

JOURNAL *of* ETHICS  
& SOCIAL PHILOSOPHY

VOLUME XXI · NUMBER 3

*March 2022*

ARTICLES

- 299 Voter Motivation  
*Adam Lovett*
- 329 Why Paternalists Must Endorse Epistocracy  
*Jason Brennan and Christopher Freiman*
- 354 Well-Being as Need Satisfaction  
*Marlowe Fardell*
- 396 An Account of Normative Stereotyping  
*Corey Barnes*
- 422 What Is the Bad-Difference View of Disability?  
*Thomas Crawley*

DISCUSSIONS

- 452 The Goal Problem in the “Now What” Problem  
*Xinkan Zhao*
- 459 Normative Uncertainty without Unjustified  
Value Comparison: A Response to Carr  
*Ron Abodi*

The *Journal of Ethics and Social Philosophy* (ISSN 1559-3061) is a peer-reviewed online journal in moral, social, political, and legal philosophy. The journal is founded on the principle of publisher-funded open access. There are no publication fees for authors, and public access to articles is free of charge and is available to all readers under the CREATIVE COMMONS ATTRIBUTION-NONCOMMERCIAL-NODERIVATIVES 4.0 license. Funding for the journal has been made possible through the generous commitment of the Gould School of Law and the Dornsife College of Letters, Arts, and Sciences at the University of Southern California.

The *Journal of Ethics and Social Philosophy* aspires to be the leading venue for the best new work in the fields that it covers, and it is governed by a correspondingly high editorial standard. The journal welcomes submissions of articles in any of these and related fields of research. The journal is interested in work in the history of ethics that bears directly on topics of contemporary interest, but does not consider articles of purely historical interest. It is the view of the associate editors that the journal's high standard does not preclude publishing work that is critical in nature, provided that it is constructive, well-argued, current, and of sufficiently general interest.

*Executive Editor*

Mark Schroeder

*Associate Editors*

|                       |                    |
|-----------------------|--------------------|
| Saba Bazargan-Forward | Hallie Liberto     |
| Stephanie Collins     | Errol Lord         |
| Dale Dorsey           | Tristram McPherson |
| James Dreier          | Colleen Murphy     |
| Julia Driver          | Hille Paakkunainen |
| Anca Gheaus           | David Plunkett     |

*Discussion Notes Editor*

Kimberley Brownlee

*Editorial Board*

|                          |                    |
|--------------------------|--------------------|
| Elizabeth Anderson       | Philip Pettit      |
| David Brink              | Gerald Postema     |
| John Broome              | Joseph Raz         |
| Joshua Cohen             | Henry Richardson   |
| Jonathan Dancy           | Thomas M. Scanlon  |
| John Finnis              | Tamar Schapiro     |
| John Gardner             | David Schmidtz     |
| Leslie Green             | Russ Shafer-Landau |
| Karen Jones              | Tommie Shelby      |
| Frances Kamm             | Sarah Stroud       |
| Will Kymlicka            | Valerie Tiberius   |
| Matthew Liao             | Peter Vallentyne   |
| Kasper Lippert-Rasmussen | Gary Watson        |
| Elinor Mason             | Kit Wellman        |
| Stephen Perry            | Susan Wolf         |

*Managing Editor*

Rachel Keith

*Copyeditor*

Susan Wampler

*Typesetting*

Matthew Silverstein



## VOTER MOTIVATION

*Adam Lovett*

SOMETIMES we vote on the issues. Consider a voter who detests gun control. They might, on this basis, vote Republican. Their opposition to gun control drives their vote choice. They vote Republican because they share the Republican Party's policy position. But sometimes we instead vote on performance. Consider a voter who loves the booming 1990s economy. They might, on this basis, vote for Bill Clinton. Their assessment of the incumbent's performance drives their vote. They vote for Clinton because, well, it's the economy, stupid. And, sometimes, we vote on group identities. Consider a Catholic from 1960. Perhaps they cannot stomach voting against their church. They might, on this basis, vote for John F. Kennedy. They vote for Kennedy because he is a Catholic, like them. Their group identities drive their vote. These voters differ in the basis on which they vote. They differ in the reasons they have for voting the way they do. Policy issues drive issue voters. Performance issues drive performance voters. Group identities drive group voters. But which type of voting is best for democracy? And how well do we do? And, finally, on what basis should each of us vote?

These questions concern a single topic: voter motivation. The first question plumbs how the prevalence of different kinds of voter motivations impact democratic values. Answering this tells us what would motivate voters in an ideal democracy. The second question plumbs how voters' actual motivations matter to such values. Answering this tells us how far from the ideal our real-world democracies are. The third plumbs what motivations should drive actual individual voters. The answer to this depends, in part, on how voters contribute to democratic values in their nonideal democracies. For each question, I will concern myself with intrinsic democratic values alone. These are the ways that democracy is valuable in itself, besides its causal consequences. Voter motivations no doubt matter to the instrumental value of a democracy. But democracy's intrinsic value is my focus. For the second and third questions, I will concern myself with American voters and American democracy alone. Much of what I will say applies elsewhere. But American democracy is my focus. Together, these questions plumb how voter motivations interact, both evaluatively and deontically,

with the intrinsic value of American democracy. That interaction is the topic of this paper.

These three types of voting have been the subject of sustained empirical investigation. But they have not been the subject of much normative investigation. When political scientists evaluate them, they do so in terms of instrumental values. For example, they explore which motivation will produce the best policy.<sup>1</sup> They ignore how these motivations matter to intrinsic democratic values. Meanwhile, political theorists have written a lot about voting but little about voter motivations. Rather, they have addressed whether citizens ought to vote in the first place. The driving problem here is that each vote has a very small chance of making a difference to an election. So, is it rational to vote at all?<sup>2</sup> Much time has been spent on this question. Little has been spent exploring what should motivate those who do vote.<sup>3</sup> Jason Brennan has investigated a connected topic.<sup>4</sup> He has examined whether those who vote ought to know about politics.<sup>5</sup> The connection, as we will see later, is that voter competence and voter motivation interact in contributing to democratic values. But voter competence, on its own, tells us little about voter motivations. Voter motivations, then, have been largely neglected: I think that that neglect is unfair. Such motivations, I will argue, matter to the intrinsic value of democracy.

Here is the plan for the rest of the paper. In section 1, I will say more about the nature and prevalence of these different types of voter motivations. In section 2, I will outline two core intrinsic democratic values: equality and self-rule. In section 3, I will identify how different types of voter motivations matter to these values. My view is that issue voting is better than performance voting and performance voting is better than group voting. This is not meant to be a radical view. It seems to me to be the conventional wisdom. But the grounds of that wisdom are not well understood. This paper identifies those grounds. In section

- 1 See, for instance, the discussion of retrospective voting in Achen and Bartels, *Democracy for Realists*, ch. 4.
- 2 For the problem, see Downs, *An Economic Theory of Democracy*, 274. For three different responses, see Parfit, *Reasons and Persons*, 73–75; Goldman, “Why Citizens Should Vote”; Guerrero, “The Paradox of Voting and the Ethics of Political Representation.”
- 3 Some working in the “public reason” tradition do discuss it. Such writers claim, or presuppose, that state action is legitimate if and only if it is supported by a justification that all reasonable people accept. Among those who think this, Rawls denies that motivation matters much to how we should vote (*Political Liberalism*, 235). Quong contends that it does matter (*Liberalism without Perfection*, 274–90). It is not clear, however, what import this discussion has for those of us who do not accept public reason presuppositions.
- 4 Brennan, *The Ethics of Voting*.
- 5 Brennan thinks so. For a reply, see Arvan, “People Do Not Have a Duty to Avoid Voting Badly.”

4, we will turn to how these types of voter motivation interact with voter competence. I will argue that voter incompetence modulates the effect of voter motivations on self-rule but leaves their effect on equality untouched. In section 5, we will see what this means for American democracy. The American voter, I will suggest, rarely votes in the ways intrinsic democratic values require. American democracy is deeply defective. Finally, in section 6, we will see what that means for how Americans should vote. The key conclusion here is that their nonideal circumstances weaken their obligations. In such conditions, they merely must avoid voting on privileged group identities.

### 1. TYPES OF VOTER MOTIVATION

The three kinds of voting we will focus on are voting on the issues, voting on performance, and voting on group identities. We focus on these not because they are the only possible motivations voters could have. Rather, we focus on them precisely because they have been the subject of such sustained empirical investigation. Issue voting is at the core of spatial modeling of voting behavior.<sup>6</sup> Early empirical researchers took it to be an influential driver of voting.<sup>7</sup> Voting on performance became a topic core to the study of voting behavior in the 1970s. A vast literature plumbs, in particular, whether and how voters respond to the economic performance of incumbents.<sup>8</sup> Voting on group identities was a preoccupation of the early empirical literature on voting behavior.<sup>9</sup> Recently, it has again become a prominent focus. Achen and Bartels claim that, in the political sphere, group identities form “the very basis of reasons.”<sup>10</sup> This empirical literature allows us to assess the prevalence of each kind of voter motivation. As we will later see, that will be essential to evaluating the quality of American democracy and the duties of American citizens. But first I will say more about each kind of voting.

We will begin with issue voting. This is voting on the basis of shared policy platforms or issue positions. Consider Democrats who voted for Barack Obama because they wanted public health care. They were issue voting. Or consider Republicans who voted for Donald Trump because they wanted to build a wall.

6 Downs, *An Economic Theory of Democracy*, ch. 8.

7 Campbell et al., *The American Voter*, 112–36.

8 Kramer, “Short-Term Fluctuations in U.S. Voting Behavior, 1896–1964”; Fiorina, *Retrospective Voting in American National Elections*.

9 Berelson, Lazarsfeld, and McPhee, *Voting*, 58–87; Campbell, Gurin, and Miller, *The Voter Decides*, 88–112.

10 Achen and Bartels, *Democracy for Realists*, 213.

They too were issue voting. Their agreement with that candidate on the issues drove their vote. They wanted certain policies enacted. These candidates said that they would enact them. This is why they voted for the candidate. How often does issue voting happen? The preponderance of evidence indicates that it does not happen very often. There are two weighty pieces of evidence for this.<sup>11</sup> The first turns on what voters say when you ask them what they like about different candidates. They rarely mention policy issues. Fewer than 20 percent mention any issue positions at all. So issue positions seem unlikely to drive vote choice. The second is that voters themselves likely lack firm positions on most issues. Their expressed issue positions are inconstant. At one time, they will say that they are all for, for example, federally provided universal employment. At another they will say that they are all against it. Voters seem to be constructing an opinion on the fly.<sup>12</sup> But opinions constructed on the fly surely do not drive vote choice. This evidence suggests that issue voting is relatively rare: it happens more often in textbooks than ballot boxes.

Not everyone is convinced by this evidence. Some people think that issue voting happens quite often. They point out that voters' issue stances correlate with their vote choice. Voters vote for the party that shares their issue stances. And so these people infer that voters' issue positions drive whom they vote for.<sup>13</sup> But, in turn, many find this argument unconvincing. The problem is that this evidence does not establish the direction of causality. People often take their issue position from the party that they are going to vote for. They conform their policy stance to the party line. So these correlations might be due to people's vote choice driving their policy preferences rather than their policy preferences driving their vote choice.<sup>14</sup> And there is good evidence that this is what is going on. In some cases, one can identify exactly when people find out that they do not share their preferred candidate's issue position. Afterward, they more often change their mind on the issue than stop liking the candidate.<sup>15</sup> So, it seems to me unlikely that issue voting happens very often.

11 The first of these pieces comes from Campbell et al., *The American Voter*. The second comes from Converse, "The Nature of Belief Systems in Mass Publics (1964)." For contemporary updates on both pieces of evidence, see Lewis-Beck et al., *The American Voter Revisited*, ch. 10; and Kinder and Kalmoe, *Neither Liberal nor Conservative*.

12 For further discussion, see Zaller, *The Nature and Origins of Mass Opinion*.

13 For an influential example of this argument, see Ansolabehere, Rodden, and Snyder, "The Strength of Issues."

14 For this reply, see Achen and Bartels, *Democracy for Realists*, 41–45.

15 The direct evidence for this is from Lenz, *Follow the Leader?* chs. 3, 8. But for supporting evidence, see Cohen, "Party over Policy"; and Berinsky, *In Times of War*. Now, one explanation of these findings is that voters have firm ideologies but do not know what policies best



Now let us turn to performance voting. This is voting based on the expected performance of the candidates.<sup>16</sup> It is voting on one's expectations about their performance at promoting widely shared goals. Think of those Democrats who voted for Clinton because they thought he would boost the economy. They were performance voting. Or consider Republicans who voted for George W. Bush because they thought he would make America safer. They too were performance voting. These people might have had no view on which policies would help with prosperity or safety. They might just have had views on which candidate would best promote such goals. Often, such views are based on assessments of prior performance in office. These are called "retrospective assessments." But they might also be based in the perceived personal qualities of candidates: their integrity, intelligence, competence, and so on. All these things can ground assessments of a candidate's expected performance.

Among political scientists, the consensus is that performance voting is extremely common. The best evidence for this involves retrospective voting on the economy. A huge number of observational studies look at such voting behavior. Incumbents suffer when the economy is diving. They flourish when it is rising.<sup>17</sup> There are also some panel-survey studies on performance voting. These

---

align with those ideologies. Yet they can identify which politicians share their ideologies. Thus, they adopt the policy stances of these politicians as a quick and easy way of adopting the policy stance most congruent with their ideologies. Popkins (*The Reasoning Voter*) and Lupia and McCubbins (*The Democratic Dilemma*) make this claim. But I doubt this for two reasons. First, I doubt that voters have firm ideologies. Kinder says, "Precious few Americans make sophisticated use of political abstractions. Most are mystified by or at least indifferent to standard ideological concepts" ("Opinion and Action in the Realm of Politics," 796). The evidence for this is, *inter alia*, that many citizens are simply unable to say much about the content of different political ideologies. Here see Kinder and Kalmoe, *Neither Liberal nor Conservative*, 11–43. Second, there are other explanations of what is going on when voters adopts elites' policy stances. The foremost explanation puts it down to motivated reasoning: partisans are strongly driven to agree with their party. They care much less about whether they have accurate political beliefs. Lab experiments cohere better with this view than the one that rests on ideology. See, for example, Petersen, et al., "Motivated Reasoning and Political Parties"; Bolsen, Druckman, and Cook, "The Influence of Partisan Motivated Reasoning on Public Opinion." So the ideology-based explanation of these findings seems to me to be dubious. But, in any case, the key point is that these findings mean correlations between issue positions and vote choice are weak evidence that the former cause the latter.

16 For the "performance" terminology, see Lenz, *Follow the Leader?* 2. We might label this kind of voting "expected performance voting" instead of performance voting. But to retain consistency with the empirical literature I prefer to simply call it "performance voting."

17 For the seminal works on this, see Key, *The Responsible Electorate*; Kramer, "Short-Term Fluctuations in U.S. Voting Behavior, 1896–1964"; Fiorina, *Retrospective Voting in American National Elections*. For a recent discussion of this classic literature, see Achen and Bartels,

studies interview the same individuals many times. This lets researchers see whether performance assessments change before vote intentions change or vice versa. Gabriel Lenz's *Follow the Leader?* is a landmark such study.<sup>18</sup> He shows that, when people think the economy is doing badly, they later reduce their approval of incumbent presidents. The former seems to be causing the latter. It is a short jump from this to the conclusion that economic perceptions also drive vote choice. Performance issues, or at least the issue of prosperity, have a pervasive impact on vote choice.

Let us turn to group voting. This is voting on the basis of group identities. Catholics voted for Kennedy. White southerners voted for George Wallace. Black people voted for Obama. It is standard to understand this in terms of social identities.<sup>19</sup> Social identities start with self-categorization: we see ourselves as members of certain groups. And they add to this an emotional charge: we care about our group memberships. How does that affect voting behavior? Well, when we have such a social identity, we are driven to achieve positive distinctiveness for it. That means we are driven to raise the status of our group relative to that of other groups: we want to "maintain or achieve superiority over an out-group."<sup>20</sup> In the electoral context, getting a group member or affiliate into office is the main way to do this. Having a president who comes from your group enhances your group's status. Thus, we often vote for fellow group members or affiliates of our groups. When I talk about group voting, I mean voting so driven by social identities.

Why construe group voting like this? Because it comports well with social identity theory. This theory is rooted in experiments conducted by Henri Tajfel in the late 1960s. Tajfel set out to plumb the origins of group conflict. He assigned people to groups arbitrarily. In one such experiment, he did this by asking them which of two abstract artworks they preferred. After picking, the subjects were told they were either in the group that liked Klee or that liked Kandinsky. He then asked them to allocate money among the other subjects. They could choose to ensure either that (1) everyone got the maximum amount of money or (2) their group got more money than the other group, but less than the maximum possible. He found that subjects favored 2. They preferred their group to

---

*Democracy for Realists*, 93–98. For an overview of the later literature, see Lewis-Beck and Stegmaier, "Economic Voting."

18 Lenz, *Follow the Leader?*

19 See, for example, Achen and Bartels, *Democracy for Realists*, 228–29; and Mason, *Uncivil Agreement*, 1–17.

20 Tajfel and Turner, "The Social Identity Theory of Intergroup Behavior," 378.

be worse off in absolute terms but better off relative to other groups.<sup>21</sup> These experiments showed, first, that it is easy to motivate people by group identities. In Tajfel's experiments, subjects never even saw members of either group. They were told only that they had similar taste in art. And they showed, second, that when driven by such identities, we do not just want our group to do well. We want it to win: we want it to be superior to other groups. These claims are at the core of social identity theory. The first makes it likely that identities are operative in political contexts. The second suggests that we should understand that operation in terms of status enhancement. Thus, this more basic psychological theory grounds our construal of group voting.

Group voting also seems to be very common. Race, religion, gender, and geography are all common bases for group voting.<sup>22</sup> But perhaps the most common type of group voting is voting on party identification. Those who identify as Democrats vote for the Democratic party. Those who identify as Republicans vote for the Republican party. Why think of this as a kind of group voting? Because party identification behaves like a social identity. It is more like Catholicism than it is like Libertarianism.<sup>23</sup> People avow their party identifications in survey interviews. They talk about their party in terms of "we." They feel attacks on their party as personal insults. They get a party identification by early adulthood. They usually stick with it for the rest of their lives. Party identification looks for all the world like a social identity.<sup>24</sup> Thus, since it has a pervasive impact on vote choice, group identities have such an impact.

In sum, on the strength of this evidence, group and performance voting happen often. Issue voting is rarer. I want to end this section with two final, clarificatory points. First, I wish to stress again that these three kinds of voting do not exhaust voters' possible motivations. Perhaps voters also vote based on candidate charisma, or on their perceived self-interest. But we have less empirical traction on these issues than on the three types of voter motivation just canvassed.<sup>25</sup>

21 For the striking original finding, see Tajfel et al., "Social Categorization and Intergroup Behaviour." It has been widely replicated. See, for example, Billig and Tajfel, "Social Categorization and Similarity in Intergroup Behaviour"; Locksley, Ortiz, and Hepburn, "Social Categorization and Discriminatory Behavior"; Gagnon and Bourhis, "Discrimination in the Minimal Group Paradigm."

22 Achen and Bartels lay out some case studies supporting this (*Democracy for Realists*, ch. 7).

23 The canonical source of this idea is Campbell et al., *The American Voter*.

24 For this evidence, see Green, Palmquist, and Schickler, *Partisan Hearts and Minds*, 32–40, ch. 3.

25 In this connection, I want to make a remark about self-interest as a voter motivation. There is a large literature, stemming from Kinder and Kiewiet ("Sociotropic Politics") on whether performance voters are *pocketbook* voters or *sociotropic* voters. Pocketbook voters vote for incumbents when they think that they personally have been doing well. Sociotropic voters

And, as the evidence I have cited indicates, many of these kinds of voter motivations clearly matter. They have a big impact on how voters behave. So, they are a good place to start. They carve out important drivers of voter behavior, the prevalence of which we have some grasp on. Thus, understanding the normative significance of these kinds of voter behavior will put us in a position to answer concrete normative questions about American democracy.

Second, many voters no doubt have multiple of these motivations. They are motivated in part by the issues, in part by performance, and in part by group identities. Sometimes, these motivations may be entangled. One might, for example, have one's policy position because of one's group identity. Perhaps one opposes gun control because one identifies as a white man.<sup>26</sup> Or, to take another example, one's group identity might lead one to prioritize certain performance issues. Perhaps one thinks terrorism is the top priority because one identifies as a Republican.<sup>27</sup> Nonetheless, we can disentangle the impact of different motivations. In theory, although rarely in practice, we can see in individual voters the relative force of these factors. We can say whether they were driven more by the issues, or by performance, or by group identities. In both theory and practice we can say, for the electorate as a whole, which of these motivations has the biggest impact on vote choice. That is what the empirical work just cited attempts to do. We will return to this issue in section 3. But that is all we will need to do to answer our normative questions. Yet, before turning to that, I must say more about what makes democracy valuable.

## 2. DEMOCRATIC VALUES

In this section, I spell out a conception of democracy's intrinsic value. This conception will be my own, but it has deep roots in democratic theory. The con-

---

vote for incumbents when they think that the national economy has been doing well. This distinction is sometimes equated with that between self-interested and altruistic voting. See, e.g., Brennan, *The Ethics of Voting*, 162–63, *Against Democracy*, 49–51, and “The Ethics and Rationality of Voting.” But this is a mistake. Sociotropic voters, as argued persuasively by Kiewiet and Lewis-Beck, may be entirely self-interested (“No Man Is an Island”). They may be voting for the candidate whom they see as good for the national economy solely because *they themselves* will do well when the national economy is doing well. Indeed, this point is made clear by Kinder and Kiewiet's initial paper on this topic. Kinder and Kiewiet stress that the “distinction between pocketbook and sociotropic politics is *not* equivalent to the distinction between a self-interested and an altruistic politics” (“Sociotropic Politics,” 132). Thus, we know frustratingly little about how much voters are driven by self-interest.

26 Melzer claims that this is common (*Gun Crusaders*).

27 This is consistent with survey data. See Jones, “Republicans and Democrats Have Grown Further Apart on What the Nation's Top Priorities Should Be.”

ception concerns intrinsic democratic values alone. These contrast with instrumental values. Intrinsic values make things valuable in themselves. Instrumental values make things valuable for their causal consequences or capacities. A good friendship is intrinsically valuable. A good hammer is just instrumentally valuable. The intrinsic democratic values we will focus on are equality and self-rule. In recent years, writers such as Niko Kolodny and Daniel Viehoff have advocated for the former.<sup>28</sup> They think that democracies are egalitarian in a way that other political systems are not.<sup>29</sup> Advocacy of the latter has a long and venerable history. This is the value tapped by Rousseau when he insists that “the people, subjected to law, ought to be its author,” and the United Nations when its treaties assert that “all peoples have the right to self-determination.”<sup>30</sup> My own view is that all noninstrumental democratic values reduce to these two values. So, the impact of voting behavior on equality and self-rule just is its impacts on the intrinsic value of democracy. But to determine this impact, we need the right conception of these values.

