. 7, 1992a, pp. 102–3) criticism that Hegel wrongly put abstract ideas before the concrete social form.

<sup>122</sup> Most prominent historically are Hume (2000, pt. II, sec. II) and Hobbes (1994, Chapter XIV), though ancients like Epicurus and Glaucon provide related accounts of justice (Epicurus, 1994, sec. 31–7; Plato, 1968, l. 338c ff, 357a ff; Thrasher, 2013, sec. 2). Recent accounts include Buchanan and Brennan (2000, Chapters 1, 7), D’Agostino (1996), Gaus (2011a, sec. 1), Gauthier (1986, Chapter V), Hayek (1978a, Chapter 1, 2007, p. 113), Schmidtz (2008), and Vanderschraaf (forthcoming, 2006, 2011).

importance of stability and conditions he lays out for the workability of a conception.

Section I provides a basic account of the functionalist project, as found in Rawls's work. Though the Rawlsian account is my point of departure, I shall argue that, contra Rawls, functionalist justice must be embodied and made determinate by social institutions and that, in a fundamental sense, these institutions are prior to principles of justice. Section II defends the claim that any functional conception of justice must be able to effectively coordinate the expectations of diverse members of society. Section III examines the problems reasonable pluralism raises for this critical coordination, and how institutions solve these problems. As section IV shows, supra-institutional principles are unable to effectively respond to this coordination problem. This leads to two conclusions in stark contrast to Rawlsian orthodoxy. First, there cannot be any adequate conception of justice for society's basic institutional structure or constitution. Since every particular conception must be rejected, we must reject the very idea of justice for the basic structure. Second, shared public reasons are needed within institutions, for interpreting and applying the requirements of those institutions, but are not needed for evaluating or reforming those institutions. Section V concludes by discussing the theoretical place of institutionally-emergent principles within a theory of justice, and how social institutions can be critically engaged or vindicated without appeal to shared principles.

## **I. Rawls's Functionalism**

Rawls (1999d, p. 394) is explicit that the aim of his theory of justice is not to discover moral truth "interpreted as fixed by a prior and independent order of objects and relations," but, instead, to locate "a *public basis* of political agreement."<sup>123</sup> Most important to the present

---

<sup>123</sup> Emphasis added. D'Agostino (2010, p. 4) likewise holds that "we should understand that claims about 'the right'

analysis is that this public basis of agreement is provided by a set of principles, along with the way these principles relate to each other, defining the criteria for resolving the disputes.<sup>124</sup> A successful conception of justice serves as a basic public charter for a society (1999a, p. 114, 1999f, pp. 306–7). Should such a charter be secured, there would be a fact about what conception of justice a society embraces, much as there is a fact about the society’s constitution and the language(s) spoken in that society. Such a conception of justice would be not simply normatively justified, but constitute a social fact. The Rawlsian moral theorist assesses and proposes possibilities for an acceptable conception of justice that may be socially adopted, and aims to understand the conditions a conception must satisfy if it is to play its “expected role in human life.” (Rawls, 1999i, p. 286)<sup>125</sup> A well-ordered society for Rawls (1999a, p. 4) is one in which there is public acceptance of such a conception. Such a society is not merely in compliance with a conception, it is *ordered by* it insofar as the citizens recognize their mutual acceptance and use their understanding of the conception’s principles to guide their political action (Rawls, 2005, p. 35).

---

are human constructs, in particular settings, and not discoveries of antecedently or independently existing facts about the world, whether of a natural or of some other kind.” Cf. D’Agostino (1996, p. 56) on the “social role” of a conception of public justification, and his emphasis that Rawls engages in social theory rather than conceptual analysis (2003, p. 60).

<sup>124</sup> For Rawls, a conception of justice would also include a justificatory basis, like the “original position” as a device of representation and the regulative ideal of society as a fair system of cooperation between free and equal people. For the present analysis we can treat a conception of justice as defined by its principles, though consideration of the other features would not significantly change the analysis.

<sup>125</sup> Rawls (1999i, p. 296) further indicates that a conception’s feasibility “is settled largely by psychological and social theory....”

The basic fact that sets the problem for such a public charter is that we have a wide variety of interests and views of morality and the good life, leading to conflicting claims on each other. A conception of justice is a public standard to adjudicate these claims. It is the role of justice to provide conclusive public adjudication of these claims, in the sense that the conception provides an answer to the dispute and no further considerations are needed to know that that is the public answer.<sup>126</sup> Effective adjudication creates settled expectations both about what others will do and what we ourselves must do, avoiding the heavy costs of conflict.

The function provides one basis for assessing proposed conceptions of justice (1999a, pp. 113, 117, 1999d, p. 394, 1999g, p. 1, 1999i, p. 291, 2005, p. 56). The proposed conception must, as D'Agostino (2003, p. 60) emphasizes, “function effectively in the circumstances which make such an understanding socially important.”<sup>127</sup> It matters, then, whether such a conception can be feasibly and stably implemented in our social conditions, characterized as they are by human limitations, inevitable disagreement, and even a fair degree of social contingency. We need to know whether the conception could really fulfill its function for people like us, given our natures as well as our traditions.<sup>128</sup> As Rawls (1999d, p. 390, emphasis added, cf. 1999f, p. 305) puts it,

---

<sup>126</sup> This ordering of conflicting claims leads D'Agostino to see Rawls as engaged in a project of making a technology for commensuration. For a related account of justice having the function of adjudicating competing claims, see Allen Buchanan (1990, pp. 241–2).

<sup>127</sup> Larmore (2013, p. 278) indicates that for an account like this, the “validity” of principles of justice “has to be judged by how well they handle the distinctive problems of political life, which are conflict, disagreement, power, and authority.” Mason (2012, p. 527) writes that for Rawls “principles of justice can be assessed in part in terms of whether they are well-adapted to serving a social purpose, that is, to governing a society.”

<sup>128</sup> Rawls (1999f, pp. 306–7) writes: “What justifies a conception of justice is not its being true to an order antecedent to and given to us, but its congruence with our deepest understanding of ourselves and our

our “social and historical conditions... profoundly affect the requirements of a *workable* conception of political justice....” Our social conditions matter, Rawls (1999d, p. 394) tells us, because a proposed conception “will only be able to *achieve this aim*” of providing a public basis of agreement if it adequately coheres with commitments in “the public political culture of a constitutional regime and... its most firmly held considered convictions.”<sup>129</sup> As with other social technologies, a society’s culture can provide both resources and obstacles for its effective use. The central point now is that according to this functional account we must judge proposed conceptions in part by their ability to function in the social conditions for which they are proposed, with some minimal degree of functionality being absolutely required. To capture this concern, we have:

*The functional requirement:* A conception of justice is acceptable only if it provides a workable tool for adjudicating disputes and facilitating social cooperation.<sup>130</sup>

Conceptions failing to satisfy the functional requirement are as such unfit to fill the role for

---

aspirations, and our realization that, given our history and the traditions embedded in our public life, it is the most reasonable doctrine for us. We can find no better basic charter for our social world.”

<sup>129</sup> As per D’Agostino’s (2003, p. 59) interpretation, we see here considered convictions as practical constraints on a conception of justice, rather than as evidence of an independent truth.

<sup>130</sup> This formulation is based on a related principle for a social morality from Gerald Gaus, (2013, p. 5). On the tradition viewing social morality as evolved and functional, see Gaus (n.d.). An important recent contribution is Philip Kitcher’s (2011, Chapters 2–3) account of ethics as invented, not discovered, and progressing through ethical-technological developments to better fulfill the function of addressing “altruism failures.” See also Brennan et al.’s (2013, p. 13) understanding of norms as tools with a function.

which they are proposed. Non-functional conceptions must, for that reason, be rejected.

Many will find a view of this sort alien, and perhaps not deserving of the name “justice.” G. A. Cohen (2008, p. 263ff), for instance, would argue that a functionalist account provides us only an account of “rules of regulation,” but not “justice.”<sup>131</sup> Cohen conceives of justice as a standard for assessing states of affairs independent of facts about motivational, informational, physical, and sometimes even moral constraints on what we could do.<sup>132</sup> Principles of justice, on Cohen’s account, do not need to be socially functional or action guiding (2008, pp. 267–8; cf. Estlund, 2011). I will not engage with that dispute, so I invite those with such a worry to understand my argument as stipulating a functionalist understanding of “justice,” and worried readers may replace it with an alternative.

## **II. Coordination as a Function-based Constraint**

The functional requirement leads to further particular constraints on a conception of justice (Rawls, 1999a, p. 113). These further requirements elaborate what it means for a conception to be workable. For instance, if the conception is to serve its social adjudicative

---

<sup>131</sup> Larmore (2013, p. 292n19) and James (2013, pp. 5, 104) each treat Cohen’s view as addressing a different concept they call “cosmic” justice. Mason (2012, p. 525) contrasts the Rawlsian project with the non-functional view that principles of justice either do not have a point or “their point consists simply in enabling us to evaluate any conceivable states of affairs in terms of the degree of justice they contain” – a task for which facts of feasibility are irrelevant. Anderson (2010), however, argues that it is Cohen who fails to address justice because his conception is unsuited for grounding our interpersonal demands.

<sup>132</sup> Cohen’s (2008, p. 7) method “investigates the shape of, and, consequently, the logical implications of, our deepest normative convictions.” Other moral considerations, even those “of nondistributive justice,” are left out of consideration.

function, it must have generality of form and produce an ordering of outcomes (compare Gaus, 2011a, pp. 294–303; Rawls, 1999a, p. 117). Conceptions that fail to satisfy those constraints are apt to fail to resolve many of our disputes, perhaps by allowing multiple conflicting answers where we need a determinate one.<sup>133</sup> Conceptions with indeterminate principles, like a mere requirement of mutual advantage, are inadequate because such principles leave conflicting claims within the scope of indeterminacy and thus fail in the task for which they are proposed. Furthermore, the principles must be usable by people in their deliberations, and thus there is “an upper bound of sorts on how complex they can be, and on the kinds and number of distinctions they draw.” (Baier, 1958, pp. 195–200; Gaus, 2011a, pp. 296–7; Mason, 2012, pp. 527–8; Rawls, 1999a, p. 114, 1999i, pp. 294–5, cf. 2005, p. 162) These function-based constraints are *requirements*, not mere standards for comparing conceptions. If proposed principles are too complex for ordinary people to use, require impractical metrics, or otherwise are unsuitable for serving the social function of the principles, then they cannot be the correct principles. Proposed principles failing these tests are invalid.

There is an underappreciated function-based condition on plausible conceptions:

---

<sup>133</sup> “[T]he principles must be specified so that they yield a determinate conclusion[.]” and “unless a political conception is complete, it is not an adequate framework of thought in the light of which the discussion of fundamental political questions can be carried out.” (Rawls, 1999a, p. 56, and 2005, p. 455) Rawls is driven by the need to not come up with public indifference too often. Consider, for instance, that Rawls (cf. 1999a, p. 59, 1999b, p. 135) rejects the Pareto efficiency criterion, for identifying only an eligible class, and thus being “at best an incomplete principle for ordering distributions.”

*The coordination requirement:* A conception of justice is acceptable only if it reliably leads to interpersonal coordination of expectations about the adjudication of claims, either on its own or in conjunction with some acceptable support technology.<sup>134</sup>

Recall that the conception is meant to provide a publicly recognized adjudicative outcome for some group with conflicting interests and views. The conception only provides such an outcome if members of the group, at least by and large, come to the same answers when applying it. If people applying the conception's principles come to different answers, they end up with conflicting claims needing a method of making a social choice among those claims. They remain stuck in the initial problem the principles were meant to solve. The coordination requirement thus follows from the functional requirement combined with the particular function of social adjudication that Rawls assigns to justice. A clear example of failing to meet this requirement is the principle that society should be organized in accordance with religious orthodoxy (or the moral truth). Among the reasons to reject the religious orthodoxy principle is that it will not coordinate expectations in the circumstances in which we need justice, for, as Locke (1983) wrote, "everyone is orthodox to himself." Each citizen would thus interpret the principle in light of her own religious views and come to systematically different beliefs about what justice requires, leaving no public resolution. Whatever categories are used in the principles cannot be the very things in dispute and leading us to need adjudication.

To further understand the coordination requirement, consider Rawls's (1999a, pp. 78–81)

---

<sup>134</sup> Cf. D'Agostino (1996, p. 16). By "support technology" I have in mind any other tool or process that may be used to coordinate expectations on various aspects, such as a process of monetization or recognized expertise.

defense of articulating principles in terms of “primary goods.” Primary goods, like monetary income, are mere means for the more normatively fundamental subjective satisfaction, well-being, project pursuit, our ability to advance what we value, or the like. But, Rawls argues, we cannot expect people to form common judgments in metrics requiring interpersonal comparisons of utility or well-being. We thus use income or some other good as our metric to provide a “feasible way to establish a publicly recognized objective and common measure that reasonable persons can accept.” (D’Agostino, 2003, pp. 74, 95–6; cf. Mason, 2012, p. 528; Rawls, 1999a, p. 81) Justice must, for functionality, deploy metrics tending to generate interpersonal agreement. Without such metrics, citizens will have divergent interpretations of justice and will not agree on the ordering of claims.

A proposed conception must tend to lead over time to people sharing an understanding and interpretation of its principles. To really adjudicate disputes, the principles must also systematically produce agreement about their application. This coordination requirement that has been overlooked, and some may remain skeptical of its basis in functionalism, so I will provide a second, largely independent, argument for it. I am going to argue that coordination is a prerequisite for solving the widely recognized assurance problem for principles of justice. As such, the many theorists who hold that a means of assurance is necessary for an adequate conception of justice much also acknowledge the coordination requirement.

The assurance problem arises, at least in Rawls’s account, because people condition their compliance with the requirements of justice on the compliance of others. On a functional analysis, there is no point in being the only person to act justly.<sup>135</sup> Rawls argues that

---

<sup>135</sup> More precisely, there is no point in doing the action that would be just if met with like action from others, though justice does not require such action when met with defection. One cannot act justly alone for much the same

“reasonable” people are motivated by a sense of justice to participate in a fair system of cooperation. This sense of justice, however, only obliges us to do our share “when we believe that others, or sufficient many of them, will do theirs.” (James, 2013, p. 104; Rawls, 1999a, p. 236, cf. 1999a, pp. 382–3, 2001b, p. 7, 2005, pp. 49–50; Weithman, 2010, p. 47)<sup>136</sup> As Rawls argues, otherwise cooperative people may be tempted to withdraw their participation because they suspect some others are not cooperating. Ominously, merely the general awareness of that temptation “may eventually cause the scheme to break down.” (Rawls, 1999a, p. 211) One test for a proposed conception of justice, then, is that members of a society with that conception will have a means of gaining mutual assurance and know that acting justly is not a lonely, and thus functionally futile and quixotic, quest. Without such assurance, the conception could not stably perform its function.

Diverse societies face an important and often overlooked difficulty. Discussions of assurance generally assume that what counts as cooperation is common knowledge. The assurance problem for justice in a diverse society, however, faces additional epistemic obstacles that compound assurance problems even among reasonable people. If an agent thinks the others are confused about what behaviors are cooperative, she should generally expect “defection” (from her understanding) even if others are cooperating by their own lights. The agents need

---

reason one cannot *co-operated* alone.

<sup>136</sup> For informative discussion of the Rawlsian assurance problem, see Thrasher and Vallier (Forthcoming). Hadfield and Macedo (2012, pp. 8, 9) write that “Rawlsian political liberalism... incorporates positive (predictive or behavioral) elements that give rise to the problem of mutual assurance...[.]” and that “in the absence of a practical, behaviorally realistic, resolution of this mutual assurance problem, a political conception cannot be stable and hence cannot be a reasonable proposal for the basis for political life.” Empirically, people vary in what they treat as a sufficient compliance from others to warrant their own (Bicchieri, 2006, p. 63ff).

shared judgments about the relevant actions as a matter of common knowledge (Chwe, 2013, p. 3; Vanderschraaf, 2008).

To clarify this dynamic, consider a simple modification of the well-known Stag Hunt Game. In the simple Stag Hunt, a stag can only be caught by the concerted efforts of the hunters, and it is assumed that there is but a single stag so all who hunt stag automatically coordinate. Were we instead considering hunting in a region with multiple stag and disagreement about which stag is the best to hunt, an important new element of assurance is introduced. Consider the game displayed in Figure 1.

	Stag #2	Stag #1	Hare
Stag #2	3, 2 *	0, 0	0, 1
Stag #1	0, 0	2, 3 *	0, 1
Hare	1, 0	1, 0	1, 1 *

Figure 1: Asymmetric Payoff, Three Equilibria (\*) Stag Game (ordinal utilities)

The possibilities for failure of coordination in a typical Stag Hunt remain, but now we have an additional complication. There are two stag, and our hunters have conflicting preferences over which stag would be best to hunt.<sup>137</sup> As such, it is not obvious for either which stag to pursue if they do hunt stag. If each goes for his or her own preferred stag, they will catch nothing and go hungry. If they each try to hunt what the other prefers, say each thinking the other is stubborn, then again they each go hungry. Until something is added to establish the

---

<sup>137</sup> The impure coordination game best models the social disagreement about justice, though the difficulties arise even in a pure coordination game. Cf. Hadfield and Weingast (2014, p. 141).

coordination point there is good reason for each to hunt rabbit. I will emphasize here that it does not seem to solve the problem for each to display their genuine preferences, or even for some to make obvious that they will certainly be engaged in some (unspecified) stag hunt.<sup>138</sup> Their mutual enthusiasm is little consolation for a starving hunter after a failure to coordinate.

Bringing the point explicitly back to justice, we need assurance that our attempts at cooperation will not be wasted, perhaps even simultaneously costly and counter-productive.<sup>139</sup> If we all agree on what justice requires, then assurance is primarily an issue of making it a matter of common knowledge that we are all trying to do our part. If, however, there are different ways to understand what justice requires – as with competing conceptions – then, like our hunters, those who would pursue justice must worry not just that others may not try, but that even those pursuing some conception of justice will be working at cross-purposes. Recall, however, that the need to solve the assurance problem is a constraint on conceptions of justice. If there are not adequate means for solving the assurance problem accompanying a conception, then that conception would neither be normatively guiding for reasonable people nor serve its social function. We thus see that the coordination requirement follows not only from the adjudicative function of justice, but also from the conditions of mutual assurance.

Coordination failures are not an idle possibility, for Rawls himself claims that a society

---

<sup>138</sup> What D’Agostino (2012, p. 340ff) calls a “shallow consensus,” or agreement only at a high level of abstraction, is insufficient for coordinating a hunt or, I argue below, for justice.

<sup>139</sup> The potential for counter-productivity arises from policy interdependency and the multi-dimensional nature of justice. (Gaus, 2012b, sec. 3.2–3; E. Ostrom, 1999, sec. 4; Sen, 2009, p. 99; Simmons, 2010, pp. 34–5) This potential can be seen as implied by the economic general theory of the second best (Lipsey & Lancaster, 1956). Leeson and Williamson (2009) and Leeson (2014, Chapter 10) illustrate the potentially radical political implications of second best theory.

well-ordered by a single conception of justice is unrealistic and that the best we can expect is social convergence on a “family” of liberal conceptions (cf. Freeman, 2006, pp. 255–6; 1997, pp. 581, 615, 2005, pp. xxxvi, lii–liii, xlvi. 6f, 223, 490). Disagreement about conceptions of justice are *inevitable* (Rawls, 2005, p. 227). Other Rawlsians likewise find competing conceptions, or perhaps competing interpretations of an abstract conception, likely or inevitable (Lister, 2013a, p. 11; 2000, p. 44, 2010, p. 30, 2011, pp. 11, 13–4, 20; Quong, 2010, p. 183). Such disagreements about justice tend to arise because of the same the burdens of judgment that lead to other disagreements (cf. Besson, 2005, pt. 1; D’Agostino, 1996, sec. 6, 2010, sec. 3.1; Rawls, 2005, p. 56).<sup>140</sup> For instance, indeterminacy and vagueness of concepts, political or otherwise, generate divergent interpretations and judgments (Kiser & Ostrom, 1999; Macedo, 2011, pp. 10–1; Quong, 2010, p. 148; Rawls, 2005, p. 223). Furthermore, significant disagreements arise about how trade-offs will be made among the relevant considerations (Arneson, 1990, sec. IV; cf. A. Buchanan, 2004, p. 183; D’Agostino, 2003, pp. 133–7; Lister, 2013a, pp. 17, 83; Quong, 2005, p. 316; Rawls, 2005, pp. 57, 335–6). And we should not overlook the way disagreements about both the conception of justice and its implications arise from complex and conflicting empirical evidence (Lister, 2013a, pp. 17, 83; Rawls, 2005, p. 56).<sup>141</sup> This diversity of conceptions of justice, however, leaves us with much the same problem that led us to need a conception in the

---

<sup>140</sup> Interpretive disputes arise even in groups sharing religion, culture, and profession, such as the Maghribis traders – see Hadfield and Weingast (2013a, p. 23ff). Muldoon et al. (2014) show that disagreement even characterizes the Rawlsian “original position.”

<sup>141</sup> Lister (2013b) highlights the results of economic disagreement, showing that F. A. Hayek and Rawls shared much at the abstract normative level, yet had significant policy disagreements derivative of empirical disagreement. Cf. Kofman (2012, p. 249).

first place. We began with conflicting claims based on non-public, idiosyncratic perspectives; if we do not concur on a conception of justice, our understanding of the social ordering of claims will be based on another idiosyncratic perspective.

So long as you believe others are likely to hold wrong views about what the conception of justice requires, you have reason to believe that they will fail to do their part to maintain justice even if they try. We can see that the problem of coordination for purposes of assurance is actually a more difficult problem than merely coordination, for even if somehow every citizen did share a judgment on what justice required, assurance requires coordination as a matter of common knowledge. In the next section we will consider how institutions can help diverse agents coordinate their expectations even under the burdens of judgment.

### **III. Institutionally Mediated Coordination**

The analysis thus far has been limited in two ways. First, we have considered fairly abstract issues concerning justice as a social tool of adjudication for which coordination is essential. Second, the idea of disagreement and coordination has been used in an intuitive, imprecise, way. We can overcome these limitations by deriving some lessons from Gillian Hadfield and Barry Weingast's impressive coordination model of the law and case studies illustrating its application. The importance of this model to analysis of public reason has been recognized and explored in an important paper by Hadfield and Macedo (2012).<sup>142</sup> This model emphasizes the importance of, and difficulties for, coordinating on classifications, and it provides a means for coordinating on public resolutions to at least some of the conflicting claims of deeply pluralistic people. I will provide a brief description of their model before discussing the

---

<sup>142</sup> I draw conflicting lessons from the model in section IV below.

implications for coordination generally. After considering their law-focused analysis, we shall re-ascend to the more abstract issue of adjudication by principle of justice.

The Hadfield-Weingast model shows the possibility of diverse agents coordinating expectations about which actions are wrongful; an equilibrium is established at which wrongful actions are deterred by the threat of decentralized punishment from diverse.<sup>143</sup> It starts with two buyers, Alf and Betty, potentially interacting with a seller, Sam, over infinite discrete periods. Buyers can contract with Sam for a good with expected value  $V$ , for price  $P$ , where  $P < V$ , and for which Sam selects a performance vector,  $\chi$ , including  $n$  potentially relevant variables like material of the good or production methods. Sam may always provide the good as the buyer expects, but also probabilistically gains an opportunity to reduce costs through a particular performance (e.g., using a different product source). Taking this opportunity will either preserve or eliminate the value for the buyer, as determined by his or her own idiosyncratic logic,  $I^i$ . The logic maps all of Sam's performances on to categories that I shall label "ethical" or "unethical." Note that the criteria in anyone's idiosyncratic logic need not be narrowly egoistic, but may regard fairness or the like. It is assumed that the risk/benefit ratio of buyers is sufficient to maintain buying despite Sam's opportunism, and that the credible threat of collective boycotts from Alf and Betty together would deter Sam from performances. Sam will refrain from performances known with certainty to trigger such boycotting. The problem for Alf and Betty is that they do not know, and it is too costly to discover, the other's idiosyncratic logic ( $I^i$ );

---

<sup>143</sup> Hadfield and Weingast (2012, p. 477n3) note that though the model is described with of sales, it "can be interpreted as applying to any setting in which there is a potential wrongdoer who may exploit a community of potential victims who have the capacity to impose some penalty on the wrongdoer." Likewise, boycotts in the model can be replaced with other forms of costly punishment.

consequently, buyers cannot coordinate their boycotts, leaving them only ineffective independent choices not to buy from Sam.<sup>144</sup> However, it is only if they act together that they can affect Sam's performances via his anticipation of boycott. This coordination problem and way the agents can overcome it are the key for my present analysis.

Alf and Betty can reach deterrence equilibrium with Sam, the model shows, through a third-party institution providing a common logic,  $R$ , for classification of Sam's performances. I consider the common classifications as "legal" and "illegal." The model assumes  $R$  is inaccessible, but the institution provides a common knowledge representation,  $\check{R}_t$ , such as precedent or incomplete rules with methods for predicting future decisions from this information. Whenever Alf or Betty believes that Sam took an unethical opportunity, that buyer may call for a ruling from the  $R$  institution on its legality. This makes possible "strategy  $R$ " of buyers purchasing goods only in periods for which no actions found illegal were taken in the prior period. This strategy would deter Sam from taking illegal actions (though not all actions categorized by some buyer as unethical).

Each buyer uses  $\check{R}_t$  to assess the probabilistic convergence,  $r_t^i$ , of his or her own ethical standard with  $R$ . A buyer only pursues strategy  $R$  when the expected value of the deterrence equilibrium is greater than the no-deterrence equilibrium. The values of the equilibria depend upon the probabilistic convergence with the buyer's own standard ( $r_t^i$ ), as well as the expected frequency and magnitude of loss from potentially deterred performances. Buyers reevaluate convergence with each new ruling ( $\check{R}_t$ ). It is critical to stress that because  $R$  is not directly accessible, the continued convergence for each agent is uncertain to the others. The deterrence

---

<sup>144</sup> Hadfield and Weingast (2012, p. 478) model each agent's idiosyncratic logics as "fully private and inaccessible to others at reasonable cost."

equilibrium based on strategy  $R$  thus depends on maintaining the belief for each agent that  $R$  remains sufficient convergent for each buyer; evidence of this is provided through boycott participation. The expected convergence for each agent must be high enough to overcome the uncertainty about the convergence for others. With discount factor  $\delta$ , and  $i$  assigning probability  $(1-p^i)$  of sufficient convergence for  $j$ , the following inequality must be satisfied:

$$r_t^i > \frac{1-\delta}{\delta} \frac{1}{1-\rho^j} \left[ \frac{(1-\theta^i)V - P}{\theta^i V} \right] \quad (\text{Hadfield \& Weingast, 2012, p. 488})$$

With cost potentially saved through an illegal opportunity  $c$ , Sam is deterred when:

$$c < \frac{2\delta(1-\rho^i)(1-\rho^j)}{1+2\delta(1-\rho^i)(1-\rho^j)} P \quad (\text{Hadfield \& Weingast, 2012, p. 489})$$

In equilibrium, Sam is deterred from (enough) illegal acts, despite engaging unpunished in acts that some label “unethical,” while every agent believes based on past behavior that  $R$  is anticipated to be sufficiently convergent for each buyer to warrant their participation in collective punishment of any illegal performance.