Let us start with equality. Democracies, many think, are distinctively egalitarian. And many spell out democratic equality as a type of relational equality.<sup>31</sup> In part, that consists in avoiding inegalitarian relationships. Paradigm examples of such relationships are those between a master and slave or the members of different castes. Both relationships are intrinsically bad. And both relationships are partly constituted by inequalities of power. Part of what it is to be a slave, or a member of a lower caste, is to lack relative power. What does democracy have to do with this? Well, we can more or less stipulatively define a democracy as a political system in which political power is equally distributed and the exercise of that power determines what government does.<sup>32</sup> On such a definition, democracy is constituted by equalities of political power. Thus, democracy helps preclude inegalitarian relationships. This is what we will call the negative aspect of democratic equality. This aspect consists in the minimization of relationships of domination, subordination, and hierarchy.

28 Kolodny, “Rule Over None I”; and Viehoff, “Power and Equality.”

29 For an older source of this idea, see Tocqueville, *Democracy in America*, 9, 14.

30 For the quotes, see Rousseau, *The Social Contract*, 2.6.10; and United Nations, “International Covenant on Civil and Political Rights,” article I. For some contemporary theories with this view, see Stilz, “The Value of Self-Determination”; and Zuehl, *Collective Self-Determination*.

31 This is the view in Kolodny, “Rule Over None I”; and Viehoff, “Democratic Equality and Political Authority” and “Power and Equality.” For a different conception of equality, see Christiano, *The Rule of the Many* and *The Constitution of Equality*. Most of what I say on the relational egalitarian conception would also go for Christiano’s conception.

32 Kolodny also opts for an essentially stipulative definition of democracy (“Rule Over None II,” 197).

But that does not exhaust the egalitarian value of democracy. This is because the mere absence of inegalitarian relationships does not exhaust relational egalitarian values. There are also intrinsically good relationships. On the small scale, friendship is the paradigm example. This is an intrinsically good, egalitarian relationship. Friendship does not just amount to non-domination: you are not friends with all those who avoid subordinating you. And it is not just instrumentally useful to have friends; it is good in itself. Now there are some similarities between good friendships and relationships of non-domination. In particular, friendship in part consists in equalities of power. Good friends do not wield asymmetric power over one another. But friendship requires more than just such equalities. Friends must be committed to preventing inequalities from arising. You are not friends with someone who would lord it over you if they had the chance. And friends must care appropriately about one another's welfare. You are not friends with someone who does not care about how your life goes. Friendship, in these ways, is a thicker relationship than mere non-domination.

For democratic theory, the critical claim is that you can also have large-scale such relationships. We will call these civic friendships.<sup>33</sup> These consist in part in non-domination. Civic friends can look one another in the eye.<sup>34</sup> They are not subservient to one another. But, for evaluating voter motivations, two further conditions on such relationships are critical. First, citizens' commitment to avoiding inegalitarian relationships is important. Imagine someone who would, given the chance, make themselves the dictator of their fellow citizens. They are not committed to avoiding inegalitarian relationships. This diminishes the positive value of their relationships with their fellow citizens. If they are completely indifferent to the equality of those relationships, I suspect that they are not in relationships of civic friendship at all. Second, citizens' care for others' welfare is important. Imagine someone who would sacrifice very little for the benefit of their fellow citizens. They do not care much about their fellow citizens' welfare. This again diminishes the positive value of their relationship with their fellows. If they are completely indifferent to that welfare, then again they are not in relationships of civic friendship at all. But when all these conditions are met, at least to a minimal extent, we have civic friendships. Democracy consists, in part, in the equalities of power necessary to these relationships. This is the positive aspect of democratic equality. It consists in democracy facilitating relationships of civic friendship.

33 The term comes from Schwarzenbach, "On Civic Friendship." For the most extensive defense of this as a democratic value, see Viehoff, "Power and Equality." Scheffler provides the underlying positive conception of egalitarian relationships ("The Practice of Equality").

34 The eyeballing metaphor comes from Pettit, *On the People's Terms*, 47.

Let us now turn to a second democratic value: self-rule. This consists in the manifestation of the people's will in their social and political affairs. The conception of this I favor hinges on joint intentions.<sup>35</sup> A joint intention is just an intention one shares with other people. When we together intend to sing a duet, paint a house, raise a child we have a joint intention. Now suppose some citizens have a joint intention to bring about some political event. This could be an action of government or an outcome of government action. And suppose their having this intention brings about this thing. Then we can say that they are self-ruling with respect to that outcome or action. The more people are self-ruling with respect to more actions or outcomes, the more the political system realizes the value of self-rule, and the more that political events manifest our joint intentions. What does democracy have to do with this? Well, for people to be self-ruling there must be a causal connection between their will and policy. On the definition above, democracy in part consists in such influence. Thus, democracy ensures that a necessary condition for self-rule is satisfied. This is another part of its intrinsic value.

This is, right now, a controversial view. Recently, some egalitarians have said that the only democratic value is an egalitarian one.<sup>36</sup> They have thought this because it is hard to give a good explanation of why self-rule is important. And without such an explanation, so they have thought, we should not think it is important. This seems to me rash. It is very intuitive that there is a democratic value in the vicinity of self-rule. Here is an example of the intuition: suppose we got rid of government by human beings and replaced it with government by algorithm.<sup>37</sup> The algorithm we replaced it with, let us stipulate, spits out perfect legislation. It institutes far superior legislation than any human government could. Yet, in this situation, citizens have no influence over the laws that govern them. It seems to me compelling that something is lost here. If we did this, we would be sacrificing something important about democracy. But that cannot be an egalitarian loss: in this case every person has equal political power (zero). Rather, it is a loss associated with lack of influence over the laws to which you are subject. So, intuitively, self-rule is valuable.

But we would still like an explanation of why self-rule is valuable. The account I favor hinges on the value of self-authorship. Being the author of your life is attractive. It is good to be responsible for what has a big impact on your life.

35 This type of account comes from Stilz, *Liberal Loyalty*; and Zuehl, *Collective Self-Determination*. But as I say in the text, I think the underlying idea has a long history. For example, we can see it in Rousseau, *The Social Contract*.

36 Kolodny, "Rule Over None II."

37 This case is from Zuehl, *Collective Self-Determination*, 18–19.



We can see this in personal cases. Compare two people. One has a conception of the good life and pursues it. They deliberately live their lives in accord with their values. The other makes few real choices. They might have a conception of the good life. But they do not pursue it: they just go with the flow. Intuitively, there is something preferable about the first life. We want mastery, not drift. We want people to be the author of their own lives.<sup>38</sup> In the personal case, what it is to be the author of things in your life is for your intending them to bring them about. You become a pilot because you intended to be; you marry your partner because they were whom you wanted to marry. But there is also a social dimension to this notion of authorship. You can, together with others, intend to bring about certain political outcomes. When this joint intention brings about those outcomes, you are their joint author. This is valuable in much the same way that single authorship is valuable. Such things have a huge impact on you. It is valuable to be partly responsible for things with such an impact on you. Self-rule helps realize this value.

So there are two parts to what makes democracy intrinsically valuable. On the one hand, democracy advances relational equality. This advancement itself has two aspects. The negative aspect amounts to the avoidance of inegalitarian relationships. The positive aspect amounts to the facilitation of egalitarian relationships. On the other hand, democracy advances self-rule. It helps make citizens joint authors of their social and political affairs. Advocacy of each value has a long history in democratic theory. It is plausible that both make democracy intrinsically valuable. Now there might be other things that make democracy intrinsically valuable. Perhaps the very act of democratic deliberation has intrinsic value. Perhaps simply resolving disagreement democratically has intrinsic value. And perhaps neither value reduces to the value of equality or self-rule. I doubt this, but I have given no evidence against it. Yet we will go forward with a focus on equality and self-rule. If there are other democratic values, then this will give us just a partial answer to how voting behavior affects intrinsic democratic values. But it will still provide an important part of the answer. So, with this caveat in mind, we can move to my first question: How does the prevalence of certain types of voter motivations affect these intrinsic democratic values?

### 3. EVALUATING VOTER MOTIVATIONS

First, we look at issue voting. Suppose everyone voted on the basis of policy issues. Imagine policy stances motivated people's vote choice. How much would this facilitate democratic values? I think the answer is: a lot. Let us start by look-

38 Raz does much more to spell out the attractiveness of this thought (*The Morality of Freedom*).



ing at how it would affect self-rule. Consider the people who, in 1932, voted for Franklin Delano Roosevelt (FDR). Imagine they did so because they wanted unemployment relief. This is a prerequisite for their having a joint intention to enact employment relief. It is a prerequisite for that intention bringing about unemployment relief. So enactment of employment relief might manifest the joint intentions of FDR voters. So, these people may be self-ruling with respect to unemployment relief. More generally, issue voting is a prerequisite for policy manifesting joint intentions. The more widespread issue voting is, the better positioned people are to be self-ruling with respect to particular policies.

Now, widespread issue voting does not guarantee such self-rule. Issue voters might not *jointly* intend to enact any policies. To see this, suppose that Bratman's account of such intentions is right. Bratman thinks that some people have a joint intention to  $\phi$  when (a) they each intend that they together  $\phi$ , (b) they have jointly compatible plans for contributing to  $\phi$ -ing, and (c) they are not coerced into  $\phi$ -ing.<sup>39</sup> Issue voters might fail to meet these conditions. They might, for example, only think of their own contributions to policy. They might not intend that they together with others enact policy. But, in truth, these conditions are not that hard to meet. FDR voters could easily have intended to bring about unemployment relief with other FDR voters. Their individuals plans to contribute to this—voting for FDR—are jointly compatible. And nobody was coerced into voting for FDR. So widespread issue voting does not ensure that voters have the joint intentions that self-rule requires. But it puts them in a good position to have such intentions. It helps enable them to be self-ruling.

Yet there is a more important way in which issue voters might fail to achieve the value of self-rule. They might be very incompetent. Suppose that they do not know much about FDR's policies. They have an inkling that he is the one offering a New Deal to the American people. But they cannot really remember. Might it not have been, they wonder, Hoover who was banging on about a deal? But, on the basis of the inkling, they vote for FDR. Here, they are not very competent voters. If they aimed to help enact the New Deal, their actions did not very reliably contribute to this goal. They could have easily voted for the candidate who would stymie it. We will talk more about such incompetence in the next section. But, for now, I will just register the belief that when issue voters are incompetent in this way, they achieve little self-rule. Voter incompetence means policies at most match, rather than manifest, voters' joint intentions. Thus, widespread issue voting aids, without assuring, the achievement of self-rule.

Let us turn to equality. Citizens need certain attitudes to achieve the positive aspect of democratic equality. They must have some care for the welfare of their

39 Bratman, "Shared Cooperative Activity."

fellow citizens. They must be committed to avoiding inegalitarian relationships. Issue voters can fall short on these commitments. Consider people who voted for Wallace in 1968 because they liked his segregationist platform. These people were issue voters. But they do not achieve the positive aspect of democratic equality. They violate both conditions. They were not sufficiently concerned for the welfare of their fellow citizens. They were not sufficiently committed to the avoidance of inegalitarian relationships. So, for issue voters to help achieve this value, they cannot vote on the basis of odious commitments. But issue voting is compatible with such abstinence. Issue voters might well vote on issues that are not odious. So, not all issue voting is consistent with the positive egalitarian value. But there is no inherent tension between issue voting and democratic equality. Issue voting, when combined with the other attitudes, does facilitate such equality.

Second, we look at performance voting. Suppose everyone votes on the basis of expected performance in office. Expected performance motivates vote choice. How much does this facilitate democratic values? We will start with self-rule. Self-rule is a little less well achieved by widespread performance voting than by widespread issue voting. That is because it is only outcomes that can now manifest people's intentions. Suppose people voted for FDR, in 1932, because they thought he would be a better economic performer than Hoover. That is a low bar, but it paid off handsomely. FDR did not just enact unemployment relief; he helped pull America out of the Great Depression. In this case, the economic upturn might well manifest the joint intentions of FDR voters. But the actual policies that FDR implemented would not have manifested these intentions. More generally, performance voting fits with outcomes, rather than policies, manifesting voters' intentions.

Why is this worse than issue voting? Well, to explain that we have to make some more assumptions about issue voting. I assume that few people want a set of policies with total disregard for the outcomes of those policies. They think that those very policies will produce some desired outcomes. So they also have the intention to produce an outcome. So, for such issue voters, both policies and outcomes manifest their joint intentions. That is why they have a leg up on performance voters. For performance voters, only the outcomes manifest the intentions. Performance voters might well be responsible for large parts of their social environment. But issue voters—at least given certain assumptions—are responsible for larger parts. But I want to be clear on my view here: the leg up is the size of a small leg. Issue voting beats out performance voting on achieving self-rule. But the margin of victory is not large. Both seem to me respectable ways of achieving this value.

Let us turn to equality. Issue voting and performance voting are in the same position when it comes to equality. Performance voting does not guarantee the achievement of the positive aspect of democratic equality. Some people performance vote on the basis of inegalitarian commitments. Their performance voting will not aid this value. Some vote on sheer self-interest. They ask not what a candidate can do for their country, but just what the candidate can do for them. This does not help the achievement of democratic equality. But performance voters need not exhibit such misbehavior. They might vote for whom they think will produce the best outcomes for all their fellow citizens. They might vote for Clinton because they think he will make everyone better off. So widespread performance voting and issue voting are consistent with democratic equality. Neither ensure it, but both can facilitate it.

Finally, we look at group voting. Suppose everyone votes on the basis of their group identities. They vote for candidates affiliated with the groups with which they identify. And they do this to boost the relative social standing of their group. How does this affect democratic values? We start with self-rule. This type of voter motivation, were it widespread, would not be good for self-rule. When you group vote, neither the policies of government nor the outcomes of those policies manifest your intentions. You did not intend to bring about any particular policies. You did not intend to bring about any particular outcomes. You voted on the basis of group affiliation. So group voters do not enjoy self-rule with respect to policies or their outcomes. Now it is not that they enjoy nothing. When they get someone affiliated with their group into office, this can count as the manifestation of their intentions. Any ensuing change in social hierarchies can also count as manifesting their intentions. But, generally, such changes are not enormous. Obama's election did not transform race relations in the United States. So this makes voters, at best, responsible for only minor changes in status hierarchies. Yet such minor changes are less important to citizens' social and political affairs than is government policy and the huge changes to social life wrought by such policy. Thus, widespread group voting would not much help the achievement of self-rule.

Now let us consider equality. Is widespread group voting consistent with the positive aspect of democratic equality? This depends on the type of group voting. There are three types. First, there is *maintaining superiority*. Suppose one identifies with a group that holds a privileged place in a social hierarchy. One votes as one does to maintain this group's elevated place in the hierarchy. This is surely incompatible with a commitment to social equality. You cannot be both committed to social equality and motivated by maintaining the status superiority of your group. This is exactly a vote motivated by a commitment to social

*inequality*. In the United States, some instances of racial voting give us concrete examples of this. The United States is a racially stratified society. It is not white people who suffer from racial oppression. So consider the case of white people who vote on the basis of their racial identity. This is a case of maintaining social superiority. If such voting is widespread, then that impairs the realization for the positive egalitarian value.

Second, there is *creating superiority*. Suppose one identifies with a group that holds neither a high nor low place in the social hierarchy. One hopes one's vote will facilitate a realignment in status hierarchies. It will help this group gain status and, in particular, become superior to other groups. This again is incompatible with a commitment to social equality. Such voting behavior is part of a commitment to social inequality. The best concrete example of this is voting on the basis of party identification. In the United States, party groups hold roughly similar levels of social status. So, consider Republicans who vote for the Republican candidate to raise the social status of Republicans. They are attempting to create social superiority. This is incompatible with a commitment to social equality. So widespread group voting of this type would also impair the positive aspect of democratic equality.

Third, there is *ameliorating inferiority*. Suppose one identifies with a group that holds a low place in the social hierarchy. One votes for a group-affiliated candidate to ameliorate the status inferiority of this group. One hopes that, if the candidate wins the election, the group will gain status. The status gain will not make that group superior to other comparison groups, but rather will make it closer to their equal. This seems completely consistent with a commitment to social inequality. The driving force here is not a desire for social superiority; it is a desire for equality. In the United States, much race-based voting exemplifies this. Consider Black voters who voted for Obama. This need not have hurt the positive aspect of democratic equality. In this case, elevating one's group's status amounted to diminishing America's racial hierarchies. This is surely a motivation compatible with egalitarian commitments. So, widespread group voting of this type is quite consistent with democratic equality.

So different kinds of group voting interact differently with democratic equality. Voting in order to ameliorate the inferiority of a group is compatible with the positive aspects of equality. One can have attractive egalitarian relationships with people moved by such motivations. But voting in order to protect or produce the superiority of a group clashes with this aspect. This type of voting manifests a lack of commitment to equality. One cannot have a civic friendship with those who wholly lack such commitments and one's civic friendships are impaired with those who have only very weak such commitments. So, how group

voting impacts the positive aspects of equality depends on the type of group voting in play. Now that does not mean that group voting impacts the negative aspect of democratic equality. I doubt it does. Group voting, by itself, never puts people into relationships of subordination. But it can prevent relationships of civic friendship. It thus impairs the positive, but not the negative, aspect of democratic equality.

Let me conclude the section by returning to an issue I raised in section 1. We have been exploring the question of how the prevalence of different voter motivations impacts democratic values. But these motivations are often combined in individuals: often, single voters are moved to some extent by all three types of motivation. How does that affect our discussion? To account for this, the key thing we need to be able to do is evaluate how much each motivation matters on average. The larger the average impact of issue voting, and to a lesser extent performance voting, the better positioned a democracy is to achieve self-rule, and the more citizens' social and political affairs can manifest their joint intentions. The larger the average impact of privileged group identities, the worse positioned a democracy is to achieve the positive aspect of democratic equality, and the more civic friendships are seriously damaged. This, in effect, answers the first question of this paper. Roughly speaking, issue voting is best, followed by performance voting, followed by group voting. And that answer puts us in a better position to assess how voters' motivations affect the value of American democracy. But we are not yet in a quite good enough position. For how these motivations matter to democratic values depends on how competent voters are. So we now turn to voter competence.

#### 4. VOTER COMPETENCE

Let us say that someone is competent with respect to a certain aim when they reliably do what promotes that aim. They do what promotes that aim in many contexts. Let us say that voters are competent insofar as they are competent with respect to the aims that underlie their vote. In this section, we will look at how voter competence modulates the contribution those aims make to democratic values. This is crucial for two reasons. First, it tightens our grip on how voter motivation and democratic values relate. It tells us when certain motivations successfully contribute to those values. Second, we need to do this to understand how voter motivation contributes to the value of American democracy. There are well-known doubts about the competence of American voters.<sup>40</sup> If voter motivation only contributes to democratic values when voters are sufficiently competent,

40 See, for example, Brennan, *The Ethics of Voting*, ch. 7.

then that matters to our assessment of that contribution. So, what is required of voter competence for voter motivation to contribute to democratic values?

It depends on the value. Let us start with self-rule. Suppose voters want to vote for the candidate who would perform best, but imagine that they are utterly incompetent. They judge candidates on the basis of good looks or how well tank helmets fit on their heads.<sup>41</sup> But head size does not predict which candidate will be the best performer. Yet suppose the lucky thing happens: a majority of voters do end up voting for the best performer. As previously noted, intuitively this means that the good performance does not manifest their joint intentions in the sense necessary for self-rule. For this type of manifestation, their vote and the good performance has to be more reliably connected. Voters, in general, have to be competent in order for the value of self-rule to be achieved. Now, that is not to say that there is a sharp cutoff at which they achieve the anointed standard of competence. Rather we should think of it in scalar terms. The more competent voters are, the more of the value of self-rule they can attain. So, when voters are quite incompetent, their issue and performance voting contribute little to self-rule.

I think this point is clear in personal cases. Imagine that you start a business. But, let us suppose, you are not a very good businessperson. You hire layabouts, invest in fads, advertise on Myspace. Left to your own devices, you would quickly run your new business into the ground. But, fortunately for you, you are a Rockefeller. And your indulgent uncle is both a very good businessman and very, very rich. He works behind the scenes to rectify your mistakes. He hires hard workers. He contacts the right politicians. He intimidates your competitors (he is a Rockefeller, too). This makes your business a moderate success. In this case, it seems to me that you are not the author of this success. That is because you were so unreliable at achieving it. You were only saved by fortuitous family connections. So, that success does not really redound to your credit. In this personal case, incompetence seems to undercut the achievement of authorship. That is evidence that, in the political case, incompetence also undercuts the achievement of authorship. When people are not competent with respect to their goals, in both cases, they are less the authors of those goals. The achievement of those goals merely matches, rather than manifests, their intentions.

Let us turn to equality. Here the key question is whether incompetent voting is incompatible with the attitudes that the positive egalitarian value requires. If you are incompetent, does that imply that you lack a commitment to equality? Does it imply that you do not care appropriately about your fellow citizens' wel-

41 Good looks do sometimes drive vote choice. For some recent evidence, see Ahler et al., "Face Value?" The import of head size turns on your take on Michael Dukakis's ill-fated presidential push.

fare? At first glance, the answer seems to be a clear no. One can have goals one is not good at achieving. Suppose you care deeply about your nephew's welfare. But your nephew lives in England and you live in the United States. You just cannot keep up with his life. The tyranny of distance defeats you. So you never get him the right Christmas presents. You get him films when he wants games, sugar candy when he wants chocolate, scarves when he wants "jumpers." You are not very good at contributing to his welfare. But that does not imply that you do not care about his welfare. You can care about things you are not very good at promoting. So, at first glance, voter competence need not matter to how voter motivations impact democratic equality.

But perhaps first glances deceive. There are cases where your incompetence does make for a lack of concern. Suppose you could easily find out what your nephew wanted. You just need to phone your sister. Then your incompetence suggests that you do not care that much about your nephew's welfare. Your unwillingness to pick up the phone in part constitutes a lack of substantial concern. Two things are going on in this case. First, it is not very costly to become competent. You just need to dial your sibling. Second, this minor cost really boosts the chances of achieving the relevant goal. Calling your sister will make you much more likely to give your nephew good presents. So, when increasing your competence is relatively easy, and would substantially improve the chances of achieving some goal, lack of competence constitutes your not putting much weight on the goal at all.

But voting meets neither condition. It is not easy to become a very competent voter. You have to spend a lot of time reading things like *Politico* and *The New York Times*. That is all time stolen from other, more valuable activities. And, more importantly, there is little chance that such competence will make a difference to the welfare of your fellow citizens. This is because there is so little chance your vote will make a difference. Rarely do individual votes decide elections. Even if you were the most competent voter in the world, that would in expectation yield a tiny benefit to your fellow citizens. So, I suspect that you can be an incompetent voter while having the attitudes that the positive egalitarian value requires. Incompetence does not constitute a failure to care enough about your fellow citizens' welfare or to be committed to equality. So self-rule is only achieved by reasonably competent voters. But the positive aspect of democratic equality imposes minimal standards of voter competence.