The model is important to us for seeing how the diverse agents, with idiosyncratic classification systems, coordinate. The agents need to coordinate expectations so the seller can avoid, and the buyers can jointly punish, actions deemed wrong by the common logic. How the agents achieve this coordination, despite their disagreements, points the way to answering the coordination problem for conceptions of justice presented in Section II. We must consider two coordination-enabling features of the common logic that can help address coordination for a conception of justice.

The first feature enabling coordination is the predictability of the logic itself. For shared understanding and mutual anticipation, particularly in novel and unforeseen circumstances, the

common logic must include a stable, impersonal method of reasoning with general rules, principles, or categories.<sup>145</sup> Though Hadfield and Weingast do not emphasize this fact, since the model assumes the idiosyncratic logics are mutually inaccessible at reasonable cost the common logic must be more accessible than an idiosyncratic logic, though not even *R* is directly accessible. In practice, the common logic must undoubtedly be simpler and more restricted in the categories it deploys. In order for the diverse individuals to use a singular common logic, it must deploy relatively simple rules to maintain interpersonally clear criteria. For instance, agents have diverse, complex, context-dependent criteria for ethical wages, but the legal requirements are relatively few and appeal to interpersonally clear criteria like a flat minimum. The common logic's clarity is particularly important for achieving the common knowledge of coordination, for there is a crucial difference between each agent privately understanding and the agents mutually recognizing that they understand.<sup>146</sup> These features should not be too surprising; our discussion of Rawls above highlighted that a conception of justice must be general, simple enough to guide deliberations, and utilize metrics lending themselves to intersubjective agreement (e.g., income).

The common logic, however, cannot be so refined as to do away with the burdens of

---

<sup>145</sup> Hadfield and Weingast (2012, p. 474, 2013a, pp. 5, 9, 2013b, pp. 34–6). In related vein, Hayek (cf. 1978a, p. 87, 1978b, p. 127) writes: “As the judge can be just only if he follows the principles of the law and disregards all the circumstances not referred to by its abstract rules (but which may be highly relevant for the moral evaluation of the action), so the rules of justice must limit the circumstances which may be taken into account in all instances.” On use of shared mental schema, see Bicchieri (2006, Chapter 2).

<sup>146</sup> In the model, Sam is only deterred when the probability of Alf and Betty coordinating a boycott is high, and Alf and Betty each have incentive to participate in the boycott if they believe both that the other will participate and that Sam will sufficiently understand *R* to be frequently deterred from illegal action. Coordination is made all the more important when we extend the model to account for the fact that rulings on cases are costly.

judgment. Agents will encounter new and difficult cases in which their interpretations of the common logic diverge, as when the rules are vague or unexpectedly conflict. On Hadfield and Weingast's (2012, pp. 479, 495) account, this is addressed by ensuring that the common logic is neither a "disembodied classification scheme" nor merely "a set of rules or principles."<sup>147</sup> Instead, and this is the second key feature for coordination, an interpretive institution provides the common logic; there is no common logic without that institution, or, more strongly, the common logic *is* the institution.<sup>148</sup> This is an absolutely fundamental point. Any shared principles, to the extent that they matter for coordination, come from the institution. In the model, the agents understand the common logic through representations of it, adjusting their understanding as new representations are provided in the resolution of disputes by the adjudicative institution. This institution enables the diverse agents to coordinate their understanding of the common logic and, thus, continue to share a single classification system to coordinate their interpersonal expectations.

Hadfield and Weingast (2012, p. 495, 2013a, p. 12, cf. 2013b, pp. 33–4) emphasize the need for an authoritative "steward": "a unique arbiter able to resolve complexities, ambiguities, and gaps" and to "adapt the common logic to changing circumstances."<sup>149</sup> A paradigmatic

---

<sup>147</sup> It is, likewise, *not* the public reason idea Macedo (2011, p. 23) emphasizes of "a shared freestanding case for our shared principles in the currency of a public moral system of ideas."

<sup>148</sup> Though Hadfield and Weingast (2012, p. 479 cf. 490) "emphasize that a logic is an institution," they elsewhere write of the institution *supplying* the common logic (2012, pp. 479, 484, 2013b, pp. 26, 30). Though Macedo (2011, p. 10) indicates that institutions "help generate, sustain, and improve the shared logic," the Hadfield-Weingast model would have us drop "help."

<sup>149</sup> In criticizing the Hobbesian view that interpretive disputes requiring an arbiter are inevitable, Ridge (1998, sec. III) focuses on the ubiquity of shared interpretation of immediate commands. His criticism, however, does not

steward was the Law Speaker of Medieval Iceland — an individual nominated to memorize, publicly recite, and be the sole authoritative interpreter of all of the laws. When interpretive disputes arose, the Law Speaker provided the definitive public answer. Other examples include a court system, village elders interpreting custom, or church hierarchies interpreting religious requirements. These are all ongoing institutions with histories, commonly known to adjudicate interpretive disputes for the relevant community.<sup>150</sup> Each of these steward systems, whether they involve an individual agent with the social role of steward or a known procedure of public adjudication, provides a means to make the common logic determinate, and so bring the agents back to coordinated expectations when confronted with the inevitable interpretive disputes. Hadfield and Weingast thus illuminate an insight also had by social contract theorists like Hobbes, Locke and Kant. Interpretive differences, like those regarding the natural law in a state of nature, undermine cooperation, but adjudicative institutions can provide a common understanding.<sup>151</sup>

---

address the application of the shared principles in novel and unforeseen circumstances that, under the burdens of judgment, allow multiple reasonable interpretations. Furthermore, the primary worry is not that such disagreements may ever arise, but that they will persist, multiply and escalate.

<sup>150</sup> The steward's history, general prestige and other features can contribute to it being a focal point for the diverse members. On the importance of focal points to cooperative interactions, see Schelling (1960, Chapter 3). Though Hadfield and Weingast emphasize formal, monocentric institutions, I believe the analysis can extend to informal institutions and polycentric order – cf. Hadfield and Weingast (2012, p. 509), V. Ostrom (1999). The key to such an extension is understanding the feedback mechanisms for bringing diverse members to common interpretations in decentralized and informal settings.

<sup>151</sup> Locke (1681, sec. 124); Kant (1999, pp. 116–9, 146); Hobbes (1994, Chapters 5.3, 26.21). As Gaus (2011b, p. 82) writes, “Hobbes, Locke, and Kant trace the problem of social disorder to conflicting claims of private

The picture we have so far, then, is a system of institutions with laws publicly adjudicating competing claims about acts, and a steward adjudicating competing interpretations of the laws. The concepts and principles within the common logic emerge from the institutions, gaining meaning and determinacy as the disputes are adjudicated. Consider, for instance, the legal principle of *volenti non fit iniuria*, according to which someone cannot legally claim harm arising from risk they voluntarily assumed. *Volenti* is a broad principle, useful for discerning legal liability or lack thereof in diverse, in many ways novel, situations. This principle has developed to include standards for assessing adequacy of information, what risks are explicitly or implicitly assumed, and so on. We likewise see that concepts like legal personhood or trespass-warranting necessity can function in a legal system and gain determinacy from the settling of disputes over time. Without adjudication, abstractly stated principles and concepts lack sufficient determinacy and would leave the agents in perpetual interpretive disputes. It is the meaning that develops from the institution's activities over time that provides the agents a common way to reason about their circumstances.

If we think of the institutions applying to individuals as first-order institutions, we can extend the model to address second-order institutions, or institutions that themselves govern other institutions. Hadfield and Weingast (2014) extend their model to constitutions, with the

---

conscience about morality, and all see the solution to be the establishment of the state as the final umpire with respect to the demands of justice.” In a related vein, Bernard Bosanquet (2001b, p. 19) argues that the scale of cooperation may always be increased, along with the diversity of members and associations, “so long as it is made clear to what system of authorities every separate human being is subject in respect of the ultimate adjustment or claims upon him.” For this, he writes, “there must always be at least some machinery for making this clear (like the Court which interprets the constitution of the U.S.A.), if civilised life is to be possible.”

main changes being that instead of classifying the actions of individuals as legal or illegal, we consider the actions of the sovereign (potentially an institutional agent) as constitutional or unconstitutional.<sup>152</sup> There remains in the model a common logic of principles, concepts, rules and the like, as well as a central role for a steward to settle the disputes that inevitably arise (e.g., a constitutional court). Here again principles emerge, though this time from the constitutional institutions whose ongoing adjudication makes the general principles determinate and publicly settles the inevitable interpretive disputes.

Institutions embodying the common logic can coordinate expectations and serve a public adjudicative function with regard to actions and lower level institutions. In this way, the common logic serves the role of a public charter providing a public basis for agreement. On this model, at some point we eventually reach a highest level of institutions. The various citizens will not have a common logic or other shared perspective for assessing the constitution or basic structure, but will instead judge it from their diverse, idiosyncratic perspectives. The Hadfield-Weingast model thus lacks any shared means to assess the common logic institutions themselves. We will see in the next section that the prospects for developing such shared means are grim.

#### **IV. A Conception for the Basic Structure**

Members of a society make divergent assessments of their social institutions. Reasoning from their own individual standards they might advance complaints, but these will seem idiosyncratic to many others. To settle disputes about the quality of the institutions and potential

---

<sup>152</sup> Buchanan and Brennan (2000, Chapter 7.IV) extend the justice of individual behavior understood through rules to the justice of rule selection through meta-rules.

reforms, they need public principles for judging institutions themselves from a non-idiosyncratic perspective (Mason, 2012, p. 525; Rawls, 1999c, p. 53). Such principles are supra-institutional; they stand above the institutions they are regulative over. For Rawls (1999f, p. 305), “[t]he social role of a conception of justice is to enable all members of society to make mutually acceptable to one another their shared institutions and basic arrangements...”<sup>153</sup> Serving this role requires that the principles be understood independently of the institutions, providing independent grounds for comparing institutional possibilities and potentially overriding existing institutions.<sup>154</sup> Suppose, for instance, that someone raises a complaint about part of the current democratic system or constitutionally entrenched rights, and proposes a reform. The current institutions do not simply vindicate themselves, through their mere existence, against proposed reforms. Instead, citizens judge the institutional possibilities, on the Rawls’s (1999a, p. 4, sec. 31) account, in accordance with the supra-institutional principles of justice.

This role for the conception of justice can be clarified by considering how it is applied to constitutions. According to Rawls (2005, p. 444), the conception is the source of “reasons for supporting fundamental political positions....” We use the conception’s principles as a guide for selecting a constitution; the principles both constrain the choice of constitution and rank political procedures (Rawls, 1999a, p. 172). Though the principles will not dictate, for instance, economic

---

<sup>153</sup> Rawls is focused on a kind of justice applying to a society’s basic structure, though there are other forms of justice. For discussion, see Neufeld and Van Schoelandt (2014).

<sup>154</sup> According to Rawls (1999a, pp. 117, cf. 116, 174), principles of justice “override the demands of law and custom, and of social rules generally. We are to arrange and respect social institutions as the principles of right and justice direct.” Likewise, a conception of justice justifies a constitution (2005, p. xlv). Macedo (cf. Hadfield & Macedo, 2012, p. 10; 2011, p. 9) writes, these principles “undergird and inform our shared law....”

policy on their own, the principles should provide the basis for discussing “the question of private property vs. socialism” and guide that decision in light of the society’s material and cultural circumstances.<sup>155</sup> Overall, the conception of justice provides an “Archimedean point” or “shared point of view from which they can resolve questions concerning the constitutional essentials,” and the resources for “agreement [on constitutional matters] among those who affirm opposing comprehensive doctrines” (Rawls, 1999a, p. 230, 2001b, pp. 32, 183, cf. 2005, pp. xvii–xlviii, 35, 48). Further questions of legislation are considered in light of the conception, the selected just constitution, and further social facts.

It is of utmost importance to see that this is a process of applying the principles, not a process of giving the principles greater determination. That is, for Rawls (1999a, p. 173), the principles must have a determinate meaning conceived prior to the selection of a constitution or other institutions, for the principles “define an independent criterion” for those institutions.<sup>156</sup> The conception of justice provides a standard by which to publicly judge a constitution as just or unjust. The conception of justice, then, must already have a determinant meaning that the members of the society share. As Rawls puts it, a “conception should express principles, standards, and ideals, along with guidelines of inquiry, such that those values alone give a reasonable answer to all, or nearly all, questions involving constitutional essentials and matters

---

<sup>155</sup> The quote is of Rawls from a correspondence quoted in DiQuattro (1986, p. 307), and the same sentiment is expressed in Rawls (2001b, p. 139). My interpretation contrasts with that of Pogge (1989, pp. 201–3).

<sup>156</sup> Rawls (2005, p. 9) similarly writes that the conception of justice “provides a publicly recognized point of view from which all citizens can examine before one another whether their political and social institutions are just[.]” and that with the conception “[s]ociety’s main institutions and how they fit together into one system of social cooperation can be assessed in the same way by each citizen....”

of basic justice.” The principles must be determinant in selecting a constitution; selecting a constitution cannot determine what the principles require of a constitution. The constitution and other institutions are essential, on Rawls’s account, for determining various derivative requirements for just legislation, and perhaps just individual action within the laws, but the constitution itself must comply with the independent principles of justice. We would, for instance, expect that for our social conditions either markets or socialism better satisfy the principles of justice for the basic structure. We would not take the content of those principles to depend upon the existing institutions, as if instituting a market creates different standards of liberty, opportunity, or wealth within the conception of justice. That the conception cannot gain determinacy from the constitution is made all the more clear by the fact that it is to “guide how the constitution should be *amended*...” and provide “ideas in the light of which the constitution as so far understood could be *changed*.” (Rawls, 2005, p. 165 emphasis added) Again, the conception is meant by Rawls to provide a shared, Archimedean, point of view providing public resolution to our disputes, in the first instance about what the constitution ought to be.<sup>157</sup>

As I noted, Hadfield and Stephen Macedo explicitly connect the Hadfield-Weingast model to the Rawlsian project. They argue (2012, p. 8) that the “quite distinct approaches” of the Hadfield-Weingast model and Rawlsian normative philosophy “converge on strikingly similar

---

<sup>157</sup> The conception may produce indifference among some options (Rawls, 1999a, p. 176, 2001b, p. 41). Though it does not select a single best option, it does give us the determinate public answer that the proposals are *equally* just from the public perspective. The interpretative disagreement that we have focused on does not produce a public indifference. When there is pluralism about justice, instead of agreeing that the matter is publicly indifferent, each holds that one of the answers is in fact superior in accordance with her understanding of justice; each holds that there is a more just institution, though they disagree about which institution that is.

accounts of public reason.”<sup>158</sup> Hadfield and Macedo (2012, p. 8) hold that diverse members of society require “reciprocal commitment to basing law on a system of shared reasons.” The members, then, have a duty in discourse regarding potential reforms to the constitution and laws to appeal only to the shared reasons, “rather than whatever reasons they personally find compelling[.]”<sup>159</sup> Mutually acceptable values, evidence, and ways of reasoning provide the basis or material for the shared conception of justice guiding our judgment and reform of institutions.<sup>160</sup> The Rawlsian conception of justice provides a system of reasoning independent from the idiosyncratic views of the diverse members of society, and is in this way like the common logic of the Hadfield-Weingast model.

There is, however, a key difference between the Hadfield-Weingast model and Rawls’s project.<sup>161</sup> A Hadfield-Weingast common logic makes institutions primary, with the principles of

---

<sup>158</sup> Though they emphasize the use of “public reasons,” like other Rawlsians these reasons are used to formulate a conception of justice. Nothing should be lost by my focus on the conception.

<sup>159</sup> Hadfield and Macedo (2012, p. 7). They also (2012, p. 14) quote Rawls (2005, p. 135) regarding endorsement of a constitution in light of shared principles.

<sup>160</sup> Hadfield and Macedo (2012, p. 16). I will note that Hadfield and Macedo (2012, p. 16) write, as many Rawlsians do, as if the conception must employ independently forceful “moral reasons,” such as those found in democratic political culture. What the Hadfield-Weingast model shows, however, is simply that a common logic has to deploy mutually intelligible concepts and reasons, but not that these be thought to be of independent value. Social institutions often deploy proxies and heuristics, with cases settled by matters that outside of the common logic seem arbitrary, such as whether a light is red or green (D’Agostino, 2003, p. 95ff).

<sup>161</sup> Though Hadfield is correct that her model is relevant to this philosophic dispute, I am here arguing that she improperly translates the model for that dispute. Though they correctly emphasize the features making a common logic predictable, they do not account for the essential adjudicative institution or stewardship of the model. Even the ablest theorist runs such risk in moving a positive political theory model to political philosophy.

the common logic always subsidiary to the institutional system of authoritative stewardship that coordinates interpretations and expectations among diverse members in complex, changing circumstances. The institutions produce the logic, and the shared system of reasoning is internal to the legal system.<sup>162</sup> The Rawlsian conception, however, gives principles primacy over the institutions, specifically providing a shared means to select and judge the institutions themselves. This difference of priority has significant consequences. As we observed at the end of section III, the Hadfield-Weingast model lacks a shared means of assessing the most basic institutions. On the other hand, the Rawlsian principles lack sufficient means to coordinate the members on a single interpretation and shared expectations, since their interpretation faces the burdens of judgment and, as noted above, the constitutional stage is not an exercise in interpretation.<sup>163</sup> The necessary institutional independence leaves the Rawlsian principles vulnerable to the previously discussed inevitable disagreements and unable to meet the coordination requirement. If a conception of justice allows for multiple and inconsistent interpretations, and does not include within itself a device to coordinate on a shared interpretation, it fails to meet the coordination requirement, and thus is inadequate.

We see that there will not be a common logic regulating the basic institutions, and so the model does not support restricting political discourse to shared reasons. The model, in fact, supports leaving political discourse open to the full range of members' values because only such

---

<sup>162</sup> Recall that Hadfield and Weingast (2012, p. 495) define “a logic not merely as a set of rules or principles but rather as the product of a third-party institution.” They extend their model to constitutions, and maintain the necessity of authoritative stewardship, as by a constitutional court (2014, pp. 125, 134–5, 149).

<sup>163</sup> Hadfield and Macedo (2012, p. 17) even repeat Rawls's view that members will hold different conceptions within a liberal family.

openness provides insight into whether or not the institutions are mutually beneficial. As Hadfield and Macedo (2012, p. 9) indicate, we must expect rational people, with their diverse evaluative commitments, to only support institutions that they expect will benefit them (Cf. Hadfield & Weingast, 2014, pp. 126, 142). As Hadfield and Weingast (2012, p. 487 emphasis original, 2013b, pp. 30–1, cf. 2014, pp. 142, 149) model this, each agent must find the common logic “*sufficiently convergent*” with her idiosyncratic classification system. Public discourse, contrary to the claims of Hadfield and Macedo, must be open to “whatever reasons [members] personally find compelling,” so that the institution can adapt to their needs and so members can tell whether or not the institutions are truly beneficial from their diverse perspectives. Even when the institutions are mutually beneficial, members can only be confident of that fact when discourse permits, rather than suppresses, voicing idiosyncratic concerns.<sup>164</sup>

These considerations of public discourse highlight an important distinction. We need publicly accessible reasons — common logics — *within* institutions to coordinate judgments about behavior, while discourse *about* institutions remains open to facilitate those institutions being, and being known to be, mutually beneficial, and so drawing on idiosyncratic logics.<sup>165</sup> We expect, for instance, that the traffic laws create a set of shared reasons for behavior, like the need to stop at red lights and stay below the speed limit. Individuals’ reasons for endorsing the traffic

---

<sup>164</sup> Gaus (1996, pp. 166–8) argues that freedom of speech is a necessary condition for confidence that out institutions are interpersonally justified. Cf. Muldoon (forthcoming) for a related Millian argument. Suppression of criticism can sustain institutions that are in fact extremely unpopular through “preference falsification” and “pluralistic ignorance” (Bicchieri, 2006, Chapter 5; Kuran, 1997).

<sup>165</sup> Rawls (1955) highlights the potential separation between justifying a practice and justifying within a practice, and I am adding that justification within a practice may be shared though justification of the practice is diverse.

laws, however, may vary as much as their destinations, and almost certainly do not include any institution-independent valuing of stopping at red lights. Drivers need a common understanding of what the traffic laws are, though they do not need to share destinations or criteria for judging the quality of those laws. In the same way, people may have diverse, even incompatible, reasons for supporting a constitutional separation of church and state. Secular citizens may want the state to not promote religion because of its falsity, while many religious members worry about the corruption of the faith by political involvement.<sup>166</sup> They have a different basis for the constitutional provision, yet within the constitutional order each can apply the shared principle of separation to rule out as unconstitutional certain policies. In more general terms, we would expect individuals to appeal to the shared reasons provided by the institutions for disputes within the institution's domain, as within the courts, though they may appeal to deeply diverse reasons when discussing how the institutions ought to be. When judging the quality of the institutions, or advocating for reform of the constitutional reform or policy within the bounds of constitutionality, we must expect people to apply diverse criteria.<sup>167</sup> The institutions provide a shared content, but do not require a shared basis.<sup>168</sup>

---

<sup>166</sup> This example is also discussed in Gaus and Van Schoelandt (forthcoming). More generally, Baier (1989, p. 775) argues that we can have social unity through a “constitutional consensus” without “consensus on a conception of justice.” Gaus (2012b, sec. 4.4.ii, 2014b) and Muldoon (2014) illuminate the possibility of convergence on institutions from different ontologies.

<sup>167</sup> Knight and Johnson (2011, p. 50), in defending democratic second-order institutions, write that “institutional choice will remain unavoidably political[,]” i.e., contentious, for “[s]ocial actors cannot reasonably expect a neutral, apolitical metric that might be applied to such decision[,]” and the fact of pluralism “suggests that there will always be space for legitimate disagreement about such matters.”

<sup>168</sup> Benjamin Barber writes: “political judgment is not the application of abstract, independent standards to political

## V. What is to be Done?

This paper has defended the coordination requirement for feasible conceptions of justice and sought to show the way that institutions can play a pivotal role in coordinating expectations. I have further argued that Rawls's proposal for a supra-institutional conception of justice must be rejected because it is unable to serve its function, at least insofar as no apparatus has yet been proposed that would effectively coordinate the expectations of the diverse people using the conception. That is, while institutions provide us principles for assessing actions within those institutions, we do not have a way to coordinate our judgments about society's most basic institutions. The problems raised were sufficiently general to make the project of effective reform at the level of supra-institutional principles discouraging. The problem cannot be solved by better principles. Thus we come to the fundamental point: a conception of justice is simply not fit to perform the function Rawls assigns it of adjudicating our competing claims on our institutions. We adjudicate claims through institutions, and cannot escape to a shared supra-institutional perspective.

Given the present analysis, it seems that if supra-institutional principles of justice cannot coordinate our expectations, we may lack adequate means to publicly vindicate our basic structure against the many potential reforms and alternative structures we could have adopted. Coordination comes at the institutional level, not in the abstract discourse of principles of justice. Competing claims about society's basic structure are an enduring aspect of the reasonable pluralism characterizing liberal societies. Of course, individuals will have judgments about how

---

actuality... [but rather] the forging of common actuality in the absence of abstract, independent standards.”

(1989, p. 209; as quoted by D'Agostino, 1996, p. 4)

society should be structured, but these private judgments cannot be thought to have the social standing of a basic social charter.<sup>169</sup> Each of these perspectives carries its own means for gaining critical leverage on the institutions, though no one wields an Archimedean lever.<sup>170</sup> In the end, functionalists evaluate institutions in many of the same ways Rawls (1999f, pp. 306–7) said we would evaluate a conception of justice. We seek institutions “that all can live with” and that are reasonable “given our history and the traditions embedded in our public life....” An overlapping consensus shows when institutions are vindicated from each of our idiosyncratic perspectives, though not from any single social perspective (1999h, 2005, pp. 386–7; cf. Weithman, 2010, p. 307).<sup>171</sup> Vindicated, that is, as good enough, or sufficiently convergent, to constitute terms of

---

<sup>169</sup> This is a central theme in the social contract tradition. For discussion, see Van Schoelandt and Gaus (forthcoming).

<sup>170</sup> It may be that some supra-institutional principles can develop setting partial orderings, particularly by ruling out some institutions. Though I lack the space to argue the point here, these principles would tend to coordinate members only regarding paradigm cases, and leave most political disputes open to idiosyncratic concerns. The potential, limits, and dynamic emergence of such principles are important topics for further research, and I thank Fred D’Agostino, Jerry Gaus, R. J. Leland, Ryan Muldoon and Danny Shahar for extensive discussion of these issues.

<sup>171</sup> Similar criteria for institutions are used by James M. Buchanan (1999a, p. 196, 1999b, p. 467), Gaus (2011a, Chapter V), Gauthier (1995, p. 42), Hadfield and Weingast (2012, p. 487, 2013a, p. 10f), Hayek (1978a, pp. 99, 105, 121; cf. Mack, 2006), Vallier (2014, Chapter 4), and Van Schoelandt (2015). As per Vanderschraaf’s (2013, p. 149) model of a governance convention, an agent’s endorsement may depend upon what institutions she believes would arise were the current ones to fall. To use D’Agostino’s (1996, p. 30) distinction, these are forms of convergence, rather than consensus, justification (Vallier, 2011; cf. Vallier & D’Agostino, 1996, sec. 2.3). The perceived distribution of benefits and historic legitimacy may be relevant to the convergence (D’Agostino, 2013, p. 147ff).

cooperation warranting our voluntary participation, though probably not as optimal. People will apply diverse criteria for assessing the quality of the basic institutions, and so will push for competing reforms to those institutions.

These diverse perspectives should not be confused with Rawls's (1997, p. 615) thought that there will be the competing conceptions of justice. In holding that competing conceptions are to be expected, Rawls gives up on the feasibility of a well-ordered society (cf. Freeman, 2006, pp. 255–6). This is a fundamental denial of the project at the heart of decades of Rawls's work. Whatever projects remain, there is no longer a shared conception of justice to adjudicate disputes about basic social institutions. It is not at all clear how members of such a society could even hold conceptions of justice in the sense Rawls had developed, for each knows that her conception cannot feasibly perform the public adjudicative task. What members have, instead, are merely private preferences about institutions based on idiosyncratic standards. Many of the role-based restrictions on a conception of justice will not apply to these standards, as they do not apply to private conceptions of the good, since they do not serve that public role.<sup>172</sup> We should expect, then, that members will appeal to their full range of values, including egoistic and altruistic consequentialist considerations, as well as deontological standards.<sup>173</sup>

In closing, let me say that to understand that place of principles of justice we must heed a warning from Friedrich Nietzsche (1976, p. 481). He called it dangerous to place the “highest

---

<sup>172</sup> Of course, members may have alternative reasons for similarly constraining their private standards.

<sup>173</sup> Deontological values applied to institutions will likely include conceptions of fairness in procedures and representation, as well as rights the individual wants the institution to respect. Also relevant will be the degree to which the institution may require the agent to violate her perceived duties, with duty violation understood as a cost but not absolute constraint (Benn, 1988, Chapter 3).

concepts” not at the end, but instead “in the beginning, *as* the beginning.” Such a placement is a philosopher’s “way of showing reverence,” expressing a commitment that, “the higher *may* not grow out of the lower, may not have grown at all.” Nietzsche of course had in mind the concept of God, but much the same can be said of justice. Many philosophers conceive of justice as having an independent, eternal existence and as exerting regulative power over our institutions and our selves. Instead, we should see that our interactions and institutions come first, and from them emerge ideas and principles of justice that must always be understood through those institutions. Our principles of justice grow out of institutions structuring our social lives.