##### 5. THE AMERICAN VOTER

We can now see how the motivations of the American voter contribute to democratic values. This tells us, in part, the extent to which American democracy



achieves these values. We first address self-rule. I have already suggested that issue voters are scarce. If this is true, then only performance voters can realize this value. How many of those are there? Well, when you ask voters what they like about candidates, about 40 percent mention performance issues. About 30 percent mention topics like the economy. Up to 10 percent mention candidates' personal qualities.<sup>42</sup> So this seems an upper bound for the number of performance voters in the American electorate. And it is a respectable upper bound: 40 percent of voters is a lot of voters.

Yet, unfortunately, I doubt these voters enjoy much of the value of self-rule. The problem is that many voters are rather incompetent. To see this, we draw from Christopher Achen and Larry Bartels's great book, *Democracy for Realists*. They argue, persuasively, that American performance voters are "myopic" and "blind." They are myopic in the sense that they only vote retrospectively on short-term performance. They are blind in the sense that they punish incumbents for things out of their control. Fixating on short-term performance and kicking incumbents for acts of God are not, I suspect, reliable ways to pick good performers. So I suspect American performance voters are not competent performance voters. Insofar as these suspicions are accurate, American voters will not achieve much of the value of self-rule.

What is the evidence for voters' myopia and blindness? Let us start with myopia. Now everyone knows that economic performance correlates with incumbent vote share. But economic performance can be different over different time periods. It might be good over four years, but less good over the last two years. So Achen and Bartels test what period of economic performance is associated with incumbent vote share.<sup>43</sup> They find that an extra percentage of real income growth in the six months before the presidential election is associated with a large increase in incumbent popular vote margin: seven-and-a-half percentage points. Income growth at other times, they find, is not associated with any change in the incumbent's vote margin. Achen and Bartels conclude that voters are just responding to economic conditions around the time they are voting. If that is right, then voters only care about what you have done for them lately. They are myopic, in the sense that they fixate on the recent past.<sup>44</sup>

Now turn to blindness. Again, the best piece of evidence for voter blindness

42 Lewis-Beck et al., *The American Voter Revisited*, ch. 10.

43 Achen and Bartels, *Democracy for Realists*, 146–76.

44 Healy and Malhotra ("Myopic Voters and Natural Disaster Policy") and Montalvo ("Voting after the Bombings") report similar outcomes with respect to natural disasters and terrorist attacks, respectively. Healy and Lenz argue that this is a manifestation of the "end" part of peak-end effects ("Substituting the End for the Whole").



comes from Achen and Bartels.<sup>45</sup> They point out that the level of rainfall matters to voters' welfare. Low rainfall means drought; high rainfall means flooding. But the weather is outside of incumbents' control. Thus, they investigate how, in the United States, incumbent vote share tracks rainfall. They find very low and very high rainfall is associated with lower incumbent vote share. They conclude that voters are punishing incumbents for something over which they have no control: bad weather. This is not a reliable way to pick good performers. Thus, insofar as Achen and Bartels are correct, American performance voters are not competent. So these voters likely achieve little of the value of self-rule. American democracy, then, can attain little of this value. The American voter, at least by the lights of the evidence I have presented, pushes that value largely out of reach.<sup>46</sup>

But what about democratic equality? In particular, does American voting behavior impair the positive aspect of democratic equality? Let us start with the impact of performance voting. Here competency matters. But the competency constraint I advanced was minimal. Indeed, I think even myopic and blind voters can meet it. After all, myopic and blind voters are not *completely* incompetent: they still managed to kick out Hoover. They just have a low level of competency. But there is a tiny chance that their vote makes a difference. So this low competency is consistent with having the attitudes that the positive egalitarian value demands. It need not mean that voters do not care appropriately about their fellow citizens or are not sufficiently committed to equality. The lack of competency evinced by American voters, then, does not much matter to democratic equality.

Let us turn to group voting. Here the outlook is much gloomier. The first problem arises from the pervasive impact of partisan identification on voting behavior. I noted above that voting on the basis of party identification involves voting in order to elevate your own social group above other social groups. It is a case of creating superiority. That is incompatible with a commitment to social equality. This is bad news for the positive value of equality in American democracy. Partisans on each side are trying to make themselves superior to those on the other. They cannot at the same time forge valuable egalitarian relationships across party lines. Substantively, that is of enormous import. Partisan identification is probably the strongest influence on voting behavior.<sup>47</sup> Since it severs

45 Achen and Bartels, *Democracy for Realists*, 116–46. The corroborating literature is now quite large. For a review, see Healy and Malhotra, “Retrospective Voting Reconsidered.”

46 I defend a similar conclusion, but on different grounds, in Lovett, “Democratic Autonomy and the Shortcomings of Citizens.”

47 For the source of this position, see Campbell et al., *The American Voter*. For recent forceful advocates, see again Achen and Bartels, *Democracy for Realists*, 232–66.

positively valuable egalitarian relationships, only a few such relationships can span party lines. Cross-partisan relationships cannot be civic friendships.

Yet things are worse than that. To see why, we have to look at some more empirical evidence. And we will need to turn to current affairs: we will need to turn to the 2016 election of Donald Trump. One of the most crucial points about Trump's rise is its connection to white identity. In the primaries, white voters more attached to their white identity were much more likely to vote for Trump. He won the general election with a majority of fifteen points among white voters. Again, white identifiers were most likely to vote for him. The reason is not obscure. His rhetoric was littered with both implicit and explicit racial appeals.<sup>48</sup> These appeals helped cement Trump as the candidate of white Americans. He swept to office on a wave of white-identity voting.<sup>49</sup> White-identity voting, as we noted above, is incompatible with civic friendship. You cannot stand in such an egalitarian relationship with someone while trying to cement your superiority over them.

Trump contributed to this wave, but he did not create it. Thirty percent to 40 percent of white Americans say that being white is very, or extremely, important to their identity.<sup>50</sup> And white-identity voting mattered well before Trump. It seems to have reduced the vote for Obama as well as for Black candidates in other elections.<sup>51</sup> For at least a decade, then, millions of white Americans have voted on the basis of protecting their lofty place in America's racial hierarchy. And white people are probably not the only members of a privileged group to vote on the basis of group identity. For example, Trump won by twelve percentage points among men. The more sexist someone was the more likely they were to vote for him.<sup>52</sup> So it seems plausible (although the evidence is less strong) that male identity also mattered to vote choice. In short, group voting in America is not the preserve of oppressed groups. Members of privileged groups often vote on the basis of their group identity.

This is even worse news for the positive aspect of democratic equality. Voting behavior rends positively valuable egalitarian relationships between partisans. And it also seems to, often, prevent them between the more and less privileged. That means those relationships cannot hold between each American citizen. Now that does not mean they cannot hold between anyone. Not every white

48 The *New York Times* keeps a list of Trump's racist comments. See Leonhardt and Philbrick, "Donald Trump's Racism."

49 The story here comes from Jardina, *White Identity Politics*, 230–47.

50 Jardina, *White Identity Politics*, 63.

51 Petrow, Transue, and Vercellotti, "Do White In-Group Processes Matter, Too?"

52 Schaffner, Macwilliams, and Nteta, "Understanding White Polarization in the 2016 Vote for President."

person votes on their white identity. Not every partisan votes on party identity, and those who do not can share positively valuable egalitarian relationships. But millions of people do vote on such bases. So the American voter strikes a blow against the positive aspect of democratic equality. That leaves the negative aspect of democratic equality untouched. It does not by itself make American citizens subordinate to their fellows.<sup>53</sup> But, all the same, it is a big blow to the intrinsic value of American democracy.

#### 6. HOW SHOULD WE VOTE?

We are now in a position to give a partial answer to the third question: How should we vote? The question here concerns how actual American citizens should vote, given the condition of American democracy. My answer will be partial. We will look at just the reasons democratic values give rise to. I think that the value of self-rule can give rise to two types of reasons with respect to voting behavior. First, it can give rise to a self-interested reason. You yourself benefit from achieving this value. But you only achieve this when your fellow citizens put you in a position to achieve it. They must have the intentions that would underpin a joint intention. And they must have formed those intentions competently. Otherwise it does not matter how you vote. The incompetence of your fellow citizens puts the value of self-rule out of reach. But neither condition is usually met in the United States. American voters, as we have seen, often lack the motivations they need to achieve the value of self-rule. They are often group voters. And those who are performance voters are rarely competent performance voters. So, in the United States, self-rule provides little self-interested reason to vote on particular motivations.

Second, the value of self-rule can give rise to an altruistic reason. Generally, we should help out our fellow citizens. If our doing something helps them achieve some good, we have reason to do the thing. One of the reasons to pay our taxes is that it helps us get good roads, parks, schools. It helps out our co-citizens. Thus, were American voters good competent issue voters, you would have reason to be such a voter yourself. This would help Americans achieve the value of self-rule. But again as we have seen, American voters are not competent issue voters. So being such a voter does not help them achieve self-rule. You can only help those who help themselves. So you lack this altruistic reason to be a competent issue voter. Thus I doubt the value of self-rule gives American voters any reason to vote in certain ways. It would in an ideal democracy. In an ideal

53 I explore a problem for the negative aspect of democratic equality in Lovett, "Must Egalitarians Condemn Representative Democracy?"

democracy it would give American voters reason to be competent issue voters. But in our deeply nonideal, real-world case, it is normatively inert.

Now one might resist this. Suppose you endorse a view like rule consequentialism. On this view, one should act in line with the rules that, were they widely accepted, would lead to the best consequences. So imagine that your college needs a million dollars to stay open. If every member of the college gave the college a thousand dollars it would stay afloat. This would be to great benefit overall. So you should give the college a thousand dollars. And you should do this even when you know you are throwing your money into the abyss; you know that your perfidious colleagues will never chip in. This sort of view says that you should be a competent issue voter despite it achieving nothing. For if everyone accepted the rule “be a competent issue voter,” then we would achieve the value of self-rule. So my position will not be congenial to people with such rule-based moral views. But I am skeptical of such views. The cases at hand are exactly those where they seem to go wrong. In these cases, following such rules seems pointless. So, the relevant cases seem like counterexamples to such views. That is not secure footing from which to resist the position I have put forward.

Let us turn to equality. This gives rise to reasons connected to the constraints on egalitarian relationships. You should not do things that sever your egalitarian relationships. Now, were America entirely devoid of egalitarian civic relationships, this too would not matter. But that is not the picture I just painted. Millions of people may vote on party identification and privileged identities. But millions also do not. You still have reason to avoid severing your egalitarian relationships with these latter people. That means you should not vote on certain group identities. Voting on party identification seems out. Voting on whiteness or masculinity is definitely out. Such voting precludes a commitment to equality. In short, you cannot be the type of group voter who votes on the basis of privileged group identities. Now that does not preclude voting on unprivileged group identities. Ninety-six percent of Black voters voted for Obama. They need not have been doing anything wrong. But it precludes much group voting all the same. So equality imposes constraints on your motivations. Does it also impose constraints on your competence? Only minimal ones. This is because acquiring competence is costly and the chances of it making a difference are low. Thus you need not hit the books to meet the requirements of democratic equality.<sup>54</sup> Equality mainly requires you to manage the motivations underlying your vote.

So we have shed some light on how we ought to vote. Insofar as achieving

54 Jason Brennan, of course, argues that voters have reasons to be competent that are not grounded in self-rule or equality. See Brennan, *The Ethics of Voting*, ch. 3. I have not engaged with his argument here.

democratic values is important, we have reason not to vote on certain motivations. In ideal democracies, this reason would be quite constraining. We would have reason to be competent issue voters. But the nonideal nature of American democracy makes a crucial difference. It means democratic values impose quite lax standards on voting behavior. As long as we do not vote on relatively privileged identities, we are likely doing all that such values require of us. Of course, many of us fall short of even these standards. Many voters vote on white identity. Many more are driven by party identity. But the standard is not, in principle, hard to meet.<sup>55</sup>

## 7. CONCLUSION

Let me sum up. We started with three questions. The first concerned how the prevalence of different kinds of voter motivations mattered to intrinsic democratic values. I have argued that issue voting would be best, followed by performance voting, followed by group voting. The second concerned how much American voters contribute to these values. I have argued: not much. The American voter often lacks the motivations, or the competence, necessary to contribute to either equality or self-rule. The third concerned how Americans should vote. I have argued that Americans need not pretend that they live in an ideal democracy. In their nonideal democracy, they only do wrong by voting on relatively privileged identities. This covers much of the territory of how voter motivation interacts, both evaluatively and deontically, with intrinsic democratic values. It also leaves much of that territory uncovered. But it suffices to show, I think, that voter motivations matter to democracy.<sup>56</sup>

*London School of Economics and Political Science*  
adam.lovett7@gmail.com

55 Now, as I have said, some people have a sunnier view of American voters than I think is accurate. See, for example, Ansolabehere, Rodden, and Snyder, "The Strength of Issues." They should think that American democracy is in a better position to achieve democratic values than I do. And they should also think that American voters are under more stringent obligations than I take them to be. If most voters are competent and vote on the issues, then each voter has reason to be competent and vote on the issues. Thus, empirical premises aside, the theoretical upshot of this section is that there are systematic dependencies between the quality of a democracy and how its voters have reason to vote.

56 For helpful comments on earlier drafts of this paper, I would like to thank Daniel Brinkerhoff-Young, Jane Friedman, Annette Martin, Samuel Scheffler, Daniel Sharp, Jake Zuehl, Daniel Viehoff, and two anonymous reviewers.

## REFERENCES

- Achen, Christopher H., and Larry M. Bartels. *Democracy for Realists: Why Elections Do Not Produce Responsive Government*. Princeton: Princeton University Press, 2016.
- Ahler, Douglas J., Jack Citrin, Michael C. Dougal, and Gabriel S. Lenz. "Face Value? Experimental Evidence that Candidate Appearance Influences Electoral Choice." *Political Behavior* 39, no. 1 (March 1, 2017): 77–102.
- Ansolabehere, Stephen, Jonathan Rodden, and James M. Snyder. "The Strength of Issues: Using Multiple Measures to Gauge Preference Stability, Ideological Constraint, and Issue Voting." *American Political Science Review* 102, no. 2 (May 2008): 215–32.
- Arvan, Marcus. "People Do Not Have a Duty to Avoid Voting Badly: Reply to Brennan." *Journal of Ethics and Social Philosophy* 5, no. 1 (2010): 1–6.
- Berelson, Bernard R., Paul F. Lazarsfeld, and William N. McPhee. *Voting: A Study of Opinion Formation in a Presidential Campaign*. Chicago: University of Chicago Press, 1954.
- Berinsky, Adam J. *In Times of War*. Chicago: University of Chicago Press, 2009.
- Billig, Michael, and Henri Tajfel. "Social Categorization and Similarity in Intergroup Behaviour." *European Journal of Social Psychology* 3, no. 1 (January/March 1973): 27–52.
- Bolsen, Toby, James N. Druckman, and Fay Lomax Cook. "The Influence of Partisan Motivated Reasoning on Public Opinion." *Political Behavior* 36, no. 2 (June 2014): 235–62.
- Bratman, Michael E. "Shared Cooperative Activity." *Philosophical Review* 101, no. 2 (April 1992): 327–41.
- Brennan, Jason. *Against Democracy*. Princeton: Princeton University Press, 2016.
- . "The Ethics and Rationality of Voting." In *The Stanford Encyclopedia of Philosophy* (Winter 2016). <https://plato.stanford.edu/archives/win2016/entries/voting/>.
- . *The Ethics of Voting*. Princeton: Princeton University Press, 2011.
- Campbell, Angus, Philip Converse, Warren E. Miller, and Donald E. Stokes. *The American Voter*. Chicago: University of Chicago Press, 1960.
- Campbell, Angus, Gerald Gurin, and Warren Miller. *The Voter Decides*. Evanston, IL: Row, Peterson, 1954.
- Christiano, Thomas. *The Constitution of Equality: Democratic Authority and Its Limits*. New York: Oxford University Press, 2008.
- . *The Rule of the Many: Fundamental Issues in Democratic Theory*. Boulder, CO: Westview Press, 1996.

- Cohen, Geoffrey L. "Party over Policy: The Dominating Impact of Group Influence on Political Beliefs." *Journal of Personality and Social Psychology* 85, no. 5 (2003): 808–22.
- Converse, Philip E. "The Nature of Belief Systems in Mass Publics (1964)." *Critical Review* 18, no. 1–3 (2006): 1–74.
- Downs, Anthony. *An Economic Theory of Democracy*. Boston: Harper and Row, 1957.
- Fiorina, Morris. *Retrospective Voting in American National Elections*. New Haven, CT: Yale University Press, 1981.
- Gagnon, Andre, and Richard Y. Bourhis. "Discrimination in the Minimal Group Paradigm: Social Identity or Self-Interest?" *Personality and Social Psychology Bulletin* 22, no. 12 (December 1996): 1289–1301.
- Goldman, Alvin I. "Why Citizens Should Vote: A Causal Responsibility Approach." *Social Philosophy and Policy* 16, no. 2 (Summer 1999): 201–17.
- Green, Donald, Bradley Palmquist, and Eric Schickler. *Partisan Hearts and Minds*. New Haven, CT: Yale University Press, 2002.
- Guerrero, Alexander A. "The Paradox of Voting and the Ethics of Political Representation." *Philosophy and Public Affairs* 38, no. 3 (Summer 2010): 272–306.
- Healy, Andrew, and Gabriel S. Lenz. "Substituting the End for the Whole: Why Voters Respond Primarily to the Election-Year Economy." *American Journal of Political Science* 58, no. 1 (January 2014): 31–47.
- Healy, Andrew, and Neil Malhotra. "Myopic Voters and Natural Disaster Policy." *American Political Science Review* 103, no. 3 (August 2009): 387–406.
- . "Retrospective Voting Reconsidered." *Annual Review of Political Science* 16, no. 1 (May 2013): 285–306.
- Jardina, Ashley. *White Identity Politics*. New York: Cambridge University Press, 2019.
- Jones, Bradley. "Republicans and Democrats Have Grown Further Apart on What the Nation's Top Priorities Should Be." Pew Research Center Blog, February 5, 2019. <https://www.pewresearch.org/fact-tank/2019/02/05/republicans-and-democrats-have-grown-further-apart-on-what-the-nations-top-priorities-should-be>.
- Key, V. O. *The Responsible Electorate: Rationality in Presidential Voting, 1936–1960*. Cambridge, MA: Belknap Press of Harvard University Press, 1966.
- Kiewiet, D. Roderick, and Michael S. Lewis-Beck. "No Man Is an Island: Self-Interest, the Public Interest, and Sociotropic Voting." *Critical Review* 23, no. 3 (September 1, 2011): 303–19.
- Kinder, Donald R. "Opinion and Action in the Realm of Politics." In *The Hand-*



- book of Social Psychology*, edited by Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey, 778–867. Boston: McGraw-Hill, 1998.
- Kinder, Donald R., and Nathan P. Kalmoe. *Neither Liberal nor Conservative: Ideological Innocence in the American Public*. Chicago: University of Chicago Press, 2017.
- Kinder, Donald R., and D. Roderick Kiewiet. “Sociotropic Politics: The American Case.” *British Journal of Political Science* 11, no. 2 (April 1981): 129–61.
- Kolodny, Niko. “Rule Over None I: What Justifies Democracy?” *Philosophy and Public Affairs* 42, no. 3 (Summer 2014): 195–229.
- . “Rule Over None II: Social Equality and the Justification of Democracy.” *Philosophy & Public Affairs* 42, no. 4 (Fall 2014): 287–336.
- Kramer, Gerald H. “Short-Term Fluctuations in U.S. Voting Behavior, 1896–1964.” *American Political Science Review* 65, no. 1 (March 1971): 131–43.
- Lenz, Gabriel S. *Follow the Leader? How Voters Respond to Politicians’ Policies and Performance*. Chicago: University of Chicago Press, 2012.
- Leonhardt, David, and Ian Prasad Philbrick. “Donald Trump’s Racism: The Definitive List, Updated.” *New York Times*, January 15, 2018. <https://www.nytimes.com/interactive/2018/01/15/opinion/leonhardt-trump-racist.html>.
- Lewis-Beck, Michael S., William F. Jacoby, Helmut Norpoth, and Herbert F. Weisberg. *The American Voter Revisited*. Ann Arbor, MI: University of Michigan Press, 2008.
- Lewis-Beck, Michael S., and Mary Stegmaier. “Economic Voting.” In *The Oxford Handbook of Public Choice*, edited by Roger D. Congleton, Bernard N. Grofman, and Stefan Voigt, 247–62. Oxford: Oxford University Press, 2018.
- Locksley, Anne, Vilma Ortiz, and Christine Hepburn. “Social Categorization and Discriminatory Behavior: Extinguishing the Minimal Intergroup Discrimination Effect.” *Journal of Personality and Social Psychology* 39, no. 5 (1980): 773–83.
- Lovett, Adam. “Democratic Autonomy and the Shortcomings of Citizens.” *Journal of Moral Philosophy* 18, no. 4 (August 2021): 363–86.
- . “Must Egalitarians Condemn Representative Democracy?” *Social Theory and Practice* 47, no. 1 (January 2021): 171–98.
- Lupia, Arthur, and Mathew McCubbins. *The Democratic Dilemma*. New York: Cambridge University Press, 1998.
- Mason, Lilliana. *Uncivil Agreement: How Politics Became Our Identity*. Chicago: University of Chicago Press, 2018.
- Melzer, Scott. *Gun Crusaders: The NRA’s Culture War*. New York: New York University Press, 2012.
- Montalvo, José G. “Voting after the Bombings: A Natural Experiment on the Ef-



- fect of Terrorist Attacks on Democratic Elections.” *Review of Economics and Statistics* 93, no. 4 (November 2011): 1146–54.
- Parfit, Derek. *Reasons and Persons*. Oxford: Oxford University Press, 1984.
- Petersen, Michael, Martin Skov, Soren Serritzlew, and Thomas Ramsoy. “Motivated Reasoning and Political Parties: Evidence for Increased Processing in the Face of Party Cues.” *Political Behavior* 35, no. 4 (December 2013): 831–54.
- Petrow, Gregory A., John E. Transue, and Timothy Vercellotti. “Do White In-Group Processes Matter, Too? White Racial Identity and Support for Black Political Candidates.” *Political Behavior* 40, no. 1 (2018): 197–222.
- Pettit, Philip. *On the People’s Terms: A Republican Theory and Model of Democracy*. New York: Cambridge University Press, 2012.
- Popkins, Samuel. *The Reasoning Voter*. Chicago: Chicago University Press, 1991.
- Quong, Jonathan. *Liberalism without Perfection*. Oxford: Oxford University Press, 2011.
- Rawls, John. *Political Liberalism*. New York: Columbia University Press, 1993.
- Raz, Joseph. *The Morality of Freedom*. Oxford: Oxford University Press, 1986.
- Rousseau, Jean-Jacques. *The Social Contract*. Harmondsworth, UK: Penguin, 1968.
- Schaffner, Brian F., Matthew Macwilliams, and Tatische Nteta. “Understanding White Polarization in the 2016 Vote for President: The Sobering Role of Racism and Sexism.” *Political Science Quarterly* 133, no. 1 (Spring 2018): 9–34.
- Scheffler, Samuel. “The Practice of Equality.” In *Social Equality: On What It Means to Be Equals*, edited by Carina Fourie, Fabian Schuppert, and Ivo Walilman-Helmer, 21–44. Oxford: Oxford University Press, 2015.
- Schwarzenbach, Sibyl A. “On Civic Friendship.” *Ethics* 107, no. 1 (October 1996): 97–128.
- Stilz, Anna. *Liberal Loyalty: Freedom, Obligation, and the State*. Princeton: Princeton University Press, 2009.
- . “The Value of Self-Determination.” In *Oxford Studies in Political Philosophy*, edited by Peter Vallentyne, David Sobel, and Steven Wall, 98–128. Oxford: Oxford University Press, 2016.
- Tajfel, Henri, M. G. Billig, R. P. Bundy, and Claude Flament. “Social Categorization and Intergroup Behaviour.” *European Journal of Social Psychology* 1, no. 2 (April/June 1971): 149–78.
- Tajfel, Henri, and John C. Turner. “The Social Identity Theory of Intergroup Behavior.” In *Political Psychology: Key Readings*, by John T. Jost and Jim Sidanius, 276–93. New York: Taylor & Francis, 2004.
- Tocqueville, Alexis de. *Democracy in America*. New York: Harper, 1988.

- United Nations. "International Covenant on Civil and Political Rights," December 16, 1966. <https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx>.
- Viehoff, Daniel. "Democratic Equality and Political Authority." *Philosophy and Public Affairs* 42, no. 4 (Fall 2014): 337–75.
- . "Power and Equality." In *Oxford Studies in Political Philosophy*, vol. 5, edited by David Sobel, Peter Vallentyne, and Steven Wall, 3–37. Oxford: Oxford University Press, 2019.
- Zaller, John. *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press, 1992.
- Zuehl, Jake. *Collective Self-Determination*. PhD dissertation, Princeton University, 2016.

## WHY PATERNALISTS MUST ENDORSE EPISTOCRACY

*Jason Brennan and Christopher Freiman*

RECENT FINDINGS from psychology and behavioral economics suggest that we are “predictably irrational” in the pursuit of our interests.<sup>1</sup> Paternalists from both the social sciences and philosophy use these findings to defend interfering with people’s consumption choices for their own good.<sup>2</sup> We should tax soda, ban cigarettes, and mandate retirement savings to make people healthier and wealthier than they would be on their own.

While there is an extensive literature arguing for paternalistic interference with people’s consumption choices, little has been said on behalf of paternalistic interference with people’s *voting* choices. Brennan’s work in defense of epistocracy, for instance, focuses on the ways in which incompetent voters wrongly harm others.<sup>3</sup> Our thesis is instead that the standard arguments offered in support of restricting someone’s consumption choices for their own good also imply support for restricting someone’s voting choices for their own good. Indeed, the case for paternalistic restrictions on voting choices is in many ways *stronger* than the case for restricting personal consumption choices. So, paternalists face a dilemma: either endorse less interference with consumption choices or more interference with voting choices. Note that we do not take a stand here on whether paternalism or epistocracy is justified; we are merely arguing that paternalists should, on pain of inconsistency, also accept a strong presumption in favor of epistocracy.

We begin with a sketch of the social scientific research on cognitive bias and its effects on decision making (section I). From there we explore how this research informs recent philosophical defenses of paternalism: due to the pervasiveness of cognitive bias, paternalists claim, the state will frequently be positioned to better advance the aims of citizens than citizens themselves (section II). Next, we show that the same considerations that purportedly count in favor

1 Ariely, *Predictably Irrational*.

2 See, e.g., Thaler and Sunstein, *Nudge*; Conly, *Against Autonomy*; and Hanna, *In Our Interest*.

3 See, e.g., Brennan, *Against Democracy*.

of paternalistic interference with citizens' consumption choices also count in favor of paternalistic interference with citizens' voting choices (section III). We then consider a variety of objections, including the claim that political liberties occupy a special status that shields them from coercive restriction (section IV). In closing, we acknowledge that the extent to which paternalists ought to endorse interference with the vote is an empirical question but insist that they are committed to such interference in principle (section V).

## I

Most of the recent arguments for paternalistic interference with consumption choices begin by identifying ways in which people fail to be competent judges and pursuers of their own interests. One issue could be a simple lack of information or the presence of misinformation. Maybe you do not save enough for retirement because you do not understand the power of compound interest. Or perhaps you drink too much Mountain Dew because you do not know the calorie count of a sixty-four-ounce Big Gulp. People fail to obtain vaccines because they are misinformed about a supposed link between vaccines and autism, or about the dangers of thimerosal and other chemicals.

Another obstacle to competent decision making is cognitive bias. A cognitive bias is a systematic deviation from rational thought. A bias prevents a person from believing what she ought to believe in light of the evidence and information she possesses. In some cases, consumers might possess the relevant information but fail to use it to form true beliefs.

Sarah Conly, for instance, discusses the problem of the optimism bias within personal choice.<sup>4</sup> People systematically underestimate their chances of being harmed by risky behavior like driving without a seat belt. As a result, we may play things less safe than we would if we made a sober assessment of the risks.

Consider also the case of "present bias"—we find it harder to endure a wait for a reward *right now* than to endure that same wait in the future.<sup>5</sup> This bias may explain why we are not as healthy and wealthy as we would like to be. For instance, if someone offers to give you \$10,000 in ten years or \$100,000 in twenty years, you will probably wait an extra ten years to get the extra \$90,000. But if someone offers to give you \$10,000 *today* or \$100,000 in ten years, many people will take the money now and not wait the extra ten years to get the extra \$90,000. This tendency would help explain why many people save too little for retirement and pile up credit card debt. The same problem applies to decisions about health.

4 Conly, *Against Autonomy*, 22.

5 See, for instance, O'Donoghue and Rabin, "Present Bias."

If you have to fill out a dinner card months before your friend's wedding, you might request "no dessert" because you figure the taste of the cake is not worth the blood-sugar spike. But once you are at the wedding and the server mistakenly puts a slice of cake in front of you, you are far more likely to eat it. Motivated reasoning can also lead us to make poor decisions. Relapsed smokers, for instance, are more likely than successful quitters to accept positive beliefs about smoking—beliefs that rationalize their unhealthy behavior.<sup>6</sup>

For many decades in the middle of the twentieth century, economists often modeled consumers and producers as perfectly rational utility maximizers. The development of behavioral economics offered an empirical challenge to this model, and by extension to some of the normative positions economists and philosophers took in response to that model. In fact, along some dimensions, individuals make predictable mistakes.

Just how robust these behavioral economics findings are, and in particular how much laboratory results spill over into genuine market behavior, is a matter of considerable debate. We take no stance here, although of course paternalists' arguments often rely upon such findings. For instance, Conly writes:

The ground for valuing liberty is the claim that we are pre-eminently rational agents, each of us well suited to determining what goes in our own life. There is ample evidence, however, from the fields of psychology and behavioral economics, that in many situations this is simply not true. The incidence of irrationality is much higher than our Enlightenment tradition has given us to believe, and keeps us from making the decisions we need to reach our goals. The ground for respecting autonomy is shaky.<sup>7</sup>

We will note that even standard neoclassical economics often incorporates some of the findings of behavioral economics by accounting for the *costs* of rationality. Consider that activities such as thinking carefully or scientifically, engaging in mathematical calculations rather than relying upon gut heuristics and shortcuts, and working to overcome bias, are expensive rather than cost-free behaviors; they require time, effort, and other resources. The neoclassical model predicts that people will "spend" on information and rationality only if the expected benefits exceed the expected costs. When decisions are immediate with all or nearly all the consequences felt immediately, then over time people will be disciplined to make better, more rational decisions. However, the more the consequences of decisions are attenuated or the power of individual choices weakened, the

6 Fotuhi, Fong, Zanna, et al., "Patterns of Cognitive Dissonance—Reducing Beliefs among Smokers."

7 Conly, *Against Autonomy*, 2.

weaker the connection between outcomes and choices. The weaker the feedback signal, the less that reality disciplines individuals to become ever more rational. This explains, perhaps, why people learn quickly how to order food they will find tasty, take far longer to learn to be rational in choosing mates rather than rely entirely on gut feelings and heuristics, and seem to be even worse at planning long term for retirement. It explains why they can and do indulge irrational beliefs about vaccines (since vaccination is partly a collective action problem) but less often irrational beliefs about driving or crossing the street.

## II

The implication of these sorts of findings is that people tend to smoke more and buckle their seat belts less often than they would if they had accurate beliefs about the risks. Paternalists then argue that policies like cigarette bans and seat-belt mandates can help people satisfy the preferences they *would* have were their beliefs accurate and if they reasoned about that information in a bias-free way. If people *really* understood how high the health costs of smoking are, very few would take up smoking. The ban simply helps people get what they really want and avoid what they do not really want.

Further, it may be possible to accommodate people who genuinely prefer to smoke despite the high health risks. A cigarette tax in effect shifts some of the high long-term costs of smoking forward, forcing the potential smoker to bear them today.<sup>8</sup> Remember, one of the predictions is that consumers will be less rational in their choices when the costs are attenuated or distant. On some margin, this reduces smoking, though it appears that very high tax rates are necessary to produce significant reductions in smoking.<sup>9</sup> Even if a polity institutes a ban on cigarette consumption, it might allow people to “opt out” of the ban by performing some relatively expensive procedure, such as having to apply for a smoking license after passing a test on the health risks of smoking.

To further motivate the case for paternalism, consider the following example from Conly.<sup>10</sup> Suppose the person next to you is about to drink antifreeze because they think it is blue Gatorade. You can try persuading them that it really is antifreeze, but if that does not work, you should forcibly prevent them from

8 For a recent, informative discussion of “sin taxes,” see Allcott, Lockwood, and Taubinsky, “Regressive Sin Taxes, with an Application to the Optimal Soda Tax.”

9 Bader, Boisclair, and Ferrence, “Effects of Tobacco Taxation and Pricing on Smoking Behavior in High Risk Populations”; Callison and Kaestner, “Do Higher Tobacco Taxes Reduce Adult Smoking?”

10 Conly, *Against Autonomy*, 3.

drinking it. Here again, you are just helping them achieve their own ends—they do not want to drink antifreeze, they want to drink Gatorade. You are enabling them to satisfy the preferences they *would* have were they to possess more accurate beliefs about the facts.

To extend that example, suppose your friend actively tries to investigate whether the drink in her hand is Gatorade or antifreeze. She goes so far as to conduct chemical tests on the drink before drinking. However, she is so bad at chemistry that she concludes the propylene glycol in the liquid is actually sucrose, and so tries to drink it. Here, she *tries* to be scientific but fails. Again, you would feel justified in swatting the drink out of her hands.

This example suggests that paternalistic measures, such as cigarette bans, that coercively prevent people from making self-harming choices may be justified. But you could endorse milder forms of paternalism, such as Richard Thaler and Cass Sunstein's "nudges."<sup>11</sup> Maybe cafeterias should put healthier food in easier-to-reach places. You can still get a donut, but you will have to work for it.

Although critics may worry that paternalism is insulting or degrading to citizens—implying that they are not competent to make decisions about their own lives—Conly argues that this is not the case. As she puts it:

We don't regard it as insulting to assume that the man on the street can't do quantum mechanics, because he can't (unless you're on a very special street). The paternalist believes it is the facts that suggest a change in the status we accord people, a change from what we might have thought about ourselves to a more realistic acceptance of our inabilities. The suggestion here is simply that we should treat people in accordance with their real abilities and their real limitations.<sup>12</sup>

It is not degrading to face up to our own fallibilities and seek to help each other overcome them when we are in a position to do so.

Jason Hanna takes a similar position as Conly. For you to tell an obese person you randomly see in Target that he needs to lose weight may indeed be insulting. However, for a doctor to calmly and politely inform a patient that his obesity is harming him is not. Further, Hanna argues, if prominent physicians declare obesity a national health crisis, on the basis of overwhelming scientific evidence, and then issue diet and exercise recommendations, it is unclear why this would have an insulting message. If the national government then issues policy changes on that basis—such as taxes incentivizing healthy choices—it is unclear why

11 Thaler and Sunstein, *Nudge*.

12 Conly, *Against Autonomy*, 41.

this would suddenly become insulting if the underlying message on which the policies were based is not insulting.<sup>13</sup>

Hanna says that the most plausible version of this kind of objection to paternalism holds that certain policies objectively express an insulting message regardless of what the policymakers' motives are or how the policies make people feel. However, Hanna argues, the expressive significance of a political policy is not written into the fabric of the universe, but rather is a contingent social construct that depends upon the meaning people attach to it. As policies become commonplace and accepted, they lose their insulting force. Adults might find it degrading at first for the government to mandate that they wear seat belts. After a generation, it feels normal. Indeed, Conly argues that paternalism can actually be quite liberating by taking tedious decisions off our plates.<sup>14</sup> By removing options from our option set, paternalistic regulation allows us to focus on those dimensions of our lives that we find more worthy of attention, or so the paternalist argues.

Although Conly may regard personal autonomy or freedom as being of largely instrumental value, other paternalists may simply hold that liberty is not of *absolute* value. If one has an absolutist view of liberty, then one would never accept a paternalistic interference with a person's freedom simply because it produces better consequences. But this seems implausible. To take John Stuart Mill's familiar case, if we stop you from crossing a bridge because you fail to see that the bridge is unsafe, it is implausible to regard that as wrongful, though we interfere with your movement.<sup>15</sup> A more moderate deontology holds that rights violations or restrictions can be justified not whenever doing so promotes overall utility, but when doing so prevents some sufficiently severe enough harm or causes some sufficiently momentous good. However, Hanna argues, regardless of where one sets the "threshold" for justifying restrictions on liberty, it will be easy to construct at least hypothetical cases where paternalism meets that threshold.<sup>16</sup> We need only imagine cases where paternalism stops someone from suffering a severe harm or helps them obtain a momentous good. Further, the *lighter* the burden a restriction imposes, the easier it is to justify. (For example, requiring adults to wear seat belts is less burdensome than requiring them

13 Hanna, *In Our Interest*, 73.

14 Hanna, *In Our Interest*, 90.

15 Mill says that we do not interfere with your liberty at all, since we do not stop you from doing what you want to do. You want to cross the bridge, not fall into the water. We do not actually stop you from crossing the bridge.

16 Hanna, *In Our Interest*, 118–44.



to save 5 percent of their income, which is less burdensome than mandating that they eat kale and quinoa.)

Of course, there are many reasons why one might reject paternalistic interference. Perhaps what appears to be cognitive bias at first blush will turn out not to be bias after all. For instance, the “sunk cost fallacy” may not be fallacious in some cases.<sup>17</sup> Suppose you are deciding whether or not to change your college major. That you have already invested in a particular course of study might supply evidence that you will find your current major to be the right one for you at some point again in the future. Similar arguments may apply to other apparent biases.<sup>18</sup>

Another strategy for resisting paternalism appeals to deontic considerations. Evidence suggests that newlyweds suffer from optimism bias when surveyed about the prospects for their marriage.<sup>19</sup> Even if the state could coercively arrange—or even nudge—marriages for citizens more successfully than the citizens themselves could, you might find such paternalistic interference with their choices to be impermissibly disrespectful of their autonomy.<sup>20</sup>

There is also a consistency objection to paternalism: that people are systematically irrational does not speak in favor of paternalism because paternalistic regulators are themselves systematically irrational.<sup>21</sup> Perhaps individuals are prone to undervalue future gains and overvalue present gains, but the same tendency presumably holds true of bureaucrats and legislators. Thus, we have at least a *prima facie* reason to doubt that paternalist intervention will succeed in correcting the biases it is introduced to correct. The emerging field of behavioral political economy argues in this vein.

We take no stand on whether consumption paternalism is ultimately justified. Rather, ours is a point about symmetry: we claim that paternalists about consumption should also favor paternalism about voting. Of course, neither form of paternalism may be acceptable. For example, if an apparent bias is not in fact a bias, then this result speaks against both consumption and epistocratic pater-

17 T. Kelly, “Sunk Costs, Rationality, and Acting for the Sake of the Past.”

18 See, e.g., Nebel, “Status Quo Bias, Rationality, and Conservatism about Value”; and Hedden, “Hindsight Bias Is Not a Bias.”

19 Lavner, Karney, and Bradbury, “Newlyweds’ Optimistic Forecasts of Their Marriage.”

20 Along similar lines, Jessica Flanigan writes, “Falling in love is often irrational and motivated by inconsistent and imprudent desires. Though some people may take steps to avoid finding themselves in this irrational state, many of us value the chance to make such an irrational decision even when we know it might have disastrous consequences. So it does not follow from the fact that a particular choice is irrational that people have rational reasons to want to avoid it, all things considered” (“Seat Belt Mandates and Paternalism,” 307).

21 Rizzo and Whitman, *Escaping Paternalism*.

nalism in that instance. The challenge for consumption paternalists who wish to avoid epistocratic paternalism is to argue for an asymmetry between the two cases. In the next section, we explain why the problem of cognitive bias is at least as troublesome for voting choices as consumption choices.

### III

There is a large literature on paternalism and a growing literature on restrictions on the vote, but the two have not come into much contact. As noted, paternalists tend to consider restrictions on consumption choices. Much of the debate on epistocratic restrictions on suffrage focuses on the threat that incompetent voters pose to others rather than to themselves.<sup>22</sup> For instance, Brennan analogizes incompetent voters to incompetent jurors or surgeons who impose serious risks on innocent people.<sup>23</sup> The right to vote not only gives you a say over your own life, but the lives of others—and this sort of authority requires justification. But may the state interfere with a citizen's vote for that citizen's own good? The considerations that speak in favor of paternalistic restrictions on consumption choices seem to speak as strongly in favor of paternalistic restrictions on voting choices.

One prominent line of argument in defense of democracy alleges that people tend to be the best judges of their own interests. Consider, for instance, these remarks from Samuel Freeman:

The rule of law, representative assemblies (elected and non-elected), separation of powers, and the convention that government acts solely as representative of the people, are all institutional expressions of the public nature of political power. Democracy, or a universal franchise with equal rights of political participation, is a natural extension of this idea; for if what affects all concerns all, and *assuming that adults are normally best situated to understand and advance their own interests*, then it is natural to conclude that each person ought to have a share of political authority to better ensure that no one's basic rights are undermined or interests are neglected in political procedures.<sup>24</sup>

Surprisingly, Conly's own confidence in democratic institutions is unshaken by

22 E.g., López-Guerra defends epistocracy (he uses the term "aristocracy") on the grounds that incompetent voters may unjustly harm others (*Democracy and Disenfranchisement*). Other defenses of epistocracy appeal to general improvements in electoral quality. See, for instance, Mulligan, "Plural Voting for the Twenty-First Century."

23 Brennan, "The Right to a Competent Electorate."

24 Freeman, "Illiberal Libertarians," 122, emphasis added.

her thorough review of the pervasiveness of cognitive bias in human decision making. She claims that paternalistic “legislation, like all legislation, is best made under a democratically elected and accountable legislature under conditions of transparency.”<sup>25</sup>

However, the assumption of voter competence is even more doubtful than the assumption of consumer competence. *A priori*, we would expect that every flaw in consumers to be worse in voters because the expected cost of an uninformed and biased consumption choice is higher than an uninformed and biased voting choice. A consumer bears most of the cost of their decision to smoke. But unlike consumers, voters never have unilateral decision-making power. Their votes are thrown in with everyone else’s. Except in very tight elections, how individual voters vote (or whether they vote at all) has almost no effect; the expected utility of voting one way is the same as voting the other.

To illustrate, imagine a professor tells her students in a five-hundred-person Econ 101 lecture that fourteen weeks from now they will take a final exam worth 100 percent of their grade. Instead of getting their individual grades, she will average all their scores together and everyone will receive the same equal score. One would expect—pending heroic efforts to overcome a collective action problem—that average grade would be an  $F$ , and that students would work less to overcome whatever biases and prejudices they harbor in economic reasoning (say, for instance, the exponential growth bias that causes them to systematically misunderstand compounding).<sup>26</sup> Students would be rationally ignorant and rationally biased, meaning that, in light of the professor’s rules, the expected costs of overcoming bias and learning the materials exceed the expected benefits. In short, investing in getting one’s answers right is instrumentally irrational given that such an investment is highly unlikely to change their grade. Democratic elections have roughly same basic structure and thus the same incentives as this final exam.<sup>27</sup> While the stakes are higher, the number of other “students” (i.e., voters) is also much higher.

A massive body of evidence, collected over seventy years, indicates that the majority of voters are uninformed. We will spare you the details, but voters tend to be ignorant of political matters ranging from their local representative, which party controls Congress, or changes in economic performance, to changes in

25 Conly, *Against Autonomy*, 39. Elsewhere she writes, “What we need is a democratically elected government, but one in which the government is allowed to pass legislation that protects citizens from themselves, just as we now allow legislation to protect us from others” (*Against Autonomy*, 2).

26 Stango and Zinman, “Exponential Growth Bias and Household Finance.”

27 See, e.g., Caplan, *The Myth of the Rational Voter*.

social indicators such as unemployment, recent changes in legislation, or the branches of government.<sup>28</sup> They are not simply ignorant; rather, voters many have systematically mistaken beliefs about both basic political facts as well as basic social-scientific issues.<sup>29</sup>

Voters, like consumers, are also subject to a variety of biases. Some biases are the same as those at play in the marketplace. Take motivated reasoning. Plenty of studies show that political partisans are selectively skeptical—they will accept evidence that confirms their preexisting policy commitments and reject evidence that threatens them.<sup>30</sup> Just as a consumer may be motivated to rationalize their preference for an expensive luxury car, voters are motivated to rationalize their preference for the platform of their favorite party. So even when they are presented with relevant information, these voters will not update their beliefs appropriately.

Consider also the present bias discussed earlier. Suppose two candidates are running for president: Sensible and Reckless. Sensible proposes the immediate installation of a carbon tax to start tackling the problem of climate change. Yes, the tax will impose short-term economic pain but it is for the sake of long-term gain. Reckless argues that no immediate action needs to be taken. He downplays the urgency of the threat of climate change and floats the idea of nonbinding, vague emissions targets to be met at some unspecified point in the future. Voters biased toward the present will tend to prefer Reckless because he promises small but immediate benefits, despite the policies working toward voters' long-term disadvantage.

Voters also suffer from availability bias.<sup>31</sup> The easier it is for us to think of something, the more common we think that thing is. The easier it is for us to think of an event occurring, the more significant we assume the consequences will be. We are thus terrible at statistical reasoning. Vivid things—plane crashes, shark attacks, terrorist attacks, Ebola—come to mind easily, so we assume these things are much more common than they are. Things that are not vivid—deaths from the flu or pneumonia—do not come to mind easily, and so we wrongly conclude these things are uncommon. This bias can cause voters to ignore less

28 See, e.g., Somini, *Democracy and Political Ignorance*; Delli-Carpini and Keeter, *What Americans Know about Politics and Why It Matters*; Brennan, *Against Democracy*; and Achen and Bartels, *Democracy for Realists*.