## Markets, Community, and Pluralism<sup>174</sup>

Many critics find that markets generate unacceptable inequalities. Others claim that markets in certain goods, such as kidneys or votes, corrode valuable relationships.<sup>175</sup> In *Why Not Socialism?*, G. A. Cohen presents a provocative and more thoroughgoing rejection of markets, insisting that market competition and market-based inequalities as such undermine conditions of community. According to Cohen, markets rest on base motives and naturally engender community-undermining inequalities. This paper critically assesses Cohen's community-based objection to markets. It shows that his objection presupposes a conception of community that is antithetical to pluralism and it argues that his characterization of the motives of market participants is too simple.

Cohen's *Why Not Socialism?* defends the moral desirability of socialism. Much debate revolves around the potential efficiency, stability and feasibility of socialist institutions. Cohen addresses some of these issues, particularly arguing that human nature is not an insurmountable obstacle and acknowledging that effective socialist institutions present an unsolved – perhaps insoluble – design problem.<sup>176</sup> The main question, however, regards whether there is something desirable about socialist relations or if instead we should be content with markets. It is an important insight that socialism could present an attractive ideal even if we do not currently

---

<sup>174</sup> A version of this paper is published as Van Schoelandt (2014).

<sup>175</sup> E.g., Elizabeth Anderson (1995, Chapter 7); Debra Satz (2010); Michael J. Sandel (2013). For a recent reply, see Brennan and Jaworski (2015).

<sup>176</sup> G. A. Cohen (2009, p. 58). For recent criticism, see Brennan (2014).

know how to bring them about.<sup>177</sup> Cohen defends a socialist equality of opportunity principle requiring correction of all inequalities of advantage for which the agents cannot be held responsible.<sup>178</sup> The key insight, however, comes from Cohen's illuminating discussion of how community ideals provide further normative considerations for assessing market and socialist society.

## **I. The Community Principle**

While this paper offers reasons to reject Cohen's community-based objection to markets, we should first appreciate the objection's theoretic importance. As mentioned, the objection is comprehensive, calling into question all markets, not just markets in particular goods. But, more importantly, the objection targets market inequalities that are justified by plausible egalitarian principles.<sup>179</sup> Cohen's objection thus speaks against not only laissez-faire market systems, but also luck egalitarian schemes that rely on markets as well as market socialist arrangements.<sup>180</sup> Beyond the justice of equal opportunity, Cohen argues for more radical anti-market conclusions through an additional “community principle” and a demand for “communal reciprocity” that contrast with repugnant market motivations.

---

<sup>177</sup> Representing contrasting views of the relevance of feasibility, see G. A. Cohen (2008, Chapter 6) and Gerald Gaus (2015).

<sup>178</sup> Cohen (2009, p. 18).

<sup>179</sup> Cohen (2009, p. 37). Cohen's preferred principle of opportunity allows inequalities arising from responsible choices, such to enter in market interactions when there is a suitable alternative (p. 32).

<sup>180</sup> Cohen (2009, pp. 69–70, 75). Cohen accepts markets as necessary until efficient socialist institutions are designed (pp. 75-7), but “[e]very market, even a socialist market, is a system of predation.” (p. 82).

Cohen doesn't provide a canonical statement of the community principle, but he makes several claims providing its content. The community principle constrains inequalities because “community is put under strain when large inequalities obtain,”<sup>181</sup> and we “cannot enjoy full community, you and I, if you make, and keep, say, ten times as much money as I do, because my life will then labor under challenges that you will never face, challenges that you could help me to cope with, but do not, because you keep your money.”<sup>182</sup> Similarly, the community principle is said to be necessary for society to realize the attractive relations that Cohen claims obtain in a good camping trip, and calls for a mode of caring “that curbs some of the inequalities that result from socialist equality of opportunity.”<sup>183</sup>

Cohen emphasizes that the community principle forbids inequalities permitted by socialist equality of opportunity and holds that the caring aspects of community require eliminating differences in outcome (in terms of welfare or other relevant “currency”). That is, community effectively requires that if I see that things are going better for me than they are for you, I should act to improve things for you even at a cost to me.<sup>184</sup> This requirement shows that community reciprocity is constituted by serving each other because of the need of the others rather than what can be gained for oneself by serving. This concern echos Karl Marx's view that

---

<sup>181</sup> Cohen (2009, p. 34).

<sup>182</sup> Cohen (2009, p. 35).

<sup>183</sup> Cohen (2009, pp. 34–5 cf. chapter 1 for the camping trip example).

<sup>184</sup> To have implications for the level of inequality permitted, the principle must specifically require the better off helping the worse off. If community simply involves people caring about and for each other in terms of satisfying each other's needs, no systematic equalizing or socialist implications follow. Though, our ordinary community often involve people helping each other without regard for their relative position.

capitalism causes “estrangement of man from man.”<sup>185</sup> Marx charged that in markets each person “is active as a *private individual*, regards other men as means, debases himself to a means,”<sup>186</sup> while a better world is one of free association for the social realization of productive abilities.<sup>187</sup>

Even if mutual concern within a community tends to generate some equalizing tendency, equality does not define community. Instead, inequalities separate people, make them face different challenges, and perhaps leave them in different social worlds unable to identify with each other's experiences. Cohen offers the example of a car owner having to ride the bus one day: “I can reasonably complain about [the day I have to ride the bus] to a fellow car-driver, but not to you [who must always ride the bus due to poverty]. ... There is a lack of community between us of just the sort that naturally obtains between me and the fellow car-driver.”<sup>188</sup> Fellow drivers, rich and poor (e.g., taxi drivers) alike, have community in shared experiences and identifying with common problems. The poor, however, systematically face different burdens than the rich.<sup>189</sup> Since inequality is not the central issue, assessing Cohen's community principle requires examining these claims about shared problems.

In the bus example, the problem is not that the driver has greater wealth or well-being. The problem was brought out by the fact that the car-driver cannot reasonably complain about

---

<sup>185</sup> Karl Marx (2011, p. 77 italics removed). This passage seems to refer to the estrangement of workers from capitalists, but a broader theme in Marx is that the capitalism estranges people from each other.

<sup>186</sup> Marx (1992b, p. 220).

<sup>187</sup> Aspects of this can be seen in Marx (2011, pp. 76–8), Marx (1992a, p. 173), and Marx and Friedrich Engels (2011, p. 53).

<sup>188</sup> Cohen (2009, p. 36).

<sup>189</sup> I believe that Cohen's criticism relies only on the fact that the poor face *different* burdens, though my arguments below stand even if Cohen means the more contentious empirical claim that the poor face *greater* burdens.

the bus to the bus-riders, though he could reasonably complain about the bus to his fellow car-drivers. The underlying issue regards what we could expect others to empathize or identify with. These differences in conditions allow community among the drivers, but block it between driver and rider. Cohen refers to the well off, even when they are responsible for their relative advantage, as being cut off “from our common life.”<sup>190</sup> While equality of opportunity regards how well your life goes and whether that can be attributed to you, community is about the sort of bonds we have and whether we can really have a common life or instead see each other as strangers, outsiders or even threats. Beyond transportation, Cohen says, wealth inequalities generate “widely different powers to care for ourselves, to protect and care for offspring, to avoid danger, and so on.”<sup>191</sup>

## II. Pluralism

It seems hard to deny that such differences in life experience strain community – at least some sense of community. But consider a case parallel to Cohen's bus example. I, being someone who does not keep kosher, regularly eat bacon, while you keep strictly kosher so never eat bacon. One day my favorite restaurant is out of bacon. It would seem we should say “I can reasonably complain about it to a fellow bacon-eater, but not to you. There is a lack of community between us of just the sort that naturally obtains between me and the fellow bacon-eaters.”<sup>192</sup> Or, mathematics comes fairly naturally to you, but not to me. Or, though we both

---

<sup>190</sup> Cohen (2009, p. 38).

<sup>191</sup> Cohen (2009, p. 36).

<sup>192</sup> Compare to Cohen (2009, p. 36).

work in demanding and onerous tasks, mine is data-entry while yours is manual labor. In all of these cases, we may have trouble identifying with each other's problems, and situations may arise that bring out the lack of communal bonds because I cannot reasonably complain to you, but can complain to others more like me. Differences that prevent identification with each other's problems are ubiquitous, at least in non-homogeneous societies.<sup>193</sup>

Cohen may reply that we are able to identify with another's problems in more general terms, such as a restaurant being out of your favorite dish, rather than bacon *per se*. But, we should expect people to identify just as well across cases grounded in economic inequalities. Cohen might argue that the point of community is not that we share every detail of a common life, like eating bacon, but instead that we overall have common bonds and exposure to common problems. It is not about the bus ride *per se*, but that the lives of the rich and poor are filled with such cases. This response will not work, however, for non-economic differences are similar. Dietary differences often arise from cultural and religious differences that pervade our lives. Occupations create different social networks and risks. Differences in beliefs, values and activities pervade our lives.

Cohen holds that “certain inequalities that cannot be forbidden in the name of socialist equality of opportunity should nevertheless be forbidden, in the name of community.”<sup>194</sup> For example, the community principle “condemns” the running of a voluntary lottery satisfying socialist equality of opportunity and displaying no injustice, if that lottery produces the kind of

---

<sup>193</sup> I here grant that Cohen is correct in his bus example, though we in fact have much more expansive sympathetic abilities.

<sup>194</sup> Cohen (2009, p. 37).

inequality that interferes with community.<sup>195</sup> Preventing differences in burdens created by pluralism would require restricting the liberty to live according to different religious or cultural traditions. If this notion of community condemns markets, it would also support these onerous personal restrictions, and thus justify a very unattractive restriction of opportunities. Even if Cohen is not endorsing forbidding economic inequalities through coercion, we should not be condemning general diversity in any strong sense.<sup>196</sup> Cohen gives no indication of why condemnation of such economic activities, but not onerous prohibitions on other sources of diversity, would be justified.

Perhaps Cohen thinks all of these community-straining differences are *prima facie* objectionable, but many of them are permitted because it would be unjust to restrict them. For instance, since any plausible conception of justice requires freedom of religion, we cannot condemn religious diversity, even when such diversity reduces community. If the protections of justice encompass considerable domains of personal liberty, but do not protect economic activities, this could suitably justify the distinction.

An appeal to justice will not distinguish economic from other community-straining differences. Cohen himself considers that it may be “an *injustice* to forbid the transactions that

---

<sup>195</sup> Cohen (2009, p. 38). More precisely, Cohen writes that “the ideal of community” does the condemning, and does not elaborate what the difference is between the ideal and principle of community.

<sup>196</sup> In other works, Cohen promotes an “egalitarian ethos” in place of legal restrictions. My argument here also applies, *mutatis mutandis*, to such an ethos. At the weakest, Cohen could endorse merely regretting or feeling disappointed about community-straining diversity. While it may be acceptable to regret the interpersonal distance arising from religious and cultural diversity, such disappointment is not damning of markets. For further discussion of an ethos of justice, see Neufeld and Van Schoelandt (2014).

generate [community straining] inequalities[.]”<sup>197</sup> This conflict between the community principle and justice arises because the inequalities in question, by stipulation, are compatible with equal opportunity (which Cohen takes roughly to characterize justice). Though equality may seem the focus, there is no reason to equalize opportunities unless opportunities themselves matter. Equal opportunity is important because we want people to have opportunities, so eliminating an opportunity, even when eliminated for all, is a cost to justice.<sup>198</sup> Cohen's community-based economic restrictions, like other restrictions on diversity, are thus unjust.

We see from the above that Cohen's community principle, given his understanding of community in terms of shared experience, unduly condemns common forms of diversity. Preventing community-straining diversity would objectionably homogenize people and prevent them from engaging in diverse pursuits, whether in the division of labor or separate cultures or belief systems. A proper conception of community, however, will not require normalizing the psychologies of diverse individuals, preventing them from having diverse religions, or arranging society to eliminate or minimize difference as such. Community, being compatible with diverse experiences and ways of life, is thus compatible with markets.

---

<sup>197</sup> Cohen (2009, p. 37). Cohen himself is non-committal about whether or not such restrictions are unjust.

<sup>198</sup> See Thomas Christiano and Will Braynen (2008).

### III. Repugnant Motives

Cohen argues that market inequalities, even just ones, undermine the shared experiences of valuable community. I have replied that other kinds of difference in our pluralistic society do as well. If we condemn market inequality, we must condemn these differences too. However, Cohen claims that markets, unlike other sources of community-impairing diversity, rely on the repugnant, egoistic, motives of greed and fear.<sup>199</sup> Other forms of pluralism sometimes also involve greed and fear, as when someone's religious practice is strictly to obtain divine rewards and avoid divine retribution. Cohen, however, thinks markets are particularly nurturing to repugnant, anti-communal, motives. We can agree with Cohen that even if markets efficiently produce public benefit, they remain problematic if they do so though private vice.<sup>200</sup> These motives contrast with his preferred “[c]ommunal reciprocity... according to which I serve you not because of what I can get in return by doing so but because you need or want my service, and you, for the same reason, serve me.”<sup>201</sup> This is, according to Cohen, “commitment to one’s fellow human beings and a desire to serve them while being served *by* them...”<sup>202</sup> Cohen rightly presents community as a form of reciprocity, rather than selflessness or absorption into an organic whole.<sup>203</sup> Community reciprocity, then, must permit some level of self-regard.

---

<sup>199</sup> Cohen (2009, pp. 58, cf. pp. 40, 76–7) holds that market societies nourish, promote and bring to prominence greed and fear, and “work on the basis of the development, and, indeed, the hypertrophy, of selfishness....”

<sup>200</sup> Cohen (2009, pp. 76–9).

<sup>201</sup> Cohen (2009, pp. 39, cf. pp. 34–5).

<sup>202</sup> Cohen (2009, p. 39).

<sup>203</sup> We can imagine people truly doing everything for the collective benefit, but the collective we thereby imagine might be outside the realm of recognizable realistic human community. *Star Trek's* Borg, for example, are a hive, but hives are not communities realizing reciprocity.

Furthermore, Cohen is sensitive to the economic necessity of “information generation and processing with respect to what should be produced”<sup>204</sup> that markets accomplish, but does not appreciate that generating relevant information about preferences requires self-regarding motives. Any realistic alternative to markets gains this information through self-regarding activity, be it purchases in a pseudo-market, submission of preferences to a central planner, in democratic deliberation, or even on idyllic camping trips, though such discussions may take place around the fire between rounds of “Kumbaya.”

A repugnant motives charge cannot focus on mere self-regard, but excessive egoism or greed. Cohen's preferred other-regard, however, is not the only alternative to greed. Exchange often facilitates ends that are not merely non-repugnant, but even virtuous.<sup>205</sup> Moreover, people have motivational constraints.<sup>206</sup> Social norms are a powerful source of motivation, as many radical socialists know.<sup>207</sup> These motives are important in markets, for markets function when people do not even consider some actions as options, like killing the competition. Market agents produce and exchange from diverse and complex motivations, including religious obligation, tradition, respect, and a sense of dignity.

The motives of market participants are empirical issues. Interdisciplinary, cross cultural research into the effects of markets on motivation lead some researchers to conclude that the “notion that the market economy makes people greedy, selfish, and amoral is simply

---

<sup>204</sup> Cohen (2009, p. 62). Cf. F. A. Hayek (1945).

<sup>205</sup> Hillel Steiner (2014).

<sup>206</sup> See David Schmitz (1995, p. 32) and Elinor Ostrom (2005, p. 139ff.). Cf. Stanley I. Benn (1988, Chapter 3) on trade-offs among commitment-based constraints.

<sup>207</sup> E.g. Joseph Carens (1981) and Paul Gomberg (2006, p. 157). On social norms, see Cristina Bicchieri (2006).

fallacious.”<sup>208</sup> Honesty, trust, reliability, fairness, and altruistic norm enforcement are ubiquitous and pre-supposed in well-functioning markets. Moreover, fairness and trust are particularly associated with market engagement rather than being human universals. If market agents lack the more demanding solidarity that Cohen endorses, this does not appear to be a problem of markets. We do not see the sort of altruism Cohen demands even in small scale, egalitarian, hunter-gatherer societies. Given the evidence, Cohen may lament human nature, but not condemn markets. Likewise, while we should condemn the greedy people in market societies, as in any society, it does not seem that markets specifically are to blame.

We learn from Cohen's discussion that community is a potential source of normative demand beyond justice. It is important that we be able to cooperate without encouraging or depending upon vicious motives. We must reject Cohen's community principle and requirement for communal reciprocity, but we should not give up on community. The notion of community appropriate for large and diverse societies cannot rest on shared experiences. Respecting people and diversity means that community cannot demand uniformity in ways of living or comprehensive conceptions of the good. Theorists must continue to explore how we may have valuable community within a diverse society.

---

<sup>208</sup> Herbert Gintis (2012). Cf. Paul Zak (2008), Joseph Henrich et al. (2004), and Virgil Storr (2009, 2010, 2012).

## References

- Abbey, R. (2007). Back Toward a Comprehensive Liberalism? Justice as Fairness, Gender, and Families. *Political Theory*, 35(1), 5–28.
- Anderson, E. (1995). *Value in Ethics and Economics* (Reprint edition). Cambridge, Mass.: Harvard University Press.
- Anderson, E. (2010). The Fundamental Disagreement between Luck Egalitarians and Relational Egalitarians. *Canadian Journal of Philosophy*, 40(sup1), 1–23.
- Aristotle. (2004). *Rhetoric*. (W. R. Roberts, Trans.) (Dover Thrift Eds edition). Mineola, N.Y.: Dover Publications.
- Arneson, R. J. (1990). Primary Goods Reconsidered. *Noûs*, 24(3), 429.
- Baier, K. (1958). *The Moral Point of View: A Rational Basis of Ethics*. Ithaca: Cornell University Press.
- Baier, K. (1989). Justice and the Aims of Political Philosophy. *Ethics*, 99(4), 771–790.
- Barber, B. (1989). *The Conquest of Politics: Liberal Philosophy in Democratic Times*. Princeton, N.J.: Princeton University Press.
- Baumol, W. J. (1990). Entrepreneurship: Productive, Unproductive, and Destructive. *Journal of Political Economy*, 98(5), 893–921.
- Benn, S. I. (1988). *A Theory of Freedom*. Cambridge University Press.
- Besson, S. (2005). *The Morality of Conflict: Reasonable Disagreement and the Law*. Portland, OR: Hart Pub.
- Bicchieri, C. (2006). *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press.
- Blair, R. (1995). A Cognitive Developmental Approach to Morality: Investigating the

- Psychopath. *Cognition*, 57, 1–29.
- Boettke, P. J., & Coyne, C. J. (2003). Entrepreneurship and Development: Cause or Consequence? *Advances in Austrian Economics*, 6, 67–87.
- Bohman, J., & Richardson, H. S. (2009). Liberalism, Deliberative Democracy, and “Reasons That All Can Accept.” *Journal of Political Philosophy*, 17(3), 253–274.
- Bosanquet, B. (2001a). The Antithesis between Individualism and Socialism Philosophically Considered. In G. F. Gaus & W. Sweet (Eds.), *The Philosophical Theory of the State and Related Essays* (pp. 324–346). South Bend, Ind.: St. Augustine’s Press.
- Bosanquet, B. (2001b). *The Philosophical Theory of the State and Related Essays*. (G. F. Gaus & W. Sweet, Eds.). South Bend, Ind.: St. Augustine’s Press.
- Bowles, S., & Gintis, H. (2011). *A Cooperative Species: Human Reciprocity and Its Evolution*. Princeton University Press.
- Boyd, R., Gintis, H., Bowles, S., & Richerson, P. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*.
- Brennan, G., Baurmann, M., Goodin, R. E., & Southwood, N. (2010). Introduction. In M. Baurmann, G. Brennan, R. E. Goodin, & N. Southwood (Eds.), *Norms and Values: The Role of Social Norms as Instruments of Value Realisation* (pp. 66–74). Baden-Baden: Nomos.
- Brennan, G., Eriksson, L., Goodin, R. E., & Southwood, N. (2013). *Explaining Norms*. New York: Oxford University Press.
- Brennan, J. (2014). *Why Not Capitalism?* (1st edition). New York ; London: Routledge.
- Brennan, J., & Jaworski, P. (2015). *Markets without Limits*. Routledge.
- Buchanan, A. (1990). Justice as Reciprocity versus Subject-Centered Justice. *Philosophy and*

- Public Affairs*, 19(3), 227.
- Buchanan, A. (2004). *Justice, Legitimacy, and Self-Determination: Moral Foundations for International Law (Oxford Political Theory)*. Oxford University Press, USA.
- Buchanan, J. M. (1999a). Positive Economics, Welfare Economics, and Political Economy. In *The Collected Works of James M. Buchanan, Volume 1: The Logical Foundations of Constitutional Liberty* (pp. 191–209). Indianapolis, Ind.: Liberty Fund.
- Buchanan, J. M. (1999b). The Constitution of Economic Policy. In *The Collected Works of James M. Buchanan, Volume 1: The Logical Foundations of Constitutional Liberty* (pp. 455–468). Indianapolis, Ind.: Liberty Fund.
- Buchanan, J. M., & Brennan, G. (2000). *The Reason of Rules* (Vol. 10). Liberty Fund Inc.
- Calhoun, C. (2003). Cognitive Emotions? In R. C. Solomon (Ed.), *What is an Emotion?: Classic and Contemporary Readings* (pp. 236–247). New York: Oxford University Press.
- Carens, J. (1981). *Equality, Moral Incentives, and the Market: An Essay in Utopian Politico-Economic Theory*. University of Chicago Press.
- Christiano, T., & Braynen, W. (2008). Inequality, Injustice and Leveling Down. *Ratio*, 21(4), 392–420.
- Chwe, M. S.-Y. (2013). *Rational Ritual: Culture, Coordination, and Common Knowledge*.
- Cohen, G. A. (2008). *Rescuing Justice and Equality*. Harvard University Press.
- Cohen, G. A. (2009). *Why Not Socialism?*. Princeton University Press.
- D'Agostino, F. (1996). *Free Public Reason: Making It Up As We Go*. Oxford University Press, USA.
- D'Agostino, F. (2003). *Incommensurability and commensuration: the common denominator*. Aldershot: Ashgate.

- D'Agostino, F. (2010). *Naturalizing epistemology: Thomas Kuhn and the "essential tension."* Houndmills, Basingstoke, Hampshire [England]; New York: Palgrave Macmillan.
- D'Agostino, F. (2012). Disciplinarity and the Growth of Knowledge. *Social Epistemology*, 26(3-4), 331–350.
- D'Agostino, F. (2013). The Orders of Public Reason. *Analytic Philosophy*, 54(1), 129–155.
- Darwall, S. (2006). *The Second-Person Standpoint: Morality, Respect, and Accountability*. Harvard University Press.
- DiQuattro, A. (1986). Rawls versus Hayek. *Political Theory*, 14(2), 307–310.
- Dreben, B. (2003). On Rawls and Political Liberalism. In S. Freeman (Ed.), *The Cambridge Companion to Rawls* (pp. 316–346). New York: Cambridge University Press.
- Dworkin, R. (1976). The Original Position. In N. Daniels (Ed.), *Reading Rawls: Critical Studies on Rawls' "A Theory of Justice."* Stanford University Press.
- Eberle, C. J. (2002). *Religious Conviction in Liberal Politics*. New York: Cambridge University Press.
- Enoch, D. (2013). *Against Public Reason*. Presented at the Oxford Studies in Political Philosophy Workshop, Tucson, Arizona.
- Epicurus. (1994). The Principle Doctrines. In B. Inwood & L. P. Gerson (Trans.), *The Epicurus Reader: Selected Writings and Testimonia*. Hackett Publishing Company.
- Estlund, D. (2008). *Democratic Authority: A Philosophical Framework*. Princeton, NJ: Princeton University Press.
- Estlund, D. (2011). Human Nature and the Limits (If Any) of Political Philosophy. *Philosophy & Public Affairs*, 39(3), 207–237.
- Faraci, D., & Shoemaker, D. (2010). Insanity, Deep Selves, and Moral Responsibility: The Case

- of JoJo. *Review of Philosophy and Psychology*, 1(3), 319–332.
- Faraci, D., & Shoemaker, D. (2014). Huck vs. Jojo: Moral Ignorance and the (A)Symmetry of Praise and Blame. In T. Lombrozo, J. Knobe, & S. Nichols (Eds.), *Oxford Studies in Experimental Philosophy* (Vol. 1). Oxford University Press.
- Freeman, S. (2006). *Justice and the Social Contract: Essays on Rawlsian Political Philosophy*. Oxford ; New York: Oxford University Press.
- Galston, W. A. (1991). *Liberal Purposes: Goods, Virtues, and Diversity in the Liberal State*. New York: Cambridge University Press.
- Gaus, G. (1990). *Value and Justification: The Foundations of Liberal Theory*. New York: Cambridge University Press.
- Gaus, G. (1996). *Justificatory Liberalism: An Essay on Epistemology and Political Theory*. New York: Oxford University Press.
- Gaus, G. (2003). Liberal Neutrality: A Compelling and Radical Principle. In S. Wall & G. Klosko (Eds.), *Perfectionism and Neutrality: Essays in Liberal Theory* (pp. 137–66). Lanham, Md.: Rowman & Littlefield Publishers.
- Gaus, G. (2010). Coercion, Ownership, and the Redistributive State: Justificatory Liberalism’s Classical Tilt. *Social Philosophy and Policy*, 27(01), 233–.
- Gaus, G. (2011a). *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*. New York: Cambridge University Press.
- Gaus, G. (2011b). The Property Equilibrium in a Liberal Social Order (or How to Correct Our Moral Vision). *Social Philosophy and Policy*, 28(02), 74–101.
- Gaus, G. (2012a). Explanation, justification, and emergent properties; an essay on Nozickian metatheory. In R. M. Bader & J. Meadowcroft (Eds.), *The Cambridge Companion to*

- Nozick's Anarchy, State, and Utopia*. Cambridge: Cambridge University Press.
- Gaus, G. (2012b). Social Contract and Social Choice. *Rutgers Law Journal*, 43, 243–76.
- Gaus, G. (2013, November). *The Egalitarian Species*. Presented at the Equality and Public Policy, Athens, Ohio.
- Gaus, G. (2014a). The Good, the Bad, and the Ugly: Three Agent-type Challenges to The Order of Public Reason. *Philosophical Studies*, 170(3), 563–577.
- Gaus, G. (2014b, June). *Is Public Reason a Normalization Project?*. Presented at the New Directions in Public Reason, Birmingham, AL.
- Gaus, G. (2015). *The Tyranny of the Ideal: A Defense of Diversity and the Open Society*. Princeton University Press.
- Gaus, G. (n.d.). The Evolution, Evaluation, and Reform of Social Morality. In S. J. Peart & D. M. Levy (Eds.), *F. A. Hayek and the Modern Economy*. Palgrave Macmillan.
- Gaus, G., & Van Schoelandt, C. (forthcoming). “Public Reason.” In *International Encyclopedia of the Social and Behavioral Sciences* (Second). Elsevier.
- Gauthier, D. (1969). *The Logic of Leviathan: The Moral and Political Theory of Thomas Hobbes*. Oxford: Oxford University Press.
- Gauthier, D. (1986). *Morals by Agreement*. Clarendon Press, Oxford.
- Gauthier, D. (1995). Public Reason. *Social Philosophy & Policy*, 12, 19–42.
- Gewirth, A. (1980). *Reason and Morality*. University Of Chicago Press.
- Gintis, H. (2012). Response: Giving Economists Their Due. *Boston Review*, (May/June). Retrieved from [http://www.bostonreview.net/BR37.3/ndf\\_herbert\\_gintis\\_markets\\_morals.php](http://www.bostonreview.net/BR37.3/ndf_herbert_gintis_markets_morals.php)
- Gomberg, P. (2006). *How to Make Opportunity Equal: Race and Contributive Justice* (1