29 Bartels, "Uninformed Votes"; Althaus, *Collective Preferences in Democratic Politics*; Caplan, *The Myth of the Rational Voter*; Gilens, *Affluence and Influence*; and Caplan, Crampton, Grove, and Somini, "Systematically Biased Beliefs about Political Influence."

30 Erisen, Lodge, and Taber, "Affective Contagion in Effortful Political Thinking"; Lodge and Taber, *The Rationalizing Voter*; Westen, Blagov, Harenski, et al., "The Neural Basis of Motivated Reasoning"; Westen, *The Political Brain*.

31 Tversky and Kahneman, "Availability."

vivid but real harms—such as the dangers of climate change or the common flu—and instead support policies aimed at less common harms. For instance, even though deaths from terrorism are rare—only about 3,500 Americans have died from terrorist attacks in the past sixty years—Americans nevertheless support a “war on terror.” The Watson Institute at Brown University estimates the total real monetary costs of the wars on terror at \$5.9 trillion.<sup>32</sup> John Mueller and Mark Stewart say that to justify the expense of the Homeland Security Administration, it would need to prevent nearly seventeen hundred major terrorist events per year, which of course it does not.<sup>33</sup> Most voters are undeterred.

Voters also are strongly influenced by “framing effects.”<sup>34</sup> How they respond to survey questions (including how they describe their own ideologies or political beliefs) and how they vote in democratic referenda depend strongly on how the questions are written. Voters can appear to change their mind, e.g., going from overwhelmingly supportive of government social insurance to opposing it, or from supporting capitalism or socialism, simply by substituting one word for a synonym in a poll. A psychologically savvy person—a pollster, newscaster, pundit, politician, moderator in a deliberative forum, or person writing up a referendum question on a ballot—can take advantage of framing effects to induce voters to support the manipulator’s favored position. The problem is so pervasive that some political scientists claim that most voters are largely “innocent” of ideology.<sup>35</sup>

To be clear, we do not mean to suggest that the problem of political bias is immutable. There may be ameliorative steps that voters can take. One suggestion is to make use of heuristics—a voter may recognize her own susceptibility to bias and therefore defer to the judgment of an impartial expert.<sup>36</sup>

We would like to register two concerns about the use of heuristics, however. First, evidence indicates that our choice of heuristic is itself susceptible to partisan bias.<sup>37</sup> More generally, there is empirical literature on the degree to which heuristics enable otherwise uninformed or irrational voters to vote well, a litera-

32 Watson Institute for International and Public Affairs, “\$5.9 Trillion Spent and Obligated on Post-9/11 Wars.”

33 Mueller and Stewart, *Terror, Security, and Money*.

34 For a thorough review of the dangers of this bias, see J. Kelly, *Framing Democracy*.

35 Kinder and Kalmoe, *Neither Liberal nor Conservative*; Achen and Bartels, *Democracy for Realists*; Converse, “The Nature of Belief Systems in Mass Publics.”

36 For a discussion of the value of shortcuts, see, e.g., Christiano, “Voter Ignorance Is Not Necessarily a Problem,” 257–60.

37 Somin, “The Ongoing Debate over Political Ignorance,” 386. For additional criticism of the appeal to heuristics, see Freiman, *Why It’s OK to Ignore Politics*, 18–21.

ture that suggests that heuristics not only often fail to overcome information and rationality deficits, but may even exacerbate them.<sup>38</sup>

Second, and more important for our argument, an appeal to heuristics does not break the symmetry between consuming and voting. If heuristics enable voters to overcome bias, they presumably could enable consumers to overcome bias too. Indeed, consider that Thomas Christiano's defense of heuristics in voting makes explicit use of analogies to consumption heuristics:

People use shortcuts in all walks of life and in every aspect of their lives. Going to the doctor is a shortcut compared to studying for the rest of my life how my body works. Going to a mechanic is a shortcut compared to learning a lot about how cars work. In a society with such a complex division of labor such as our own, economic life and political life would grind to a halt if it were required that people know a lot about the things they depend on. It is well known that people are strikingly ignorant of what is in their toothpaste, their cars, their financial arrangements, and their bodies, just to start an endless list. Does this mean that they act on the basis of no information? No. It implies that they act on the basis of other people's beliefs and statements about these matters while not knowing or even understanding the bases of those beliefs.<sup>39</sup>

Perhaps heuristics do correct for bias and ignorance in voting. But then we should even more strongly expect them to correct for bias and ignorance in consumption as well, thereby undermining the case for consumption paternalism in the first place.

Just as consumer incompetence opens the door for paternalistic interference, so too does voter incompetence. Conly claims coercive paternalism is justified when four conditions are met: (1) the activity proscribed must genuinely be against individuals' long-term interests according to the individuals' own values; (2) the coercive interference must tend to succeed; (3) it must survive cost-benefit analysis; and (4) noncoercive interventions are not as effective.<sup>40</sup> Hanna offers similar criteria. He claims that meeting condition 1 provides a presumptive justification for a "pro-paternalism attitude"; after that, it is an empiric-

38 Dancy and Sheagley, "Heuristics Behaving Badly"; Bartels, "Uninformed Votes"; Kuklinski and Quirk, "Reconsidering the Rational Public"; Fiske and Taylor, *Social Cognition, from Brains to Culture*; Kahneman, and Frederick, "Representativeness Revisited"; Nyhan and Reifler, "When Corrections Fail"; Kuklinski, Quirk, Jerit, and Schwieder, "Misinformation and the Currency of Democratic Citizenship"; Nadeau and Niemi, "Educated Guesses."

39 Christiano, review of *Against Democracy*.

40 Conly, *Against Autonomy*.

ical, social-scientific question whether paternalism will succeed and be worth implementing, all things considered.

So, the paternalist claims that if someone who decides to smoke *would not* have started smoking if they had made an accurate appraisal of the costs and benefits, and if the harms of smoking are sufficiently high, this provides presumptive grounds for interfering with their decision. Whether the government should then implement a smoking ban, impose cigarette taxes, or do nothing depends on matters of political economy, such as to what degree paternalist policies will be captured by special interests or how effective the bans will be. Similarly, if someone who decides to vote for a candidate or policy *would not* cast that vote if they had an accurate appraisal of the costs and benefits, and if the harms of doing so are sufficiently high, then the paternalist should say the same about interfering with those voters.

What might this look like in practice? A number of social scientists have provided accounts of the public's "enlightened" policy preferences.<sup>41</sup> We could, for instance, empower a regulatory body with the authority to veto legislation that conflicts with the public's enlightened economics and political preferences.<sup>42</sup> As Bryan Caplan puts it, "In the enlightened preference approach, one estimates what a person would think if you increased his level of political knowledge to the maximum level, keeping his other characteristics fixed."<sup>43</sup> To calculate such enlightened preferences, voters are asked to (1) express their opinions on a wide range of issues, (2) provide their demographic information (since this influences policy preferences), and (3) take a quiz of basic political information. With such data (all of which can be made public), it is possible to statistically estimate what a demographically identical voting public *would* have wanted if they were fully informed. It is easy to check for the robustness of the results if the questions had been changed in various ways. Indeed, this method is the way political scientists today estimate the independent effects of demographics on policy preferences while controlling for knowledge, or estimate the independent effects of knowledge while controlling for demographics.<sup>44</sup>

41 Bartels, "Uninformed Votes"; Althaus, *Collective Preferences in Democratic Politics*; Caplan, *The Myth of the Rational Voter*; and Gilens, *Affluence and Influence*.

42 See, e.g., Brennan, *Against Democracy*; and Caplan, *The Myth of the Rational Voter*. You could go for more moderate forms of epistocratic paternalism. Maybe you could nudge the voters by designing the ballot order to increase the chances of voting no. We take no stand on the particulars here.

43 Caplan, *The Myth of the Rational Voter*, 55. See also Althaus, *Collective Preferences in Democratic Politics*.

44 Brennan advocates for this kind of epistocracy, which he calls "government by simulated oracle" or, more recently, "enlightened preference voting" (*Against Democracy*). He offers



Crucially, the enlightened-preference approach to policy is similar to Conly's and other paternalists' approach to consumption. In neither case are we contemplating the perfectionist notion of imposing alien values on the public. Rather, the government is providing people with what they *would* prefer if they possessed accurate beliefs about the facts.

A fringe benefit of epistocratic paternalism is that it would spare us from the tedium of politics, just as Conly proposes that consumption paternalism would spare us from the tedium of researching car safety.<sup>45</sup> Indeed, it is probably a political philosopher's conceit to believe that Americans do or should care more about electoral politics than their cars. A stable finding in political science is that most citizens find politics uninteresting. The minority who do find it interesting tend to be more active and better informed, but also extremely biased.<sup>46</sup>

#### IV

At first glance, the philosophical justification for paternalistic intervention in consumer choices looks like an even stronger presumptive argument for paternalistic intervention in voter choices. Voters appear to be even more strongly beset by biases than consumers, in part because the feedback mechanism in democracy is far weaker than almost all market decisions. However, some will claim that these two cases are disanalogous, or that paternalism against voters faces special problems. We will turn now to considering those objections.

One objection holds that there is an asymmetry between voting badly and making bad consumer choices. Your individual consumer choices are individually efficacious. If you decide to smoke a pack of cigarettes daily, you in fact do so. Your individual vote is not efficacious. How you vote has a tiny chance of making any difference. How *we* vote matters, but how any one of us votes does not. (This, remember, explains why the problems of ignorance and bias are worse in voter choices than with consumer choices.) Thus, one might worry, the case for paternalistic interference with individual voters is weaker. How can we coerce Bob to vote better if Bob's vote does not matter?

In response, consider a variation of Mill's famous bridge case. Suppose one hundred marathoners are trying to cross a bridge all at once. The bridge can safely hold ten people, but will collapse under eleven or more people. Again,

---

an account of how to choose the questions on the "knowledge quiz" portion to avoid or reduce special-interest manipulation.

45 The idea that Conly's argument about self-regulation could carry over to the vote was first suggested by Aaron Ross Powell.

46 Mutz, *Hearing the Other Side*.



suppose you cannot warn the marathoners or convince them to cross in small groups. Here, it seems plausible, on paternalistic grounds, that one may interfere to stop them from passing, even though in this case no individual person crossing makes any significant difference. Even though an individual's choice to join the marathon will not make or break the collapse of the bridge—and it would be better for a particular individual to be permitted to run while the others are blocked—it looks like the reason to endorse paternalism in Mill's original bridge case justifies paternalistic interference here: the enlightened preferences of the marathoners would be satisfied by fencing off the bridge.

A closely related variation on this objection holds that “intelligence” can be an emergent feature of the collective decision maker. Certain mathematical theorems imply that, in some conditions, a group can make smart decisions as a whole even if the individuals within that group are not so smart. The miracle of aggregation theorem holds that ignorant voters might make random errors that cancel each other out. Condorcet's jury theorem claims that, in certain conditions, if the mean reliability of individuals within a group is greater than chance ( $>0.5$ ), then as the group becomes larger the probability it will make a correct decision approaches 1. (The theorem also says that if mean reliability  $<0.5$ , then as the size of the group increases the probability they will make the wrong decision approaches 1.) The Hong-Page theorem says that increasing the “cognitive diversity” of a group improves collective decision making more than increasing the average reliability of individuals within the group.<sup>47</sup>

Now, there is a long debate about just when and whether those conditions are met in actual democratic decisions.<sup>48</sup> We will not try to settle that debate here. Instead, we note that everyone agrees that if voters make systematic errors and mistakes, then the Hong-Page and miracle theorems do not apply, while Condorcet's jury theorem instead implies that democracies will always make bad choices. We note that both sides agree that the theorems apply to real-life democracies only under special conditions, and many real-life democratic decisions do not meet those conditions.<sup>49</sup> We also note that both sides of the debate seem to agree that sometimes the theorems apply and sometimes democracies make systematic errors. The debate concerns how *prevalent* these problems are.

47 Brennan, *Against Democracy*, 172–203; and Landemore, *Democratic Reason*.

48 Caplan, *The Myth of the Rational Voter*; Brennan, *Against Democracy*; Landemore, *Democratic Reason*; Somin, *Democracy and Political Ignorance*; and Achen and Bartels, *Democracy for Realists*.

49 E.g., Brennan claims that the Hong-Page theorem does not apply to most actual votes, because the conditions of the theorem are not met (*Against Democracy*). Landemore seems to concede this but then argues that we should change democratic decision procedures in order to better fit the theorem and take advantage of cognitive diversity (*Democratic Reason*).

Here, the paternalist can just say, “Sure, when there is emergent collective intelligence, the case for paternalism disappears, but when there is collective folly, it remains. Now it is an empirical question how often we have collective wisdom or folly. But that is no different from the problem of consumer choice. Sometimes consumers tend to make wise choices; sometimes they make systematic errors.”

Another objection to our argument alleges that the vetoed legislation would inevitably have benefited *some* citizens, thereby harming a few to help the many. For instance, tariffs might benefit some domestic producers, but an epistocratic veto would probably disallow them.

The problem with this objection is that it counts against consumption paternalism as well. As Conly herself says,

do these [paternalistic] laws mean that some people will be kept from doing what they really want to do? Probably—and yes, in many ways it hurts to be part of a society governed by laws, given that laws aren’t designed for each one of us individually. Some of us can drive safely at 90 miles per hour, but we’re bound by the same laws as the people who can’t, because individual speeding laws aren’t practical. Giving up a little liberty is something we agree to when we agree to live in a democratic society that is governed by laws.<sup>50</sup>

Similarly, a smoking ban would harm the small group of smokers who do have accurate beliefs about the risks of smoking and proceed to smoke anyway. But this implication leaves coercive paternalists undeterred.

A Rawls-inspired objection might allege that the right to vote is a political liberty that is protected as basic, meaning that it may not be infringed upon except for the sake of other basic liberties. This implies that Steve’s right to vote may not be restricted to promote Steve’s welfare, for example. This objection would rule out paternalistic interference with the vote.

Before we address this objection in depth, it is important to reiterate that our aim is not to defend paternalism about consumer choices or voting; rather our aim is to argue that those who endorse paternalism in consumer choice have grounds at least as strong to endorse paternalism in democratic choice. Sarah Conly, Jason Hanna, and Cass Sunstein and Richard Thaler, among other paternalists, do not seem to accept the Rawlsian theory of basic liberty. However, perhaps some readers might think that, even if Conly and Hanna are not Rawlsians, nevertheless the Rawlsian theory explains why paternalism in consumer choices is less objectionable than in democratic choices.

We contend that the basic liberty objection fails to break the symmetry be-

<sup>50</sup> Conly, “Three Cheers for the Nanny State.”

tween voting choices and consumption choices. Paternalistic interference with consumption choices will frequently involve infringing upon liberties that Rawlsians consider basic. For instance, the right of bodily autonomy is a basic liberty and it would appear, on the most obvious reading, to protect a person's right to smoke a cigarette or eat sugary foods. They are your lungs and it is your waistline, after all. Similarly, liberals believe in freedom of speech, yet paternalistic interventions such as mandatory calorie counts and health warnings on cigarettes are plausibly construed as forms of compelled speech. If the state may nevertheless interfere with your choices regarding your body or your speech, we see little reason why it may not interfere with your choices regarding your vote.

You could reply that the right of bodily autonomy or free speech does not protect *all* uses of one's body or one's speech. Perhaps, in a Rawlsian spirit, only those uses that are required for the *adequate* development and exercise of your sense of justice and conception of the good are protected. In a debate with John Tomasi, Samuel Freeman elaborates that a liberty is basic only if it is necessary for *all* citizens to possess that liberty in order to adequately develop and fully exercise their two moral powers.<sup>51</sup> However, elsewhere he recognizes that there will always be exceptions. Perhaps some "peripatetic ascetic" is able to adequately develop their moral powers despite living in extreme deprivation.<sup>52</sup> So, Freeman probably means that a liberty qualifies as a basic liberty just in case it is an essential social condition for *most* people to adequately develop and fully exercise the two moral powers.

However, here the Rawlsian runs into serious trouble. As an empirical matter, it seems very little liberty is strictly speaking essential for most people to develop the two moral powers. People in moderately illiberal, deeply authoritarian, or even totalitarian regimes may have a harder time than we do in accessing the proper evaluative horizons for them to develop the moral powers, but even in such countries, it is not impossible, or even all that difficult to develop these powers. Most citizens do. To develop the two moral powers, you do not need much (let alone extensive) freedom of speech, freedom of marriage rights, much freedom of association, or much political liberty.<sup>53</sup> You do not need to have the right to vote or run for office. You do not need to live in a society that completely realizes the rule of law. You do not need to have the unlimited right to choose your own occupation. Surely you can adequately develop and exercise your moral powers without being permitted to smoke. But here again, people can ade-

51 Freeman, "Can Economic Liberties Be Basic Liberties?"

52 Freeman, *Rawls*, 56.

53 See Flanigan, "All Liberty Is Basic"; and Freeman and Thrasher, "The Right to Own the Means of Production," 194.

quately develop and exercise their moral powers while having the results of their vote subject to an epistocratic veto or enlightened-preference calculations.

This last point is worth stressing. Depriving citizens of the right to smoke is compatible with their freedom to make plenty of other choices about what to do with their bodies. Thus, one's right of bodily autonomy may well be adequately respected despite being truncated by paternalistic intervention. But presumably the same point holds with respect to democratic rights. Subjecting citizens' votes to enlightened-preference calculations is compatible with respecting plenty of other participatory rights. Citizens may still participate in politics via phone banking, door-to-door campaigning, writing op-eds, and more.

The Rawlsian might agree, but then say that while people can *develop* their two moral powers despite having significant paternalistic interference with consumption choices or one's vote, nevertheless, one cannot fully exercise one's moral powers without such a right. Perhaps not—we take no stand on this point. Here we would only once again reiterate that our thesis concerns the symmetry between consuming and voting. If citizens are unable to fully exercise their moral powers when their democratic rights are abridged, then presumably they are unable to fully exercise their moral powers when their right of bodily autonomy is similarly abridged. At a minimum, we think the burden of justification rests with those who would assert an asymmetry.

Another reply to our argument may appeal to the *intrinsic* value of democracy and self-governance. Perhaps the right to make unabridged democratic choices is simply good in itself. This could be true, but it is unclear whether this reply is available to consumption paternalists given their views about the value of consumer freedom. Conly, for instance, suggests that we ought to downgrade the value of autonomy in light of the finding that we use it to make systematically bad choices. She writes, "Autonomy is not all that valuable; not valuable enough to offset what we lose by leaving people to their own autonomous choices. The truth is we do not reason very well and in many cases there is no justification for leaving us to struggle with our own inabilities and to suffer the consequences."<sup>54</sup> Any given paternalist might assign unencumbered choice *no* intrinsic value, or perhaps simply sufficiently low intrinsic value such that it is outweighed by the value of the welfare benefits of paternalist interference. Of course, paternalists might be wrong—maybe self-governance *is* quite intrinsically valuable. But in this case, the challenge of explaining why uninhibited choice is more intrinsically valuable in the political realm than the economic realm remains.

Perhaps less drastic, non-epistocratic institutional reforms could ameliorate the problems of voter ignorance and bias. In his discussion of affluence and ac-

54 Conly, *Against Autonomy*, 1.

cess to political information, Christiano writes, “Education is a good place to start with, but it will not solve the problem of political information. What is needed are institutions that disseminate what Downs calls ‘free information’ to ordinary people.”<sup>55</sup> Yet if the dissemination of free information obviates the need for paternalism in voting, it should also obviate the need for paternalism in consumption. Moreover, paternalists themselves are skeptical that information alone will make consumers more rational. For instance, people are fairly well informed of the dangers of cigarette smoking and yet they continue to smoke too much for paternalists’ liking.<sup>56</sup>

We acknowledge that objections to epistocracy may remain that we lack the space to consider. However, we remind readers that we are not here defending paternalism or epistocracy. We instead say that paternalists face a dilemma: they should either *also* accept epistocracy or give up paternalism about consumer choice. Their reasons for endorsing paternalism in the market are at least as strong as reasons to be paternalists about politics as well. (We thus acknowledge that if you have no interest in being a consumption paternalist, our arguments do not apply to you.) Thus, a good objection to our argument must find a *disanalogy* between the two cases. It must not simply be an argument that defeats the case for paternalism *simpliciter*. Rather, it must be an objection that defeats paternalism about political choices but that does not also refute paternalism about consumer or personal choices.

## v

Last, there is a practical objection to paternalistic regulation of the vote: state agents might abuse their new powers. They are influenced by self-interest and bias like anyone else. Special interest groups and political groups might engage in rent seeking with the goal of capturing administrative agencies or paternalistic laws for their own benefit. Perhaps government failure would be so severe that the paternalistic interventions into voter behavior would not on net promote good outcomes.

This sounds right to us, but at first glance the point applies equally to paternalistic regulation of consumption choices.<sup>57</sup> Indeed, paternalistic regulation of consumption has proven to be corruptible time and again. Consider, for instance, the role of “bootleggers and Baptists” in institutionalizing alcohol prohibition, sugar and corn syrup manufacturers’ influence on the US government’s campaign

55 Christiano, review of *Against Democracy*.

56 Conly, *Against Autonomy*, 3.

57 For a more detailed exploration of this worry, see Rizzo and Whitman, *Escaping Paternalism*.

against fat and in favor of carbohydrates, the influence of opiate manufacturers, alcohol producers, and police unions in preserving drug criminalization, Big Tobacco's attempts to stifle vaping through regulations, and casino owners' fight against online gambling. In these cases and more, protecting people from themselves has simply served as a convenient pretext for self-interested rent seeking.

There are also some reasons to think government failure in paternalism about voter choice may be less severe than government failure in consumer choice. The media keeps constant vigil on issues of gerrymandering, purported voter fraud, campaign finance, Russian hacking, and the like. In contrast, the media largely ignores cases of blatant, socially destructive rent seeking in consumer markets, such as Archer-Daniels-Midland's corn subsidies, even though such cases are routinely used in economics textbooks as examples. Epistocratic paternalism will be more closely monitored than consumer paternalism.

And remember that people have far stronger incentives to acquire accurate beliefs to inform their consumption choices than their voting choices, because their consumption choices are "decisive" but their voting choices are not. If you choose to buy a particular house, you will get it. Thus, you had better make sure you have done your homework to ensure that it is in a safe neighborhood, zoned for good schools, and in reasonably decent shape. If you are wrong, you will suffer the costs. On the other hand, your voting choice will never be decisive—your choice to vote for Candidate *X* will never cause Candidate *X* to win the election. Thus, the cost of casting a careless vote is dramatically smaller than the cost of making a careless purchase. One implication of the comparative thoughtfulness of consumption choices is that states may have less occasion to paternalistically interfere with them. Further, the *moral* cost of interference seems greater when the choice in question is more thoughtful and informed. Perhaps you should be less willing to interfere with your friend's reflective decision to handle snakes as part of a religious ritual than his kneejerk decision to grab a cobra at the zoo. The former decision is more expressive of his values and commitments and may therefore be deserving of more respect.

At any rate, note that Conly and Hanna, among others, already hold that whether a particular paternalistic intervention is warranted should be decided on a case-by-case basis, taking such factors as government failure and rent seeking into account. Conly, for one, thinks all-things-considered this means that alcohol prohibition, while *prima facie* justifiable according to paternalistic reasoning, is not worth pursuing, while a ban on cigarettes is. For such paternalists, the issue is only whether a suitable institutional framework for any particular paternalistic intervention can be developed. Conly, Hanna, and other paternalists should thus apply the same reasoning to paternalistic interventions in voting.

*Prima facie*, they should regard such interventions as justified. They might accept some interventions and reject others on the final analysis if, in light of problems of political economy, they fail cost-benefit analysis.

It may very well turn out that unfettered democracy simply performs better than the available institutional alternatives in terms of promoting social trust, securing stability, and so on. In this case, democracy should be preferred to epistocracy on purely instrumental grounds. This conclusion is consistent with our argument, which is that paternalists should have no in-principle opposition to epistocracy and therefore no objection to exploring whether paternalistic restrictions on democratic performance can be made to work. But note that this conclusion—perhaps to the dissatisfaction of some democratic theorists—holds the viability of democracy hostage to our cost-benefit calculations.<sup>58</sup>

In closing, let us issue a reminder that our claims are appropriately modest: paternalists should endorse the permissibility of epistocratic paternalism in principle. The extent to which their principles commit them to paternalist interference with the vote in practice is an empirical question. As of now, no one in political science or economics has done sufficient empirical work to merit the conclusion that epistocratic paternalism cannot work while consumer paternalism can and does. Our argument may make coercive paternalists uneasy. As we have seen, even Conly does not waver in her commitment to the competence of voters. And many political philosophers want to reject epistocratic interference with the vote *in principle*. As a sociological matter, paternalists about consumption vastly outnumber epistocrats. But if we are right, philosophers should be either less paternalistic or more epistocratic.

Georgetown University  
jb896@georgetown.edu

William & Mary  
cafreiman@wm.edu

#### REFERENCES

Achen, Christopher, and Larry Bartels. *Democracy for Realists*. Princeton: Princeton University Press, 2016.

58 For helpful feedback on earlier versions of this paper, we owe thanks to an audience at the 2018 meeting of the Philosophy, Politics, and Economics Society as well as two anonymous referees for this journal.



- Allcott, Hunt, Benjamin B. Lockwood, and Dmitry Taubinsky. "Regressive Sin Taxes, with an Application to the Optimal Soda Tax." *Quarterly Journal of Economics* 134, no. 3 (August 2019): 1557–1626.
- Althaus, Scott. *Collective Preferences in Democratic Politics*. New York: Cambridge University Press, 2003.
- Ariely, Dan. *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. New York: Harper Perennial, 2010.
- Bader, Pearl, David Boisclair, and Roberta Ferrence. "Effects of Tobacco Taxation and Pricing on Smoking Behavior in High Risk Populations: A Knowledge Synthesis." *International Journal of Environmental Research and Public Health* 8, no. 11 (November 2011): 4118–39.
- Bartels, Larry M. "Uninformed Votes: Information Effects in Presidential Elections." *American Journal of Political Science* 40, no. 1 (February 1996): 194–230.
- Brennan, Jason. *Against Democracy*. Princeton: Princeton University Press, 2016.
- . "The Right to a Competent Electorate." *Philosophical Quarterly* 61, no. 245 (October 2011): 700–24.
- Callison, Kevin, and Robert Kaestner. "Do Higher Tobacco Taxes Reduce Adult Smoking? New Evidence of the Effect of Recent Cigarette Taxes on Adult Smoking." *Economic Inquiry* 52, no. 1 (January 2014): 155–72.
- Caplan, Bryan. *The Myth of the Rational Voter*. Princeton: Princeton University Press, 2007.
- Caplan, Bryan, Eric Crampton, Wayne Grove, and Ilya Somin. "Systematically Biased Beliefs about Political Influence: Evidence from the Perceptions of Political Influence on Policy Outcomes Survey." *PS: Political Science and Politics* 46, no. 4 (October 2013): 760–67.
- Christiano, Thomas. Review of *Against Democracy*. *Notre Dame Philosophical Reviews*, May 19, 2017. <https://ndpr.nd.edu/reviews/against-democracy/>.
- . "Voter Ignorance Is Not Necessarily a Problem." *Critical Review* 27, nos. 3–4 (2015): 253–69.
- Conly, Sarah. *Against Autonomy: Justifying Coercive Paternalism*. New York: Cambridge University Press, 2014.
- . "Three Cheers for the Nanny State." *New York Times*, March 24, 2013. <https://www.nytimes.com/2013/03/25/opinion/three-cheers-for-the-nanny-state.html>.
- Converse, Philip. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, edited by D. E. Apter. London: Free Press of Glencoe, 1964.
- Dancy, Logan, and Geoffrey Sheagley. "Heuristics Behaving Badly: Party Cues and Voter Knowledge." *American Journal of Political Science* 57, no. 2 (April 2013): 312–25.



- Delli-Carpini, Michael X., and Scott Keeter. *What Americans Know about Politics and Why It Matters*. New Haven: Yale University Press, 1996.
- Erisen, Cengiz, Milton Lodge, and Charles S. Taber. "Affective Contagion in Effortful Political Thinking." *Political Psychology* 35, no. 2 (April 2014): 187–206.
- Fiske, Susan, and Shelley Taylor. *Social Cognition: From Brains to Culture*. New York: McGraw-Hill, 2008.
- Flanigan, Jessica. "All Liberty Is Basic." *Res Publica* 24, no. 4 (November 2018): 455–74.
- . "Seat Belt Mandates and Paternalism." *Journal of Moral Philosophy* 14, no. 3 (June 2017): 291–314.
- Fotuhi, Omid, Geoffrey T. Fong, Mark P. Zanna, Ron Borland, Hua-Hie Yong, and K. Michael Cummings. "Patterns of Cognitive Dissonance-Reducing Beliefs among Smokers: A Longitudinal Analysis from the International Tobacco Control (ITC) Four Country Survey." *Tobacco Control* 22, no. 1 (January 2013): 52–58.
- Freeman, Samuel. "Can Economic Liberties Be Basic Liberties?" *Bleeding Heart Libertarians*, June 13, 2012. <https://bleedingheartlibertarians.com/2012/06/can-economic-liberties-be-basic-liberties>.
- . "Illiberal Libertarians: Why Libertarianism Is Not a Liberal View." *Philosophy and Public Affairs* 30, no. 2 (Spring 2001): 105–51.
- . *Rawls*. New York: Routledge Press, 2006.
- Freiman, Christopher. *Why It's OK to Ignore Politics*. New York: Routledge, 2020.
- Freiman, Christopher, and John Thrasher. "The Right to Own the Means of Production." In *Economic Liberties and Human Rights*, edited by Jahel Queralt and Bas van der Vossen, 182–99. New York: Routledge, 2019.
- Gilens, Martin. *Affluence and Influence*. Princeton: Princeton University Press, 2012.
- Hanna, Jason. *In Our Interest*. New York: Oxford University Press, 2018.
- Hedden, Brian. "Hindsight Bias Is Not a Bias." *Analysis* 79, no. 1 (January 2019) 43–52.
- Kahneman, Daniel, and Shane Frederick. "Representativeness Revisited: Attribute Substitution in Intuitive Judgment." In *Heuristics and Biases: The Psychology of Intuitive Judgment*, edited by Thomas Gilovich, Dale Griffin, and Daniel Kahneman, 49–81. New York: Cambridge University Press, 2002.
- Kelly, James. *Framing Democracy*. Princeton: Princeton University Press, 2012.
- Kelly, Thomas. "Sunk Costs, Rationality, and Acting for the Sake of the Past." *Notus* 38, no. 1 (March 2004): 60–85.
- Kinder, Donald, and Nathan Kalmoe. *Neither Liberal nor Conservative: Ideolog-*

- ical Innocence in the American Public*. Chicago: University of Chicago Press, 2017.
- Kuklinski, James H., and Paul J. Quirk. "Reconsidering the Rational Public: Cognition, Heuristics, and Mass Opinion." In *Elements of Reason: Cognition, Choice, and the Bounds of Rationality*, edited by Arthur Lupia, Matthew D. McCubbins, and Samuel L. Popkin, 153–82. New York: Cambridge University Press, 2000.
- Kuklinski, James H., Paul J. Quirk, Jennifer Jerit, David Schwieder, and Robert F. Rich. "Misinformation and the Currency of Democratic Citizenship." *Journal of Politics* 62, no. 3 (2000): 790–816.
- Landemore, Hélène. *Democratic Reason*. Princeton: Princeton University Press, 2012.
- Lavner, Justin, Benjamin R. Karney, and Thomas N. Bradbury. "Newlyweds' Optimistic Forecasts of Their Marriage: For Better or for Worse?" *Journal of Family Psychology* 27, no. 4 (August 2013): 531–40.
- Lodge, Milton, and Charles Taber. *The Rationalizing Voter*. New York: Cambridge University Press, 2013.
- López-Guerra, Claudio. *Democracy and Disenfranchisement*. New York: Oxford University Press, 2014.
- Mueller, John, and Mark Steward. *Terror, Security, and Money: Balancing the Risks, Benefits, and Costs of Homeland Security*. New York: Oxford University Press, 2011.
- Mulligan, Thomas. "Plural Voting for the Twenty-First Century." *Philosophical Quarterly* 68, no. 271 (April 2018): 286–306.
- Mutz, Diana. *Hearing the Other Side*. New York: Cambridge University Press, 2006.
- Nadeau, Richard, and Richard G. Niemi. "Educated Guesses: The Process of Answering Factual Knowledge Questions in Surveys." *Public Opinion Quarterly* 59, no. 3 (Fall 1995): 323–46.
- Nebel, Jacob M. "Status Quo Bias, Rationality, and Conservatism about Value." *Ethics* 125 (January 2015): 449–76.
- Nyhan, Brendan, and Jason Reifler. "When Corrections Fail: The Persistence of Political Misperceptions." *Political Behavior* 32, no. 2 (June 2010): 303–30.
- O'Donoghue, Ted, and Matthew Rabin, "Present Bias: Lessons Learned and to Be Learned." *American Economic Review* 105, no. 5 (May 2015): 273–79.
- Rizzo, Mario, and Glen Whitman. *Escaping Paternalism*. New York: Cambridge University Press, 2019.
- Somin, Ilya. *Democracy and Political Ignorance*. Stanford: Stanford University Press, 2013.

- . “The Ongoing Debate over Political Ignorance: Reply to My Critics.” *Critical Review* 27, nos. 3–4 (2015): 380–414.
- Stango, Victor, and Jonathan Zinman. “Exponential Growth Bias and Household Finance.” *Journal of Finance* 64, no. 6 (December 2009): 2807–49.
- Thaler, Richard, and Cass Sunstein. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven: Yale University Press, 2008.
- Tversky, Andrew, and Daniel Kahneman. “Availability: A Heuristic for Judging Frequency and Probability.” *Cognitive Psychology* 5, no. 2 (September 1973): 207–33.
- Watson Institute for International and Public Affairs. “\$5.9 Trillion Spent and Obligated on Post-9/11 Wars.” Brown University, November 16, 2018. <https://watson.brown.edu/research/2018/59-trillion-spent-and-obligated-post-911-wars>.
- Westen, Drew. *The Political Brain: How We Make Up Our Minds without Using Our Heads*. New York: Perseus Books, 2008.
- Westen, Drew, Pavel S. Blagov, Keith Harenski, Clint Kilts, and Stephan Hamann. “The Neural Basis of Motivated Reasoning: An fMRI Study of Emotional Constraints on Political Judgment during the U.S. Presidential Election of 2004.” *Journal of Cognitive Neuroscience* 18, no. 11 (November 2007): 1947–58.

## WELL-BEING AS NEED SATISFACTION

Marlowe Fardell

NEED-SATISFACTION theories of well-being are rare in philosophy.<sup>1</sup> When considered at all they are commonly dismissed as unviable.<sup>2</sup> However, I argue that such neglect and dismissal owe to mistaken preconceptions about needs' essential nature. Here I make a start on refuting these, defend a new theory of well-being as satisfying certain needs, and discuss its significant practical ramifications. Crucially, a need-satisfaction theory primarily aims to do something different from the theories that philosophers most commonly discuss. Those theories usually either detail what well-being consists of (e.g., pleasure, achievement, friendship), explain why the things that are good for people are good for people (e.g., because people desire them), or both. By contrast, a need-satisfaction theory's distinctive purpose is to specify how the constituents of well-being—whatever they are—are *structured*. As a result, a need-satisfaction theory does not necessarily conflict with theories of other types, and indeed may complement them.

The chief motivation of the need-satisfaction theory I defend is to vindicate a belief about well-being's structure integral to many people's self-understanding. This is that certain of their engagements are irreplaceable to the good of their lives. Some of the central constituents of their well-being appear to them to be non-substitutable: no other goods can make up for them if they are lost or forsaken. Examples of such engagements, for some people, include commitments to certain projects, tasks, or vocations; to maintaining communities, cultural practices, environments, or relationships; and to maintaining integrity of character or devotion to belief systems or causes. In academic and policy con-

- 1 In the evaluative context at hand (cf. section 4.1), a person's "well-being" is the state (realizable to different extents) in which they have, are doing, and are being things that are *finally good for them* to have, do, and be. Something is "finally good" if it is worth having, doing, or being *for its own sake*, not for any further purpose (cf. section 1 and Crisp, *Reasons and the Good*, 100). More exactly, a person's well-being also comprises their "ill-being": ways their life is bad that are not simply absences of good things. However, I will not discuss ill-being in this paper.
- 2 E.g., Griffin, *Well-Being*, 41–47.

texts, however, a very common assumption is that well-being is structured in a way that directly contradicts this appearance. It is often assumed that well-being comes (or can be represented as if it came) in generic *amounts*, contributed in varying degrees by the good things in a person's life. In other words, it is assumed that well-being can at least be represented in terms of a unidimensional, homogeneous currency.<sup>3</sup> Anything bearing or yielding the same amount of this well-being currency as another could substitute for it with no loss. I call this view "structural monism" about well-being.<sup>4</sup> Where structural monism fails to describe many people's well-being as they understand it, the need-satisfaction theory I defend accounts for the phenomena by building non-substitutability into the structure of well-being.

The first three sections of this paper defend an account of the needs relevant here. Section 1 presents an original analysis of the concept of "categorical" needs and defines the subtype of these I am especially concerned to defend, "personal needs." In section 2, against prevailing views assuming that categorical needs are exclusively minimal, universal, and moralized, I argue that personal needs are also categorically necessary despite being non-minimal, particular to individual persons, and nonmoralized. In section 3, I explain how personal needs possess this necessity, namely by being the inescapable practical demands of a person's commitments, specially defined. The next three sections develop a theory of well-being that adopts this account of needs, well-being as personal need satisfaction (WAPNS). In section 4, I seek to avert possible misunderstandings by further clarifying this theory's relations to other types of well-being theory. Sections 5 and 6 each defend one of its two central claims, both of which arise from personal needs' and commitments' twofold non-substitutability: by non-needs and by each other. Both aspects owe to personal needs' inescapability and commitments' centrality. WAPNS would vindicate the possibility of non-substitutable constituents of well-being and refute structural monism. In section 7, I discuss its considerable implications for aggregating and measuring well-being. Section 8 concludes.

A caveat. This paper aims to convey the potential appeal and fruitfulness of a theory that has many interrelated components. This necessitates covering a lot of ground, and means I cannot discuss and defend each component in as much

- 3 It is irrelevant which scale or units are chosen. For this paper's purposes it also does not matter whether a person's (a) *absolute total* well-being, or else only (b) *changes* in their well-being, are assumed to be representable by sums of a well-being currency. (The latter is less demanding, since it does not require any level of zero well-being to be defined.)
- 4 "Structural" since it need not involve a robust metaphysical commitment to there "really existing" only one constituent of well-being. As I will later explain, a strength of the view is indeed that the currency may be a formal construction (section 3.3).

detail as it ideally merits. Accordingly, the paper is better read as a detailed outline than a definitive statement.

### 1. WHAT ARE CATEGORICAL NEEDS?

On my proposal the concept of a personal need is a subtype of a more general need concept, so I begin by analyzing the latter. The latter needs are variously called “non-contingent,” “fundamental,” “absolute,” and “categorical” needs.<sup>5</sup> I use “categorical.” They are distinct from mere instrumental preconditions for further ends, for example, such as to have a knife in order to cut something. Rather, they are conditions somehow necessary in their own right for a person to satisfy. On my analysis, this general concept is underdetermined, possessing what can be called a “modular structure.”<sup>6</sup> It has several essential conditions, three with fixed content (“A modules”), and three others that, while also essential, are open to different specifications (“B modules”). In other words: (i) all categorical needs share certain essential features (A modules); (ii) different types of categorical needs exist, varying in other respects; and (iii) this variation is nevertheless further conditioned or restricted (by B modules). What this means will become clearer as I proceed.

The A modules of categorical need are as follows:

*Final* (A<sub>1</sub>): Categorical needs are conditions that it is finally necessary for the person who has them to fulfill.

*Real* (A<sub>2</sub>): There is a fact of the matter about what a person categorically needs independent of the person’s actually, presently apprehending it.

*Inescapable* (A<sub>3</sub>): A person’s categorical needs are not subject to their will, in that a person cannot simply decide what they do and do not need (at least not directly).

Let me unpack these. The first A module, *Final*, says that categorical needs are conditions it is necessary to fulfill for their own sake, irrespective of whatever

5 Reader, *Needs and Moral Necessity*; Thomson, *Needs*; Wiggins, “Claims of Need”; Frankfurt, “Necessity and Desire.”

6 I adapt this approach and terminology from Ingrid Robeyns’s mapping of the relations between various specific capabilities *theories* to the capabilities approach to development as a whole, which is underspecified (Robeyns, *Wellbeing, Freedom and Social Justice*). My suggestion here is that this modular representation can also help with understanding *concepts* that are underspecified but that, analogously to Robeyns’s approach, nevertheless entail some restrictions on how they can be further specified while remaining species of the same general concept.

other purposes satisfying them might also serve. They are “finally necessary.”<sup>7</sup> As before, they differ from needs that are necessary only as prerequisites for fulfilling further conditions, such as means to further ends. As I will discuss at length, various accounts of categorical needs’ final necessity exist. The personal needs I defend are finally necessary due to their constitutive relation with what matters to a person. Other needs’ final necessity is held to relate to moral requirement.

Modules Real and Inescapable underlie the commonly drawn needs/wants distinction. That distinction is confused if it depicts needs and wants as exclusive opposites. Some goods can be both wanted and needed. Other things might be good for a person, but neither wanted nor needed (reading some potentially edifying but unexciting book, say). Still, this popular contrast points to a kind of “objectivity,” expressed by A2 and A3, that categorical needs possess but that desires for non-needed things lack: their existence is not contingent on whatever a person happens to *believe* or *want* to be the case.<sup>8</sup>

On my analysis, conditions A1–A3 are necessary and sufficient for a good or interest to be a categorical need. Nevertheless, however a specific subtype of categorical needs fulfills these, it cannot avoid also fulfilling the following modules B1–B3 under some specification or another:

*Normative alignment* (B1): The needs correspond to or derive from normative considerations of some kind.

*Scope* (B2): The needs are shared across persons to some greater or lesser extent.

*Extent* (B3): The needs correspond to some more or less expansive or minimal extent of attainment.

I do not propose that these B modules exhaust the respects in which different categorical need concepts might vary. These are just those that are relevant to distinguishing the two types of categorical needs that I discuss in this paper.

In most contemporary accounts, categorical needs correspond directly to standards of just minimal provision or assistance. A person’s needs are considered to be the necessary constituents of a relatively minimal, socially acceptable

7 This is like “finally good,” where, as before, something is finally good if and only if it is good to have or pursue or be for its own sake. It is common to identify final value with intrinsic value; however, since “intrinsic” tends to carry further connotations, it is best to separate the final/nonfinal and intrinsic/extrinsic distinctions (Korsgaard, “Two Distinctions in Goodness”; O’Neill, “The Varieties of Intrinsic Value”; Kagan, “Rethinking Intrinsic Value”; Rabinowicz and Rønnow-Rasmussen, “A Distinction in Value”). I eschew “intrinsically necessary for similar reasons.

8 Cf. Griffin, *Well-Being*, 41; Wiggins, “Claims of Need,” 6.



living standard, or indispensable requirements for elementary forms of human social functioning.<sup>9</sup> Common to these accounts is the idea that certain attainments—those possessing categorical necessity—are intimately connected with moral or political requirements on moral agents to ensure that other persons satisfy them. Theorists of such needs “want an analysis of needs such that they turn out to be morally compelling.”<sup>10</sup> Clearly, these needs must connect with some at least partial conception of a good life. However, only some of a person’s interests, specially interpreted, are picked out as relevant; typically, lacking what one needs is identified with suffering *harm* (a morally freighted concept). Thus these *essentially moralized* needs, as I call them, specify the B modules of the general concept of categorical need as follows:

*Moralized* (B<sub>1</sub>\*): If a person lacks an essentially moralized need, then others are *pro tanto* morally obligated to assist them in attaining it. Moreover, persons’ essentially moralized needs extend only so far as the conditions others are *pro tanto* morally obligated to ensure they can attain.

*Universal* (B<sub>2</sub>\*): Essentially moralized needs are necessary for human persons as such (at least in some society), and hence are shared universally by such persons (at least within that society).<sup>11</sup>

*Minimal* (B<sub>3</sub>\*): Essentially moralized needs correspond to a relatively minimal standard of attainment.

Yet accounts of essentially moralized needs typically do not propose these conditions as specifications of B modules. Explicitly or otherwise, they imply that Moralized (B<sub>1</sub>\*), Universal (B<sub>2</sub>\*), and Minimal (B<sub>3</sub>\*) are in fact A modules of categorical needs. That is, they appear to hold that something is a genuinely cat-

9 Miller, *Principles of Social Justice*; Wiggins, “Claims of Need”; cf. Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*; Braybrooke, *Meeting Needs*; Doyal and Gough, *A Theory of Human Need*; Brock, “Morally Important Needs.” Another approach to moralizing needs, which I will not discuss, identifies needs specific to “private morality” (Brock and Reader, “Needs-Centered Ethical Theory”; Reader and Brock, “Needs, Moral Demands and Moral Theory”; Reader, *Needs and Moral Necessity*). That approach effectively weakens conditions B<sub>2</sub>\* and B<sub>3</sub>\* slightly, bringing such needs somewhat closer to the personal needs I defend (cf. esp. Reader, *Needs and Moral Necessity*, 65–66), but nevertheless keeps tight hold of B<sub>1</sub>\*.