- edition). Malden, MA: Wiley-Blackwell.
- Gross, J. (1990, March 31). California Execution Stayed: State Is Appealing to High Court. *The New York Times*. Retrieved from <http://www.nytimes.com/1990/03/31/us/california-execution-stayed-state-is-appealing-to-high-court.html>
- Hadfield, G. K., & Macedo, S. (2012). Rational Reasonableness: Toward a Positive Theory of Public Reason. *Law and Ethics of Human Rights*, 6(1), 7–46.
- Hadfield, G. K., & Weingast, B. R. (2012). What is Law? A Coordination Account of the Characteristics of Legal Order. *Journal of Legal Analysis*, 4(2), 471–514.
- Hadfield, G. K., & Weingast, B. R. (2013a). Law without the State: Legal Attributes and the Coordination of Decentralized Collective Punishment. *Journal of Law and Courts*, 1(1), 3–34.
- Hadfield, G. K., & Weingast, B. R. (2013b). *Microfoundations of the Rule of Law* (SSRN Scholarly Paper No. ID 2342882). Rochester, NY: Social Science Research Network. Retrieved from <http://papers.ssrn.com/abstract=2342882>
- Hadfield, G. K., & Weingast, B. R. (2014). Constitutions as Coordinating Devices. In S. Galiani & I. Sened (Eds.), *Institutions, Property Rights, and Economic Growth: The Legacy of Douglass North* (pp. 121–150). New York: Cambridge University Press.
- Harman, E. (2011). Does Moral Ignorance Exculpate? *Ratio*, 24(4), 443–468.
- Hayek, F. (1945). The Use of Knowledge in Society. *American Economic Review*, XXV(4), 519–530.
- Hayek, F. A. (1978a). *Law, Legislation and Liberty, Volume 1: Rules and Order*. Chicago: University of Chicago Press.
- Hayek, F. A. (1978b). *Law, Legislation and Liberty, Volume 2: The Mirage of Social Justice*.

- Chicago: University of Chicago Press.
- Hayek, F. A. (2007). *The Road to Serfdom*. (B. Caldwell, Ed.) (Vol. 2). University of Chicago Press.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., & Gintis, H. (Eds.). (2004). *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies* (1 edition). Oxford New York: Oxford University Press.
- Hobbes, T. (1994). *Leviathan*. (E. Curley, Ed.). Hackett Publishing Company.
- Hume, D. (2000). *A Treatise of Human Nature*. (D. F. Norton & M. J. Norton, Eds.). New York: Oxford University Press.
- Hurley, E. A., & Macnamara, C. (2011). Beyond Belief: Toward a Theory of the Reactive Attitudes. *Philosophical Papers*, 39(3), 373–399.
- James, A. (2013). *Fairness in practice: a social contract for a global economy*.
- Kant, I. (1999). *Metaphysical Elements of Justice*. (J. Ladd, Trans.) (2nd ed.). Hackett Pub Co Inc.
- Kavka, G. S. (1986). *Hobbesian Moral and Political Theory*. Princeton University Press.
- Kiser, L. L., & Ostrom, E. (1999). The Three Worlds of Action. In M. D. McGinnis (Ed.), *Polycentric Games and Institutions: Readings from the Workshop in Political Theory and Policy Analysis* (pp. 56–88). University of Michigan Press.
- Kitcher, P. (2011). *The Ethical Project*. Harvard University Press.
- Knight, J., & Johnson, J. (2011). *The Priority of Democracy: Political Consequences of Pragmatism*. New York : Princeton N.J.: Princeton University Press.
- Kofman, D. (2012). How Insensitive: Principles, Facts and Normative Grounds in Cohen's

- Critique of Rawls. *Socialist Studies / Études Socialistes*, 8(1), 246–266.
- Korsgaard, C. (2004). Fellow Creatures: Kantian Ethics and Our Duties to Animals. Retrieved from <http://dash.harvard.edu/handle/1/3198692>
- Kukathas, C. (2003). *The Liberal Archipelago: A Theory of Diversity and Freedom*. Oxford: Oxford University Press.
- Kuran, T. (1997). *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Cambridge, MA: Harvard University Press.
- Larmore, C. (1990). Political Liberalism. *Political Theory*, 18(3), 339–360.
- Larmore, C. (2013). What Is Political Philosophy? *Journal of Moral Philosophy*, 10(3), 276–306.
- Leeson, P. T. (2007). An-arrgh-chy: The Law and Economics of Pirate Organization. *Journal of Political Economy*, 115(6), 1049–1094. <http://doi.org/10.1086/526403>
- Leeson, P. T. (2009). *The Invisible Hook: The Hidden Economics of Pirates*. Princeton, N.J.: Princeton University Press.
- Leeson, P. T. (2014). *Anarchy Unbound: Why Self-Governance Works Better Than You Think* (1st edition). Cambridge University Press.
- Leeson, P. T., & Williamson, C. R. (2009). Anarchy and Development: An Application of the Theory of Second Best. *The Law and Development Review*, 2(1), 77–96.
- Leland, R. J. (2014). Political Liberalism and Political Community. In *Justifying Political Liberalism*. Stanford, Calif.: dissertation manuscript.
- Lipsey, R. G., & Lancaster, K. (1956). The General Theory of Second Best. *The Review of Economic Studies*, 24(1), 11–32.
- Lister, A. (2013a). *Public Reason and Political Community*. New York: Bloomsbury.
- Lister, A. (2013b). The “Mirage” of Social Justice: Hayek Against (and For) Rawls. *Critical*

- Review*, 25(3-4), 409–444.
- Locke, J. (1681). *Second Treatise of Civil Government*. In P. Laslett (Ed.), *Two Treatises of Government*. Cambridge: Cambridge University Press.
- Locke, J. (1983). *A Letter Concerning Toleration*. (J. Tully, Ed.). Indianapolis: Hackett Pub. Co.
- Locke, J. (2010). *A Letter Concerning Toleration*. In M. Goldie (Ed.), *A Letter concerning Toleration and Other Writings*. Indianapolis: Liberty Fund. Retrieved from <http://oll.libertyfund.org/title/2375/224886>
- Lomasky, L. E. (1987). *Persons, Rights, and the Moral Community*. New York: Oxford University Press.
- Lumet, S. (1976). *Network*. Drama, MGM.
- Macedo, S. (2000). In Defense of Liberal Public Reason: Are Slavery and Abortion Hard Cases? In R. P. George & C. Wolfe (Eds.), *Natural law and public reason* (pp. 11–50). Washington, D.C.: Georgetown University Press.
- Macedo, S. (2010). Why Public Reason? Citizens' Reasons and the Constitution of the Public Sphere. Retrieved from <http://papers.ssrn.com/abstract=1664085>
- Macedo, S. (2011). *Political Education and Public Reason*. SSRN Scholarly Paper, Rochester, NY. Retrieved from <http://papers.ssrn.com/abstract=1900380>
- Mack, E. (2006). Hayek on justice and the order of actions. *The Cambridge Companion to Hayek*, 259–286.
- Mack, E. (2011). Peter Pan Strikes Back. In J. Kuznicki (Ed.), *For a New Liberalism (Cato Unbound)*. Cato Institute.
- Mackenzie, M. M. (1985). *Plato on Punishment*. Berkeley: University of California Press.
- Macnamara, C. (2013). Reactive Attitudes as Communicative Entities. *Philosophy and*

*Phenomenological Research*, 88(1).

- Macnamara, C. (2015). Blame, Communication and Morally Responsible Agency. In R. Clarke, M. McKenna, & A. Smith (Eds.), *The Nature of Moral Responsibility* (pp. 211–36). New York: Oxford University Press.
- Marx, K. (1977). *Critique of Hegel's "Philosophy Of Right."* (J. O'Malley, Ed., A. Jolin, Trans.). Cambridge England; New York: Cambridge University Press.
- Marx, K. (1979). *A Contribution to the Critique of Political Economy*. (M. Dobb, Ed.). Intl Pub.
- Marx, K. (1992a). *Capital: Volume 1: A Critique of Political Economy*. (E. Mandel, Ed., B. Fowkes, Trans.) (Reprint edition). London ; New York, N.Y: Penguin Classics.
- Marx, K. (1992b). On the Jewish Question. In L. Colletti (Ed.), R. Livingstone & G. Benton (Trans.), *Early Writings* (Reprint edition, pp. 211–42). Harmondsworth; New York: Penguin Classics.
- Marx, K. (2011). *Economic and Philosophic Manuscripts of 1844*. (M. Milligan, Trans.). Wilder Publications.
- Marx, K., & Engels, F. (2011). *The German Ideology, including Theses on Feuerbach* (Paperback edition). Prometheus Books.
- Mason, A. (2012). What Is the Point of Justice? *Utilitas*, 24(04), 525–547.
- McKenna, M. (1998). The Limits of Evil and the Role of Moral Address: A Defense of Strawsonian Compatibilism. *The Journal of Ethics*, 2(2), 123–142.
- McKenna, M. (2008). The Limits of Evil and the Role of Moral Address: A Defense of Strawsonian Compatibilism. In M. McKenna & P. Russell (Eds.), *Free Will and Reactive Attitudes : Perspectives on P.F. Strawson's "Freedom and Resentment* (pp. 201–219). Aldershot: Ashgate.

- McKenna, M. (2011). *Conversation and Responsibility*. New York: Oxford University Press.
- Moehler, M. (2014). The Scope of Instrumental Morality. *Philosophical Studies*, 167, 431–451.
- Morain, D. (1992, April 21). From Birth to Death Row, Violence Surrounded Harris : History: Abused as a child, he became an abuser. He killed a neighbor before the murders of two teen-agers brought him to face the gas chamber. *Los Angeles Times*. Retrieved from [http://articles.latimes.com/1992-04-21/news/mn-608\\_1\\_robert-harris](http://articles.latimes.com/1992-04-21/news/mn-608_1_robert-harris)
- Morris, C. (1991). Moral Standing and Rational-Choice Contractarianism. In P. Vallentyne (Ed.), *Contractarianism and Rational Choice: Essays on David Gauthier's Morals By Agreement* (pp. 76–98). Cambridge: Cambridge University Press.
- Muldoon, R. (forthcoming). Expanding the Justificatory Framework of Mill's Experiments in Living. *Utilitas*.
- Muldoon, R. (2014). Perspective-Dependent Harm. *Working Paper*.
- Muldoon, R., Lisciandra, C., Colyvan, M., Martini, C., Sillari, G., & Sprenger, J. (2014). Disagreement behind the veil of ignorance. *Philosophical Studies*, 170(3), 377–394.
- Nagel, T. (1991). *Equality and Partiality*. New York: Oxford University Press.
- Nelkin, D. K. (2015). Psychopaths, Incurable Racists, and the Faces of Responsibility. *Ethics*, 125(2), 357–390.
- Neufeld, B., & Van Schoelandt, C. (2014). Political Liberalism, Ethos Justice, and Gender Equality. *Law and Philosophy*, 33(1), 75–104.
- Nietzsche, F. W. (1976). Twilight of the Idols. In W. Kaufmann (Trans.), *The portable Nietzsche*. New York: Penguin Books.
- Nunner-Winkler, G., & Sodian, B. (1988). Children's Understanding of Moral Emotions. *Child Development*, 59(5), 1323–1338.

- Nussbaum, M. C. (2001). *Upheavals of Thought: The Intelligence of Emotions*. New York: Cambridge University Press.
- Okin, S. M. (2005). “Forty Acres and a Mule” for Women: Rawls and Feminism. *Politics, Philosophy & Economics*, 4(2), 233–248.
- Ostrom, E. (1999). An Agenda For the Study of Institutions. In M. D. McGinnis (Ed.), *Polycentric Games and Institutions: Readings from the Workshop in Political Theory and Policy Analysis* (pp. 89–113). University of Michigan Press.
- Ostrom, E. (2005). *Understanding Institutional Diversity*. Princeton University Press.
- Ostrom, V. (1999). Polycentricity, Part I. In M. D. McGinnis (Ed.), *Polycentricity and Local Public Economies: Readings from the Workshop in Political Theory and Policy Analysis* (pp. 52–74). Ann Arbor: University of Michigan Press.
- Panzram, C. (2002). *Panzram: A Journal of Murder*. (T. E. Gaddis & J. O. Long, Eds.). Amok Books.
- Pettit, P. (2001). *A Theory of Freedom: From the Psychology to the Politics of Agency*. New York: Oxford University Press.
- Pietrasik, W. S., Video by Tom. (2014, February 12). Kony’s child soldiers: “When you kill for the first time, you change.” Retrieved from <http://www.telegraph.co.uk/news/worldnews/africaandindianocean/uganda/10621792/Kony-child-soldiers-When-you-kill-for-the-first-time-you-change.html>
- Plato. (1968). *The Republic Of Plato*. (A. Bloom, Trans.). Basic Books.
- Plato. (2005). *The Laws*. (T. J. Saunders, Trans.). Princeton, N.J.: Penguin Classics.
- Pogge, T. W. (1989). *Realizing Rawls*. Cornell University Press.
- Quong, J. (2005). Disagreement, Asymmetry, and Liberal Legitimacy. *Politics, Philosophy and*

- Economics*, 4(3), 301–330.
- Quong, J. (2010). *Liberalism Without Perfection*. New York: Oxford University Press.
- Quong, J. (2012). Liberalism Without Perfection: Replies to Gaus, Colburn, Chan, and Bocchiola. *Philosophy and Public Issues*, 2(2), 51–79.
- Quong, J. (2014). What is the Point of Public Reason? *Philosophical Studies*, 170(3), 545–553.
- Rawls, J. (1955). Two Concepts of Rules. *The Philosophical Review*, 64(1), 3–32.
- Rawls, J. (1997). The Idea of Public Reason Revisited. In S. Freeman (Ed.), *Collected Papers* (pp. 573–615). Cambridge, MA: Harvard University Press.
- Rawls, J. (1999a). *A Theory of Justice* (Revised). Cambridge, MA: Belknap Press.
- Rawls, J. (1999b). Distributive Justice. In S. Freeman (Ed.), *Collected Papers* (pp. 130–153). Cambridge, MA: Harvard University Press.
- Rawls, J. (1999c). Justice as Fairness. In S. Freeman (Ed.), *Collected Papers* (pp. 47–72). Cambridge, MA: Harvard University Press.
- Rawls, J. (1999d). Justice as Fairness: Political, Not Metaphysical. In S. Freeman (Ed.), *Collected Papers* (pp. 388–414). Cambridge, MA: Harvard University Press.
- Rawls, J. (1999e). Justice as Reciprocity. In S. Freeman (Ed.), *Collected Papers* (pp. 190–224). Cambridge, MA: Harvard University Press.
- Rawls, J. (1999f). Kantian Constructivism in Moral Theory. In S. Freeman (Ed.), *Collected Papers* (pp. 303–358). Cambridge, MA: Harvard University Press.
- Rawls, J. (1999g). Outline of a Decision Procedure for Ethics. In S. Freeman (Ed.), *Collected Papers* (pp. 1–19). Cambridge, MA: Harvard University Press.
- Rawls, J. (1999h). The Idea of an Overlapping Consensus. In S. Freeman (Ed.), *Collected Papers* (pp. 421–448). Cambridge, MA: Harvard University Press.

- Rawls, J. (1999i). The Independence of Moral Theory. In S. Freeman (Ed.), *Collected Papers* (pp. 286–302). Cambridge, MA: Harvard University Press.
- Rawls, J. (1999j). The Sense of Justice. In S. Freeman (Ed.), *Collected Papers* (pp. 96–116). Cambridge, MA: Harvard University Press.
- Rawls, J. (2001a). *Justice as Fairness: A Restatement*. (E. Kelly, Ed.). London; Cambridge, MA: The Belknap Press of Harvard University Press.
- Rawls, J. (2001b). *Justice as Fairness: A Restatement*. (E. Kelly, Ed.). Cambridge: Harvard University Press.
- Rawls, J. (2005). *Political Liberalism*. New York: Columbia University Press.
- Raz, J. (1988). *The Morality of Freedom*. Oxford: Clarendon Press.
- Raz, J. (1998). Disagreement in Politics. *American Journal of Jurisprudence*, 43(1), 25–52.
- Raz, J. (1999). *Practical Reason and Norms*. Oxford University Press.
- Raz, J. (2006). The Problem of Authority: Revisiting the Service Conception. *Minnesota Law Review*, 90, 1003–1044.
- Ridge, M. (1998). Hobbesian Public Reason. *Ethics*, 108(3), 538–568.
- Rose, D. C. (2011). *The Moral Foundation of Economic Behavior*. New York: Oxford University Press, USA.
- Rosen, G. (2002). Culpability and Ignorance. *Proceedings of the Aristotelian Society*, 103(1), 61–84.
- Rosen, G. (2015). The Alethic Conception of Moral Responsibility. In R. Clarke, M. McKenna, & A. M. Smith (Eds.), *The Nature of Moral Responsibility: New Essays* (1 edition, pp. 65–88). New York, NY: Oxford University Press.
- Sandel, M. J. (2013). *What Money Can't Buy: The Moral Limits of Markets* (Reprint edition).

- New York: Farrar, Straus and Giroux.
- Satz, D. (2010). *Why Some Things Should Not Be for Sale: The Moral Limits of Markets*. Oxford University Press, USA.
- Schelling, T. (1960). *The Strategy of Conflict*. Harvard University Press.
- Schmidtz, D. (1995). *Rational Choice and Moral Agency*. Princeton University Press.
- Schmidtz, D. (1996). Justifying the State. In J. T. Sanders & J. Narveson (Eds.), *For and Against the State: New Philosophical Readings*. Rowman & Littlefield Publishers.
- Schmidtz, D. (2008). The Institution of Property. In *Person, Polis, Planet: Essays in Applied Philosophy* (pp. 193–210). Oxford: Oxford University Press.
- Schofield, M. (2012). Injury, Injustice, and the Involuntary in The Laws. In R. Kamtekar (Ed.), *Virtue and Happiness: Essays in Honour of Julia Annas* (pp. 102–114). Oxford University Press. Retrieved from <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199646043.001.0001/acprof-9780199646043-chapter-7>
- Schwab, D., & Ostrom, E. (2008). The Vital Role of Norms and Rules in Maintaining Open Public and Private Economies. In P. J. Zak (Ed.), *Moral Markets: The Critical Role of Values in the Economy* (pp. 204–227). Princeton: Princeton University Press.
- Sen, A. (2009). *The Idea of Justice*. Cambridge, MA: Harvard University Press.
- Sensen, O. (forthcoming). Kant. In G. Gaus & P. Turner (Eds.), *Public Reason in the History of Political Philosophy*. Routledge.
- Shabo, S. (2012). Where Love and Resentment Meet: Strawson's Intrapersonal Defense of Compatibilism. *Philosophical Review*, 121(1), 95–124.
- Shoemaker, D. (2007). Moral Address, Moral Responsibility, and the Boundaries of the Moral

- Community. *Ethics*, 118, 70–108.
- Shoemaker, D. (2011a). Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility. *Ethics*, 121(3), 602–632.
- Shoemaker, D. (2011b). Psychopathy, Responsibility, and the Moral/Conventional Distinction. *The Southern Journal of Philosophy*, 49, 99–124.
- Shoemaker, D. (2013a). On Criminal and Moral Responsibility. In M. Timmons (Ed.), *Oxford Studies in Normative Ethics* (pp. 154–178). Oxford University Press. Retrieved from <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199685905.001.0001/acprof-9780199685905-chapter-8>
- Shoemaker, D. (2013b). Qualities of Will. *Social Philosophy and Policy*, 30(1-2), 95–120.
- Shoemaker, D. (2015). *Responsibility from the Margins*. Oxford University Press.
- Simmons, A. J. (1999). Justification and Legitimacy. *Ethics*, 109(4), 739–771.
- Simmons, A. J. (2010). Ideal and Nonideal Theory. *Philosophy & Public Affairs*, 38(1), 5–36.
- Skarbek, D. (2010). Putting the “Con” into Constitutions: The Economics of Prison Gangs. *Journal of Law, Economics, & Organization*, 26(2), 183–211.  
<http://doi.org/10.2307/40925597>
- Skarbek, D. (2011). Governance and Prison Gangs. *American Political Science Review*, 105(4), 702–716.
- Skarbek, D. (2014). *The Social Order of the Underworld: How Prison Gangs Govern the American Penal System*. Oxford University Press.
- Sleat, M. (2013). Coercing Non-liberal Persons: Considerations on a More Realistic Liberalism. *European Journal of Political Theory*, 12(4), 347–367.
- Smith, A. (2009). *The Theory of Moral Sentiments*. Indianapolis: Liberty Fund Inc.

- Smith, A. M. (2012). Moral Blame and Moral Protest. In D. J. Coates & N. A. Tognazzini (Eds.), *Blame: Its Nature and Norms* (1 edition, pp. 27–48). Oxford ; New York: Oxford University Press.
- Solomon, R. C. (1973). Emotions and Choice. *The Review of Metaphysics*, 27(1), 20–41.
- Southwood, N. (2010). Norms, Laws and Social Authority. In M. Baurmann, G. Brennan, R. E. Goodin, & N. Southwood (Eds.), *Norms and Values: The Role of Social Norms as Instruments of Value Realisation* (pp. 75–91). Baden-Baden: Nomos.
- Steiner, H. (2014). Greed and Fear. *Politics, Philosophy and Economics*, 13(2), 140–150.
- Storr, V. (2009). Why the Market? Markets as Social and Moral Spaces. *Journal of Markets & Morality*, 12(2), 277–296.
- Storr, V. (2010). The Social Construction of the Market. *Society*, 47(3).
- Storr, V. (2012). *Understanding the Culture of Markets* (1 edition). Abingdon, Oxon ; New York: Routledge.
- Strawson, P. F. (1961). Social Morality and Individual Ideal. *Philosophy*, 36(136), 1–17.
- Strawson, P. F. (2003). Freedom and Resentment. In G. Watson (Ed.), *Free Will* (pp. 72–93). Oxford; New York: Oxford University Press.
- Strawson, P. F. (2008). Freedom and Resentment. In M. McKenna & P. Russell (Eds.), *Free Will and Reactive Attitudes : Perspectives on P.F. Strawson's "Freedom and Resentment* (pp. 19–37). Aldershot: Ashgate.
- Stringham, E. P. (2011). Embracing Morals in Economics: The Role of Internal Moral Constraints in a Market. *Journal of Economic Behavior and Organization*, 78(1), 98–109.
- Taylor, R. S. (2011). *Reconstructing Rawls: the Kantian foundations of justice as fairness*. University Park, Pa: Pennsylvania State Univ. Press.

- Thomas, L. (1998). The Grip of Immorality: Child Abuse and Moral Failure". In J. B. Schneewind (Ed.), *Reason, Ethics, and Society: Themes From Kurt Baier, With His Responses* (pp. 144–67). Chicago: Open Court.
- Thrasher, J. (2013). Reconciling Justice and Pleasure in Epicurean Contractarianism. *Ethical Theory and Moral Practice*, 16(2), 423–436.
- Thrasher, J., & Vallier, K. (Forthcoming). The Fragility of Consensus. *European Journal of Philosophy*.
- Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology*, 46(1), 35–57.
- Vallier, K. (forthcoming). On Quong's Sectarian Political Liberalism. *Criminal Law and Philosophy*.
- Vallier, K. (2011). Consensus and Convergence in Public Reason. *Public Affairs Quarterly*, 25(4), 261–279.
- Vallier, K. (2014). *Liberal Politics and Public Faith: Beyond Separation*. New York: Routledge.
- Vallier, K., & D'Agostino, F. (1996). Public Justification. Retrieved from <http://plato.stanford.edu/archives/fall2013/entries/justification-public/>
- Vanderschraaf, P. (forthcoming). *Strategic Justice*. Oxford University Press.
- Vanderschraaf, P. (2006). The Circumstances of Justice. *Politics, Philosophy and Economics*, 5(3), 321–351.
- Vanderschraaf, P. (2008). Common Knowledge. In *Stanford Encyclopedia of Philosophy*.
- Vanderschraaf, P. (2011). Justice as Mutual Advantage and the Vulnerable. *Politics, Philosophy & Economics*, 10(2), 119–147.
- Vanderschraaf, P. (2013). A Governing Convention? *Rationality, Markets and Morals*, 4, 131–

156.

Van Schoelandt, C. (2014). Markets, Community, and Pluralism. *Philosophical Quarterly*, 64(254), 144–151.

Van Schoelandt, C. (2015). Justification, coercion, and the place of public reason. *Philosophical Studies*, 172(4), 1031–1050.

Wallace, R. J. (1998). *Responsibility and the Moral Sentiments* (Reprint edition). Cambridge, Mass.: Harvard University Press.

Wall, S. (1998). *Liberalism, Perfectionism and Restraint*. Cambridge University Press.

Wall, S. (2002). Is Public Justification Self-Defeating? *American Philosophical Quarterly*, 39(4), 385–394.

Wall, S. (2010). On Justificatory Liberalism. *Politics, Philosophy & Economics*, 9(2), 123–149.

Wall, S. (2013). Political Morality and Constitutional Settlements. *Critical Review of International Social and Political Philosophy*, 16(4), 481–499.

Watson, G. (1987). Responsibility and the Limits of Evil: Variations on a Strawsonian Theme. In F. D. Schoeman (Ed.), *Responsibility, Character, and the Emotions: New Essays in Moral Psychology* (pp. 256–86). New York: Cambridge University Press.

Watson, G. (1996). Two Faces of Responsibility. *Philosophical Topics*, 24(2), 227–248.

Watson, G. (2008). Responsibility and the Limits of Evil: Variations on a Strawsonian Theme. In M. McKenna & P. Russell (Eds.), *Free Will and Reactive Attitudes: Perspectives on P.F. Strawson's "Freedom and Resentment"* (pp. 115–148). Aldershot: Ashgate.

Weithman, P. (2002). *Religion and the Obligations of Citizenship*. Cambridge University Press.

Weithman, P. (2010). *Why Political Liberalism?: On John Rawls's Political Turn*. Oxford: Oxford University Press.

- Williams, B. (1981). *Moral Luck*. New York: Cambridge University Press.
- Wolf, S. (1987). Sanity and the Metaphysics of Responsibility. In F. D. Schoeman (Ed.), *Responsibility, Character, and the Emotions: New Essays in Moral Psychology* (pp. 46–62). Cambridge University Press.
- Wolterstorff, N. (2007). The Paradoxical Role of Coercion in the Theory of Political Liberalism. *Journal of Law, Philosophy and Culture*, 1, 135.
- Zak, P. J. (Ed.). (2008). *Moral Markets: The Critical Role of Values in the Economy*. Princeton: Princeton University Press.