10 Hooker, “Fairness, Needs, and Desert,” 185–86; cf. Doyal and Gough, *A Theory of Human Need*, 51.

11 Proponents of essentially moralized needs do hold that different people need different resources, depending on their physiologies (e.g., differences in metabolism, sex, mental or physical (dis)ability). But this variation in specific goods is required only to satisfy the same needs under general descriptions (Doyal and Gough, *A Theory of Human Need*).



egorical need if and only if it is essentially moralized, that is, also satisfies  $B_1^* - B_3^*$ . I consider rationales for this view in section 2. However, in this paper I argue that essentially moralized needs are not the only kind of categorical needs. The “personal needs” I defend are likewise categorical, but by contrast relate directly to central parts of well-being. Personal needs are categorical by virtue of likewise fulfilling the foregoing A modules, but in doing so they specify the B modules  $B_1 - B_3$  differently from essentially moralized needs:

*Practical* ( $B_1^{**}$ ): Personal needs are practical requirements on a person entailed by commitments that matter to them personally. These commitments are not necessarily moral commitments. Neither are personal needs necessarily conditions that other people are morally required to help a person satisfy.<sup>12</sup>

*Particular* ( $B_2^{**}$ ): Personal needs are particular to individual persons, since other people will share these needs (be practically compelled in the same ways) if and only if they have the same commitments. Moreover, some persons’ commitments and personal needs may be ones that only some persons or even no other person shares.

*Expansive* ( $B_3^{**}$ ): Satisfying personal needs constitutes a major part of the well-being of persons who have them.<sup>13</sup>

To summarize, personal needs are defined as the objective, inescapable practical requirements entailed by a person’s particular commitments. A commitment here is specially defined as a personal engagement consisting of the personal needs it entails; it is a constellation of related personal needs.

I discuss commitment and inescapable practical requirement in detail in section 4 and yet further in section 7, but we can see already that personal needs and essentially moralized needs are very different. Essentially moralized needs define a standard of just provision that is supposed to guide moral and political distribution. By contrast, fulfilled personal needs constitute part of well-being. Specifically, they constitute part of well-being in a personal, agential context of

12 I allow for the possibility that moral requirements form a subset of a person’s personal needs—presumably among *every* moral agent’s personal needs. But this is by no means essential to the account.

13 Other expansive accounts of needs define them as the necessary constituents of a state of full human flourishing (Kropotkin, *The Conquest of Bread*; Stewart, *Basic Needs in Developing Countries* and “Basic Needs Approach”; Grix and McKibbin, “Needs and Well-Being”). While interesting, these cannot be evaluated independently of the accounts of human flourishing in question (cf. Anscombe, “Modern Moral Philosophy,” 7, 18), which I cannot do in this paper.

evaluation—how it is best for a person to *live*, at some specific time and place, in respect of what matters personally to that person. By “what matters personally to that person” I mean the ends particular to the person that determine what their living well consists of. Well-being in this context is partially analogous to the concept of utility or welfare in utility theory, which is a function of an agent’s preferences over outcomes (or “consequences”), and where, “in the description of a consequence is included all that the agent values.”<sup>14</sup> So, needs in this agential context concern not moral or political entitlement, but rather how an individual *needs to act*. Beyond needs to *have* or *receive* things, a person also and most importantly needs to *do* and *be* certain things and ways.<sup>15</sup> Yet despite stark differences between personal needs and essentially moralized needs, I argue that both kinds are genuinely categorical needs.

While section 3 makes the positive case for personal needs being categorically necessary, I first address the deep skepticism my proposal might already have provoked: it may seem obviously impossible for genuinely categorical needs to be personal. I consider two arguments expressing this reaction in different ways. These do not yet threaten personal needs having a role in a theory of well-being, but rather my proposal’s adequacy as an account of needs. This discussion also serves to illustrate the difference between the personal needs I defend and other accounts of categorical needs.

## 2. TWO OBJECTIONS TO PERSONAL NEEDS

### 2.1. *Skepticism about Inescapability*

The first of these objections is that personal needs cannot be sufficiently *inescapable* to count as categorical needs; in other words, they will not be able to fulfill A<sub>3</sub> (Inescapable). Here critics may acknowledge that what matters personally to a person can be particular to themselves, but doubt that it can generate binding commitments. Lying behind this skepticism is the notion that persons *control* their personal ends. Critics allow that persons need certain things in order to

14 Arrow, “Exposition of the Theory of Choice under Uncertainty,” 254.

15 Cf. Max-Neef, Elizalde, and Hopenhayn, “Human Scale Development”; Wiggins, “An Idea We Cannot Do Without.” For the same reasons, and as will already have been apparent, commitments and personal needs are not restricted to narrow, self-involving interests. Separating a person’s other-regarding interests from their well-being is often justified for pragmatic reasons in interpersonal evaluative contexts (cf. section 4.1)—namely, to avoid double counting—but it makes little sense in the personal, agential context. A person’s narrow self is only one of their interests. This is again similar to utility theory, which (unless those are screened off for interpersonal purposes) counts a person’s other-regarding preferences among those that contribute to their welfare when satisfied.

attain these ends (i.e., instrumentally), but they think that persons choose for themselves which ends these are. For example, to be a painter a person needs enough money for canvasses, paints, brushes, and so on. That is true. But do they need to be a painter? As Harry Frankfurt puts the thought, although a person may “need the object, since it is indispensable to an end that he desires ... his need for it is his own concoction.”<sup>16</sup> Ends that matter personally to a person, on this view, can neither constitute nor entail categorical needs, because such ends are contingent on the person’s willing them to matter to themselves.

Some accounts of needs do appear vulnerable to this critique. Joseph Raz defends needs he also calls personal needs: “the conditions necessary to enable a person to have the life he or she has set upon”; that is, what a person needs in order to pursue and fulfill their goals. Not satisfying these “will make impossible the continuation of the life the agent has.”<sup>17</sup> The early David Miller similarly proposed “intrinsic needs” entailed by a person’s “life plan” (the “definite and stable idea of the kind of life that he wants to lead”).<sup>18</sup> In holding that some personally valuable aims may count as needs, Raz’s and Miller’s views are on the right track. However, they do not appear to survive the present objection: as Raz and Miller describe them, a person’s aims seem excessively subject to the person’s control. True, once a person has “set upon” a particular course of life, they need certain things in order to continue pursuing it. Yet this condition does not ensure that the person will not later simply change their mind and set upon something else. It is common for people to set upon careers and other projects without truly being committed to them, and later give them up. So, even if a person is entirely sincere, claiming to be committed to an aim does not entail enough inescapability to make it a need.<sup>19</sup> However, although I agree that this objection applies to Raz’s and early Miller’s accounts, that is because they concede too much to the voluntarist view. My account does not share the same weakness, since, as I argue in section 3, it rejects that view: despite being personal, engagements that qualify as “commitments” are not so subject to a person’s will. Before discussing that, however, I consider an objection a defender of essentially moralized needs might make. It partly depends on voluntarism, but also fails for other interesting reasons.

## 2.2. Categorical Needs as Necessarily Essentially Moralized

As mentioned, prevailing accounts of categorical needs seem to assume that,

16 Frankfurt, “Necessity and Desire,” 111; cf. Thomson, *Needs*, 88. See also Gillian Brock’s objections to Frankfurt’s own account, in “Morally Important Needs.”

17 Raz, *The Morality of Freedom*, 152–53, 377.

18 Miller, *Social Justice*, 128–35.

19 Reader, *Needs and Moral Necessity*, 63; cf. Doyal and Gough, *A Theory of Human Need*, 50–54.

necessarily, all categorical needs are essentially moralized needs. On my account, however, essentially moralized needs and personal needs are not necessarily in competition. Each of these subtypes of categorical need has a different function, suitable to different contexts of evaluation. Here I briefly explain why the contrary assumption is incorrect. Ultimately, my aim is not to critique essentially moralized needs, but rather to further distinguish them from personal needs and explain why for my purposes I can justifiably set them aside.<sup>20</sup>

Here is a drastically simplified rationale for essentially moralizing needs. It starts with the idea that certain losses or shortfalls constitute intrinsically morally salient harms to a person, harms that other people who are able to assist are morally required to prevent or ameliorate. Thus Moralized is the initial premise. It is precisely because of Moralized that Universal and Minimal allegedly hold. Universal would hold because morality and justice demand that obligation is impartial and hence uniform. Minimal would hold because people's obligations to promote others' well-being are limited. I say more about Minimal in a moment.

From here, the objection to personal needs must assume that the same concept of categorical needs will be appropriate to all contexts of evaluation. It follows that if the concept of personal needs is inadequate in moral and political contexts, then it is inadequate as a concept of categorical needs, period. That is to say, personal needs would fail to be categorical needs at all. Several related reasons might be given for why personal needs are inadequate in moral and political contexts. These center on the fact that people's aims are idiosyncratic and require differing amounts of resources to pursue, and it is assumed that people's aims are what generate personal needs. One concern is that if people were morally or politically entitled to the satisfaction of their personal needs, they would require an unjustly unequal distribution of resources. Different people would get different shares just because they have different personal aims. Moreover, since some aims are especially resource intensive, satisfying needs related to these aims might require an especially unjust extent of redistribution away from those people whose aims are modest. Furthermore, people might have a personal prerogative to privilege their own interests, which limits how much they are morally obligated to promote others' aims in the first place. Now add the earlier voluntarist skepticism about personal needs' supposed inescapability. If personal needs in fact fail to be inescapable, then far from financing people's aims whatever they are, it seems that justice demands instead that people *change* their aims to ones they can afford within their fair share of entitlements.<sup>21</sup> Similar considerations

20 My thanks to an anonymous reviewer for suggesting that I expand this discussion.

21 On this last idea, that justice expects people to limit their ambitions, see Rawls, "A Kantian Conception of Equality," 97, cited by Scanlon, "Preference and Urgency," 663–64.

led the later David Miller to explicitly recant his earlier account in favor of an essentially moralized conception of needs. He argues that intentions and plans are too “contingent and alterable” to be suitably inescapable, and that only interests that society “validates” qualify as genuine needs.<sup>22</sup> Further strengthening the case that Minimal will be a feature of categorical needs: since the relevant standards are likely to be socially contested, the validated set of needs might extend only so far as “a kind of least common denominator” required for social consensus.<sup>23</sup> The thrust of these objections is that personal needs cannot be genuine needs, because, supposing they were, they would imply implausible moral and political obligations. Worse, some critics allege, claiming to have expansive needs is really just a device for special pleading, a disingenuous means of dressing up mere desires as politically important goals.<sup>24</sup> If there are categorical needs at all, it seems to some, their inescapability is inseparable from a moral/political imperative to ensure some universal minimum.

My account sidesteps these objections. Even setting aside the incorrect voluntarist assumption that personal needs are simply aims, it denies the crucial premise that the concept of categorical needs appropriate to moral and political contexts will have to be appropriate to every other evaluative context. A concept of needs operating in the agential context is not answerable to the same requirements of moral or political adequacy. (I return to the relevance of context in section 4.1.) I reemphasize that I am not aiming to refute essentially moralized needs—this paper neither affirms nor denies they exist. My argument here is only that they need not be the only needs that can fulfill the necessary and sufficient conditions of categorical needs A1–A3. This does not entail that personal needs have no significant political implications—they do. But these are less direct than those that essentially moralized needs allegedly have. As I discuss in section 7, personal needs will not normally be an appropriate standard of well-being evaluation for public purposes, but they have strong implications for which public standards of well-being evaluation *are* appropriate. In any case, I can now safely leave essentially moralized needs behind.

### 3. THE INESCAPABILITY OF PERSONAL NEEDS

This section explains how personal needs’ inescapability can stem from a normative authority that neither derives from nor entails moral obligation. As I soon

22 Miller, *Principles of Social Justice*, 209. David Braybrooke makes precisely this criticism of the early Miller (*Meeting Needs*, 200, 308–9).

23 Goodin, “The Priority of Needs,” 624.

24 Flew, “Wants or Needs, Choices or Commands,” 216; cf. Wiggins, “Claims of Need,” 5.

explain, this inescapability takes the form of “practical necessity,” a special stringency attaching to some of the nonmoral as well as moral practical considerations a person faces, which often manifests phenomenologically. By “practical consideration” I mean a consideration that counts in favor of a particular person’s acting a certain way at some time and place. A practically necessary practical consideration becomes a “practical requirement.” Such considerations and requirements might otherwise be termed “reasons for action,” but I want to avoid any possible connotation that they must derive from a faculty of reason and/or that every rational agent faces the same considerations, even when persons are identically situated. They are *inputs* to practical reason, which may have sources that are arational and contingent upon people’s variable psycho-physical constitutions. As I argue below, however, they need be no less normatively authoritative for that.

The proposal that nonmoral personal interests might entail practical requirements contradicts the voluntarist view of nonmoral personal interests introduced in section 2.1. According to that view, people’s personal interests are nothing more than freely adopted and pursued goals, aims, or projects. However, while that is typically true of people’s less important ongoing engagements (e.g., hobbies or leisure activities, or careers pursued solely for instrumental reasons), the voluntarist view ignores the fact that people often also find themselves with allegiances and under demands they have not chosen but cannot deny. In the special sense I used in section 2, many people also have “commitments,” in the form of personal engagements that are or become beyond their control, and that they cannot rid themselves of at will. Such commitments are in this way inescapable. Common examples constituting commitments so defined, for some people, include special roles, certain projects, ideals, religious and cultural traditions, identities, relationships with family, friends, partners, communities, and so on, causes, and vocations. However, I cannot specify which concrete engagements are commitments and which others count only as hobbies, pastimes, or dispensable projects. For one thing, people’s hobbies and commitments vary from person to person; second, the same type of engagement may be only a hobby for one person and a commitment for another. Some people may paint, or garden, just because they enjoy doing so. Such pursuits might dominate a person’s leisure time, but by assumption they could give them up and do something equally or more pleasurable. For some dedicated artists or horticulturists, by contrast, it may be that they cannot but paint or garden; they pursue painting or gardening in particular for its own sake, and nothing can do as well. If that is the case, painting or gardening is among the second persons’ commitments. The courses of action these commitments dictate are practical requirements. The practical requirements a person’s commitments generate are the person’s personal needs;

they are the things the person needs to do with their life.<sup>25</sup> However, I do not need to make the stronger claim that every person has any commitments at all. To some people, things may only ever be generically good, with nothing in particular they need to do in life. That people may vary not only in their commitments but also in this respect owes to the “subject dependence” of commitments and personal needs that condition B2\*\* (Particular) expresses, and which I focus on in section 4.3.

As we will see is very important (section 6.4), commitments, and personal needs themselves, may be more or less specific. A person’s commitments might be more general than painting: to a form of personal creative expression, perhaps, that might be pursued through different branches of art. A commitment to gardening might lead a person to a career with a botanical garden or arboretum, or otherwise perhaps to dedication to the upkeep and improvement of a community garden. In the other direction, in some cases it might be the creation of highly particular works or projects that are compelling. Close personal relationships tend to be similarly radically specific.

In supposing that at least some people have commitments, this proposal differs substantially from the voluntarist view. However, this is not to say that adopted engagements and commitments are mutually exclusive. An engagement might initially be entirely freely chosen, only subsequently becoming a commitment once the person is entangled with it. Moreover, the difference between voluntary choice and necessity might itself be vague. Even ostensibly willingly

25 The voluntarist view appears to be a common feature of a liberal outlook (e.g., Rawls, “A Kantian Conception of Equality,” 97; Nagel, *The View from Nowhere*, 168). By contrast, this account might seem to have a communitarian flavor in its insistence on limits to free choice and inclusion of social engagements among common examples of commitments. However, while persons’ most important commitments are indeed often interpersonal relationships, this account describes a value relation that is not necessarily communal—commitments may be personal engagements that do not essentially require others’ involvement. Moreover, while it is true that commitments phenomenologically manifest a sense of externality, the limits they entail are not “other” to the person; they exist by virtue of certain objects mattering specifically to them personally in a special way (more on this soon). Thus commitments do not include socially enforced demands that fail to correspond to what truly matters to individuals. (Not to say that some of a person’s truly own commitments cannot conflict with and oppress other commitments the person has—a tyrannical love, perhaps.) So my account should not be read as a critique of liberalism as such. Indeed, in fairness to Rawls, he similarly rejects as unrealistic an image of the person without “devotion to specific final ends and adopted (or affirmed) values” (“Social Unity and Primary Goods,” 181). The compatibility of Rawls’s view on aims and projects with my account would depend on how binding such devotion is upon the person. Moreover, my account’s defense of incommensurability (section 6) is highly congenial to pluralist liberals, e.g., Berlin and Williams, “Pluralism and Liberalism.”



initiated engagements are often not the result of dispassionately considering alternative possible applications of one's talents, or comparing one's prospective compatibility with various possible friends or partners, for example. Our central engagements are often thrust upon us, where it can seem, as James Griffin observes, "they choose us."<sup>26</sup>

Indeed, the evidence that some people have commitments, not just hobbies and freely dispensable projects, is phenomenological—it comes from how some people experience some of the things that matter to them. In this experience, practical considerations can sometimes have the vivid character of absolute demands; they can *feel* binding. These are experiences of the "practical necessity" I mentioned, the term with this sense coming (I believe) from Bernard Williams.<sup>27</sup> Williams argues that some of the best support for *moral* obligations "independent of the will and inclination" comes from such a sense that some actions (including inaction) are impossible.<sup>28</sup> But as Williams also claims, I am arguing that *nonmoral* considerations can also be experienced by a person as practically necessary.<sup>29</sup> Importantly, however, this experience does not seem merely psychological, like an unreasonable brute urge or overwhelming aversion. Although it may sometimes present as irresistible, it is unlike the difficulty one would have with putting one's hand into a meat grinder, no matter the stakes.<sup>30</sup> On the contrary, if a person's will is weakened, then it may be only too easy for them to ignore demands that feel compelling—or else fail to find the necessary motivation. Rather, the compulsion has a normative aspect, the person acknowledging it as having authority over themselves, as reflecting a truth about what they (in particular, and there and then) should or should not do.

Moreover, although cases of moral impossibility are helpfully dramatic illustrations of practical necessity, it is not always so dramatic, tragic, or dilemmical. People's experiences and their negotiation of them is typically quite mundane. Helping a friend in some way or doing something for a community may feel, perhaps quite gently, like something that the person really must do even if not

26 Griffin, *Well-Being*, 54.

27 Williams, *Ethics and the Limits of Philosophy*, 196.

28 Williams, "Conflicts in Values," 75. Harry Frankfurt suggests that the "volitional necessity" he discusses can have a similar character ("Rationality and the Unthinkable," 182). He discusses what he calls "Luther cases," after Martin Luther's declaration: "Here I stand, I can do no other" (Frankfurt, "The Importance of What We Care About," 86; cf. Watson, "Volitional Necessities," 100–101).

29 Williams, *Ethics and the Limits of Philosophy*, 196.

30 Véronique Munoz-Dardé considers what persons need for a flourishing life, and likens the impossibility of foregoing such needs to a psychological impossibility of this kind ("In the Face of Austerity," 232–33).



morally demanded. A person's sense that they should get on with carrying out a personal project, though subtle, may also have the character of a demand. Even so, in unfortunate circumstances the sense of these demands' necessity may come through more powerfully. There remains the constant possibility that neglecting even otherwise mundane requirements can lead to deeply regrettable mistakes and irrecoverable losses.<sup>31</sup>

Yet an experience of practical necessity is neither a necessary nor sufficient condition for having a commitment. Another point of commitments' divergence from voluntary aims and hobbies is that the former are "objective" in the minimal sense of  $A_2$  (Real) as stated earlier. That is to say, whereas a person is conscious of their aims and hobbies, which commitments a person has is independent of whatever they actually, presently desire, care about, or believe matters to them. It is something a person may learn about themselves, and may forget. Indeed, this is another of the ways commitments and personal needs are resistant to free choice: like many normative truths, which a person has cannot be changed simply by their actually believing differently. So, although a person's experiences of practical necessity are perceptions of their personal needs when veridical, and so can be revelatory, they are also fallible. I am claiming only that experiences of practical necessity are good evidence for the existence of commitments, not that the former necessarily entail the latter.

The existence and nature of practical necessity (moral and nonmoral) calls for further investigation, but this paper has a different task.<sup>32</sup> Accepting that some of some people's interests appear to take the form of commitment, it considers how such people's well-being could be structured in a way that respects these appearances.

#### 4. WHAT KIND OF THEORY IS THIS?

This section aims to prevent misunderstandings of the objectives of well-being as personal need satisfaction, especially how WAPNS differs from other types of well-being theory. It makes a second pass over three key features I introduced earlier, situating them now within the philosophy of well-being: that personal needs are specific to an evaluative context (section 4.1); that WAPNS is primarily a structural thesis (section 4.2); and that personal needs are both personal and objective (section 4.3).

31 Cf. section 6.2.

32 For a penetrating discussion of Frankfurt's similar "volitional necessity," see Watson, "Volitional Necessities."

#### 4.1. Contextually Situated

WAPNS does not purport to apply to all contexts of evaluation. While many scholars are unused to specifying any particular evaluative context for their theories of well-being, I take seriously recent work arguing that doing so is indispensable. Some philosophers have noted this idea relatively briefly.<sup>33</sup> It is accepted as natural by investigators of well-being in other disciplines.<sup>34</sup> However, to date it is most fully developed in the work of Anna Alexandrova.<sup>35</sup> On Alexandrova's account, different concepts of well-being are appropriate to different contexts of evaluation, and these contexts and hence concepts are many and diverse. Crucially, context is not defined as some place and time of evaluation. Rather, it is determined by properties of the evaluating agent.<sup>36</sup> These properties include the evaluator's purpose in making the evaluation and the normative relationship between the evaluator and the subject (e.g., clinician to patient, scientist to subject, maternal/paternal, impartial/moral, and government to citizen). The meanings of evaluators' utterances of "well-being" and its cognates vary as these properties and circumstances vary. Alexandrova finds support for the context sensitivity of "well-being" in the actual practices of evaluators across different medical, social-scientific, and psychological fields, which do not aim to theorize well-being as something perspectiveless and all embracing—unlike the ostensibly all-purpose theories that philosophers tend to formulate. Rather, in these practices well-being is narrowly defined, relating to specific theoretical and practical purposes. In addition, often appropriate concepts of well-being are impacted by such practical factors as which forms of measurement are desirable and possible.<sup>37</sup> Other philosophers also express doubts about whether "well-being" is univocal.<sup>38</sup> They observe similar variation across contexts of ethical reasoning, with different well-being concepts relating to different moral, political, partial, and first-personal contexts.<sup>39</sup>

This paper is not the place to defend an account of well-being contextualism. My point is that need concepts appear to vary analogously by context, and that

33 Griffin, *Well-Being*, 1; Scanlon, *What We Owe to Each Other*, 110–43.

34 See, e.g., Gasper, "Human Well-Being" and "Understanding the Diversity of Conceptions of Well-Being and Quality of Life"; Veenhoven, "Subjective Measures of Well-Being."

35 Alexandrova, *A Philosophy for the Science of Well-Being*.

36 The view shares this feature with contextualist accounts of knowledge; in both cases the context-appropriate concept is "speaker relative."

37 Alkire, "The Capability Approach and Well-Being Measurement for Public Policy."

38 Kagan, "Me and My Life"; Campbell, "The Concept of Well-Being"; Feldman, "Two Visions of Welfare."

39 Scanlon, *What We Owe to Each Other*, ch. 3.

adapting such an account to the case of needs is promising. It would be no surprise if the semantics of “well-being” and “needs” were similar in this respect. Even when satisfied needs are not held to constitute aspects of well-being, I take it that needs are always at least essentially related to well-being. The need-satisfaction theory I propose here is specifically about the concept of well-being as a person’s living well according to what matters to them personally—that is, in the agential context. It does not claim to describe public concepts of well-being, or other concepts of advantage, appropriate for various moral or political purposes. As before, among other theories about those concepts, personal needs are not necessarily in competition with existing accounts of categorical needs.

My view of concepts and theories of needs differs a little from Alexandrova’s regarding well-being. Alexandrova thinks that any common core shared by well-being concepts would at most be very minimal.<sup>40</sup> By contrast, on my analysis categorical needs have a determinate common core: it is  $A_1$ – $A_3$ , with different evaluative contexts calling for specific concepts that differently specify  $B_1$ – $B_3$ .<sup>41</sup> Alexandrova is also skeptical about the prospects of a “master theory” of well-being, which would determine and map narrow concepts and intermediate-level theories to specific evaluative contexts. She is not opposed to one, but doubts that one really is possible, and in any case thinks we need not await one in order to address the theoretical needs of specific evaluative contexts. I similarly doubt that a master theory of categorical needs is possible, though my more determinate analysis of the concept may provide some guidance.<sup>42</sup>

There are not only explanatory reasons favoring a different concept of categorical need operating in personal, agential well-being evaluation. Not only is it possible to adequately theorize this concept without at once theorizing every other well-being and categorical need concept. Theorizing well-being specifically in the context of individual agency *matters*. Only a theory that is motivated in that context is capable of resisting the strongest challenges to the existence of non-substitutable interests. As I later discuss (section 6.3), one of these is initially motivated in precisely this same agential context. Furthermore, I argue that the nature of well-being in that context *prevents* the same concept of well-being from being simply exported to other evaluative contexts (section 7).

40 She floats this idea: “well-being is a summary value of goods important to the agent for reasons other than moral, aesthetic and political” (Alexandrova, *A Philosophy for the Science of Well-Being*, 153).

41 Further specifications and combinations thereof are possible besides the ones I have discussed. For example, in some contexts of public well-being measurement an appropriate concept of needs might specify  $B_1$  nonmorally,  $B_2$  universally, and  $B_3$  expansively.

42 See also section 7.2.

#### 4.2. *Primarily a Structural Thesis*

Some philosophers propose that theories of well-being fall into two categories, depending on their theoretical aims: enumerative theories and explanatory theories.<sup>43</sup> Enumerative theories specify which things are good for people, such as pleasure, achievement, knowledge, and friendship. Explanatory theories explain why the things that are good for people are good for people: for example, they are good because people desire them. These two theoretical aims can be pursued separately. A theory may enumerate the good things without providing any explanation for why they are good for people. A theory may explain why things are good for people without saying which those are. A theory of well-being may also do both, conjoining enumeration and explanation. As I said at the outset, WAPNS's primary aim is neither of these: primarily it is a theory of how well-being is *structured*. I propose that structural theories compose a third aim-sorted category of well-being theories—and that these may be developed separately from enumeration and explanation. Two major subcategories are fundamentally monistic theories and irreducibly pluralistic theories. WAPNS is in the latter.

That this third category of theory exists may be obscured by the fact that philosophers usually embed structural claims within otherwise enumerative and/or explanatory theories. For example, John Stuart Mill's theory that some pleasures are "higher" and others "lower" both enumerates two types of good *and* attributes a certain structure to well-being: namely, that some goods are somehow so much better than others that no number of the lower goods can be as good as a single higher good.<sup>44</sup> James Griffin's theory of well-being is enumerative, explanatory, and structural.<sup>45</sup> He both argues that well-being is the fulfillment of informed desires and enumerates a number of goods that a person will desire if they are sufficiently informed. Nothing yet follows about structure; absent further theory, desire and list views are silent on whether well-being is monistically or pluralistically structured. Indeed, even though Griffin enumerates several qualitatively different goods, he argues that the possibility of rational trade-offs between them entails the existence of a unidimensional scale of well-being after all.<sup>46</sup> I discuss this inference in more detail in section 6.3. The structural position Griffin defends is ultimately monistic, then, with a relatively shallow pluralistic

43 Crisp, *Reasons and the Good*, 102; Fletcher, "A Fresh Start for the Objective-List Theory of Well-Being"; Woodard, "Classifying Theories of Welfare."

44 Mill, *Utilitarianism*. Mill's ostensible monism (only utility matters) thus coexists with a form of structural pluralism.

45 Griffin, *Well-Being*.

46 Albeit one that is frequently incomplete. Griffin, *Well-Being*, 90 and ch. 6.

vener.<sup>47</sup> This is unlike, for example, John Finnis's and Martha Nussbaum's accounts.<sup>48</sup> These likewise enumerate and explain, but also embed the claim that well-being's multiple elements are deeply plural: goods of different kinds cannot be substituted in value, and hence a unidimensional scale is unavailable.

Nevertheless, structural theses have been formulated independently of explanatory and enumerative positions, as in the notable example of John Broome's account.<sup>49</sup> Broome formally derives a constructed monism (similar to Griffin's informally inferred position)—while abstaining as far as possible from any particular account of which things are good or why they are good for people.<sup>50</sup> Such abstention is advantageous and deliberate. Broome extracts and defends only a structure that is commonly assumed by preference-satisfaction theories (notably orthodox welfare economics) and consequentialist moral theories, which allows him to avoid complications and objections specific to those theories. WAPNS is likewise primarily a structural thesis because it defines commitments and personal needs in terms of a certain *form* that practical considerations can take. It likewise leaves open both which commitments people have and why, ultimately, people have commitments at all. In my case, only a structural theory is necessary to achieve my motivating goal: accounting for the apparent non-substitutability of some of at least some people's interests. By focusing on structure my account is less open to criticism for unrelated reasons.<sup>51</sup> In the following section I do discuss some highly contentious possible supplements to WAPNS; however, these are optional, and not part of WAPNS itself. Indeed, I offer WAPNS in an ecumenical spirit: something philosophers of different enumerative and explanatory persuasions might accept and build into their own more committed theories. WAPNS's primary adversaries are not those philosophers' accounts, then, but other structural theses—above all the formal structural monism of the type Broome defends (section 6.3). The two theories have precisely the same structural aspirations but draw precisely opposite conclusions.

47 Cf. Mason, "Value Pluralism."

48 Finnis, *Natural Law and Natural Rights*; and Nussbaum, *Creating Capabilities*. Finnis and Nussbaum frame their accounts in terms of a person's "basic values" and "central capabilities," respectively. They are nevertheless describing essential elements of persons' good or well-being specific to the evaluative contexts they address.

49 Broome, *Weighing Goods*.

50 Broome, *Weighing Goods*, 18–20, 32.

51 Rawls has called this strategy the "method of avoidance" ("Justice as Fairness," 231), which we might also call the "don't pick fights you don't have to" principle.

### 4.3. Enumerative and Explanatory Connections

Although the substance and explanatory basis of well-being are not WAPNS's main targets, it is not entirely neutral about them. Moreover, some theories may more naturally complement WAPNS than others.

One possible explanatory complement is an idealized-attitude-satisfaction theory. Such a theory holds that well-being is the satisfaction of certain attitudes the person (or a close counterpart) would have when placed in suitably idealized circumstances. Different versions select different attitudes, such as what the person would desire, or value, or care about; different idealizations include being fully informed, fully rational, and having undergone "cognitive psychotherapy."<sup>52</sup> Such theories allow that a person's well-being can diverge from what the person actually desires or believes it to be. So they secure the minimal objectivity that condition Real expresses. At the same time, the person's well-being remains dependent on what *they* (or at least a close counterpart) would desire, believe, care about, or value. So they also secure the "subject dependence" that Particular expresses, that is, that a person's personal needs and commitments depend on features of the particular person who has them. An example taking this suggestion could hold that a person's commitments are what the person would care about in the right circumstances. Conjoining WAPNS adds the supposition that such caring has the character of practical necessity, entailing practical requirements that are inescapable.

An idealized-attitude-satisfaction explanatory theory could fit well with WAPNS in these respects. Many philosophers are happy with such theories, so perhaps they would be happy with this combination. However, I think other reasons favor a different sort of explanatory theory.<sup>53</sup> Claims that well-being depends on people's attitudes—even idealized attitudes—face the objection that evaluative attitudes have a sort of "objective feel." When a person considers an object *worthy* of desiring or caring about, that seems like something about the object to be discovered—something the person forms a judgment *about*, that they respond *to*. Correct evaluation seems to depend not on the *person's dispositions to value* a thing, but rather on what the thing is *really like*, prior to the person's being disposed to appreciate that.<sup>54</sup> Practical necessity as I interpret it shares this character. On the other hand, some think that a good's objective feel strongly suggests that its value is "impersonal," that is, independent of facts about particular subjects. But as concerns the good specifically of persons, this con-

52 E.g., Brandt, *A Theory of the Good and the Right*.

53 I thank an anonymous reviewer for suggesting that I expand on the view to follow.

54 Cf. Brink, *Moral Realism and the Foundations of Ethics*, 225; Griffin, *Value Judgement*, 28–29.

clusion is antithetical to personal needs. WAPNS agrees with other philosophers who argue that impersonal accounts of persons' good implausibly attenuate the sense that their good is *theirs*.<sup>55</sup> If commitments were subject independent then the needs they entail would fail to be personal. So, WAPNS might appear forced to choose between objectivity and subject dependence.

However, that is not so. As Peter Railton notably recognizes, attitude dependence is not the only form of subject dependence. Furthermore, I add, the objective-feel objection only tells against attitude dependence. Railton partially agrees with attitude-dependence accounts. He argues that what matters to a person, the person's good, aligns with what a certain idealized counterpart to the person would endorse for them.<sup>56</sup> However, these endorsements do not *determine* the person's good. Rather, the determinants of a person's good are those natural facts about the person and the world that would *give rise* to those endorsements (i.e., cause and explain them).<sup>57</sup> These are facts about the particular person's psychological and physical constitution, as well as the circumstances they occupy.<sup>58</sup> Railton's view would underwrite Real and Minimal just as well as ideal-attitude-satisfaction views, and I think has the added advantage of accounting for evaluative judgments' objective feel. Similar to Railton's view of idealized endorsements generally, experiences of practical necessity are most plausibly not what determine a person's commitments and personal needs. These experiences are more likely similarly "indicators" of what matters to the person, serving a "heuristic function."<sup>59</sup> The real determinants, which underlie practical necessity, are more likely the person's actual psycho-physiological state and circumstances. I believe practical necessity has an objective feel because through it a person confronts who they really are, which, though not immutable, cannot be changed at will.

What might such determinants be? I suspect psychological drives play a large role. I have in mind drives to altruism, understanding, self-expression, ar-

55 Although I cannot defend it here, I also accept the "internalist requirement" that, lest a person's good be "alien" to them, it necessarily must engage or resonate with them in suitable circumstances (Railton, "Facts and Values," 9; Rosati, "Internalism and the Good for a Person"). I think it must be possible for a person to experience their personal needs' practical necessity.

56 Railton's usages appear to accord with my understanding of "what matters to a person" (section 1): the ends that determine what the person's good or living well consists of. A person's good is living in accordance with what these ends recommend or require of the person.

57 The endorsements' "reduction basis."

58 Railton, "Facts and Values," 25, and "Moral Realism," 175–76. This does not entail that a person's attitudes are never among these.

59 Railton, "Facts and Values," 25.



tistic creation, technical mastery, athletic excellence, maintaining traditions or practices, and possibly finer-grained drives than these. More specific commitments may partly be determined by the way other facts about the person and their circumstances channel these drives toward their most complete fulfillment. Among the facts channeling a person's vocational-type commitments will be their endowments, abilities, and opportunities.<sup>60</sup> A person may discover what "they were made to do"—no nonnaturalistic teleology required. Commitments to personal and social relationships would have different causes. The necessities involved are those of love, true friendship, loyalty, and allegiance, which are radically particular and often accidental. In these cases, facts about a person's *history* would be central: paths along which attachments, dependencies, vulnerabilities, and acknowledged obligations develop.

I sketch this contentious and underdeveloped account only to indicate how I would seek to complement WAPNS. I reemphasize that it is no part of WAPNS itself, which is open to other possibilities. Theories proposing universally shared elements of well-being, for example, might also usefully incorporate WAPNS (though admittedly deleting the second condition in B2\*\*), by formally representing those elements as commitments that all persons necessarily have. These might be explained by facts about some uniformly shared human or rational nature. For others who wish to account for cross-personal variation, "commitments" can serve as a label for values that vary across persons but that are nevertheless objective, in the sense that they do not vary due to differences in mere preferences or other subjective attitudes. That WAPNS does not commit either way in this regard may be advantageous. With it, people with different metaethical inclinations may be able to agree that well-being is objective to an extent, and that it has a certain structure—without needing to agree on well-being's deeper nature.

Briefly now regarding enumeration, it is true that WAPNS claims that certain kinds of things constitute commitments for some people (special roles, tasks, vocations, special relationships, connections with traditions, cultures, etc.). This is unlike the typical philosophical enumerative theory, however, which purportedly describes the well-being of the human being as such. WAPNS does not claim that persons have any particular commitments. WAPNS does not claim that the above kinds of things necessarily formally constitute commitments for every person. WAPNS does not even claim that every person has commitments at all. Those assumptions are not mandatory for anyone, and my theoretical goals do not require them. I hope it is now clear enough what those goals broadly are.

60 Compare Railton, "Facts and Values," 26–28.

5. FIRST STRUCTURAL CLAIM:  
NON-SUBSTITUTABILITY OF NEEDS BY NON-NEEDS

5.1. *Satisfying Needs Is Strongly Superior to Satisfying Non-needs*

WAPNS distinguishes between commitments on the one hand and non-need goods (“non-needs”) on the other. To be clear, non-needs are not just any unnecessary items. A necessary means to a worthless end is likewise worthless, so not a good and not a non-need in my terms. A non-need here is finally good, but is not finally necessary. WAPNS’s first major structural claim is that the division between commitments and non-needs is irreducibly one of kind, not quantity. This proposal resembles the familiar idea that some of a person’s interests possess a qualitatively distinct significance, lacked by other interests. Specific proposals have called such interests “simply, important,” “heavyweight,” “central,” and “global.”<sup>61</sup> More than this, WAPNS mirrors the common further claim that there is no number of noncentral goods that it is better for a person to have than satisfying their central interests. In contemporary terms the latter are “strongly superior” to the former.<sup>62</sup> WAPNS holds that, in any context of decision, no number of satisfied non-needs can take priority over a person’s satisfying their needs. The language of decision and priority suits WAPNS better, because it more clearly avoids any suggestion that commitments and personal needs can be evaluated by how good or better they are in the abstract (see section 6). Whichever terms we use, the claim that well-being is structured by such a difference between types of good is not novel.

WAPNS’s contribution is how it supports such a division. A tenable strong-superiority claim cannot allow the value difference between central and noncentral interests to be ultimately reducible to a magnitude, however large. It requires an account of why the value difference between the goods is irreducibly nonquantitative—why else is one type superior to the other, if not that it is simply greater in value?<sup>63</sup> WAPNS’s distinction between commitments and non-needs fulfills this requirement. Commitments’ place in a person’s well-being lies not in their incrementally increasing it, but in their imposition of inescapable requirements (personal needs). So, not only does this make commitments and non-needs irreducibly qualitatively distinct parts of well-being, the inescapability of personal

61 Williams, “A Critique of Utilitarianism” and *Ethics and the Limits of Philosophy*, 182–83; Griffin, *Well-Being*, 45–46; Nussbaum, *Creating Capabilities*; Dorsey, “Headaches, Lives, and Value.”

62 Arrhenius, “Superiority in Value”; Dorsey, “Headaches, Lives, and Value.”

63 Compare Julius Schönherr’s (“Still Lives for Headaches”) effective critique of Dale Dorsey’s (“Headaches, Lives, and Value”) strong-superiority proposal.

needs also overrides the mere attractions of non-needs. Section 6 expands on both points in detail.<sup>64</sup>

Although WAPNS concentrates on defending the existence of commitments/personal needs, it does not rule out other interests besides commitments in some way also being finally necessary in at least some people's lives—such things as being truly understood by others and sexual fulfillment, perhaps. In the following subsection I consider whether another need might be some degree of pleasure.<sup>65</sup>

### 5.2. Needs and Pleasure

Just as on the issue of what, specifically, people's commitments are, WAPNS itself does not take a stand on which goods are non-needs. That is the role of an enumerative theory. Still, the distinction may seem undermotivated without a concrete example, even if it is no part of WAPNS proper.

I think many people's non-needs include pure pleasure and sensory gratification in general, although my discussion will focus on the former. This is not to deny that such experiences are *good*, nor that other things equal it is better to have more of them. The point is that their attractions lack the sense of normative authority bestowed by practical necessity. Baldly stated, this position may nevertheless seem implausibly ascetic or puritanical; however, considering the nature of pleasure and of the needs in question suggests it is not.

First, the qualification *pure* is important. Pure pleasure is pleasure unattached to otherwise valuable ends; in its case pleasure itself is the end, and the aim pursued is worthwhile just insofar as it yields pleasure. This pleasure is unlike the satisfactions gained from fulfilling independently worthwhile ends. In these cases the end is not sought for the sake of the pleasure it brings; the plea-

64 There is an important objection to strong superiority that I take very seriously but cannot discuss in this paper. This is that alleged strong superiorities are far less plausible when the benefits and costs at stake are risky, and in real life risk is omnipresent. People seek relatively trivial goods despite risking extreme losses (e.g., death)—and it seems perfectly rational to do so if the risk is sufficiently small. This, critics conclude, is inconsistent with the serious good risk being strongly superior to the trivial goods. They allege that risk-based arguments such as this decisively support aggregating individually trivial benefits/costs together with serious ones, at least in risky cases (Norcross, "Comparing Harms"; Bailey, "Is It Rational to Maximize?"; Fried, *Facing Up to Scarcity*; Horton, "Aggregation, Risk, and Reductio"). A response I intend to make in future work points to the typical *necessity* of such risks, when justified, in view of the person's life as a whole—needs are risked for other needs. I will show that strong superiority is much more defensible when normative necessities embrace not only minimal goods and moralized harms, but also final ends playing a central and organizing role in people's lives.

65 Thanks to an anonymous reviewer for recommending that I discuss this.

sure is *taken in* the fulfillment, a reflection of the end's final value. So prioritizing independently worthwhile ends such as personal needs ahead of pure pleasure does not entail sacrificing pleasure *per se*. A life satisfying personal needs will be full of pleasures, often deep, taken in those satisfactions. Similarly, interpersonal relationships valued for their own sake may be categorically more important than pure pleasure, but nevertheless partially constituted by shared pleasures. Consider the purpose of celebrations, and the difference between going to a fair-ground with friends or family and going alone.

Second, *complex* pleasures, even if valued partly purely for their own sake, are also the intelligible objects of commitments. Aiming to be a connoisseur of certain goods or experiences might constitute a genuine commitment for some person. Actively seeking out complex experiences, critiquing them, and gathering a store of knowledge about them is not just passively gratifying, but rather an actively pursued project that might plausibly entail personal needs on the person. Complex pleasures are also often bound up with traditions. Distinguishing "higher" from "lower" pleasures (as on Mill's view), which I endorse here, is sometimes regarded as elitist.<sup>66</sup> However, that danger is avoided if we allow that different things objectively matter to different people (section 4.3). Being an aficionado of science-fiction B-movies might be as much a more than merely gratifying project for one person as cultivating a taste for and knowledge of fine whiskies is for another.

Third, even if no single episode of pure pleasure could be categorically necessary, Richard Arneson might be right that some sufficient amount of "cheap thrills" is somehow essential to a person's well-being.<sup>67</sup> So, the relative priority in some situation of a commitment *vis-à-vis* pure pleasure might partly depend on whether the person already has enough pure pleasure in their life. An extreme case: a person faces a fork in their life between an unremittingly dull grind necessary to live up to some commitment and a more pleasant life that abandons that commitment. Here we would have a need confronting another need, and we have not yet considered how conflicts among needs might play out. It is also worth noting that it could make a difference in varying but related cases whether having enough pleasure is *finally* necessary or else only instrumentally required—say for one's sanity and capacity to pursue other ends that *are* final. If the more pleasant life won out, but *only* because the dull grind was psychologically crushingly difficult, then we would know the following: if the dullness were at least bearable, and a genuine commitment were at stake, then in that case abandoning the commitment would not be the better choice.

66 Mill, *Utilitarianism*.

67 Arneson, "Human Flourishing versus Desire Satisfaction," 120.

6. SECOND STRUCTURAL CLAIM:  
NON-SUBSTITUTABILITY OF NEEDS BY OTHER NEEDS

6.1. *Commitments Are Irreducibly Plural*

The following two possible structural features are compatible: (i) well-being being divided into two or more qualitatively irreducible types of good (e.g., needs and non-needs), and (ii) the goods within each type ultimately being only quantitatively distinct in final value from each other. In other words, irreducibly separate types of good could still be quantitatively measurable along scales of final value, their separateness meaning only that they lie on different scales. Applied to the present case: despite their separateness from non-needs, personal needs might all still lie on a scalar dimension of their own. Personal needs' place on the scale could depend on how much of a common value they possessed or realized—degree or strength of “neededness.” If that were the case, then the satisfactions of different personal needs could substitute for each other in final value, and WAPNS would be false.<sup>68</sup>

That is not the case, however. Commitments are not only irreducibly separate from non-needs; they are also irreducibly separate from one another. This claim is not an independent addition to WAPNS. It follows from the same reason that commitments and personal needs are separate from, and superior to, non-needs: namely, the finality and inescapability of personal needs' necessity. That *each* commitment's demands are inescapable means that the latter are not answerable to *anything* else that might be achieved by acting differently. So, the attractions of non-needs are irrelevant to the bindingness of a given commitment's demands—yet equally so are the demands of the person's *other* commitments. Commitments' satisfactions are *severally* finally necessary, each for its own sake and no other. They are not jointly necessary for any further purpose that they might together serve. This, then, is what precludes placing personal needs on a single dimension of neededness. Such a dimension would falsely presuppose that personal needs' values derive from something they all possess or yield when satisfied, or a purpose that commitments serve in common.

Personal needs and commitments could sensibly be called “incommensurable” with each other. Yet this status differs from the forms of incommensurability philosophers today more frequently discuss, such as items' or outcomes' relative values or orderings being *vague*, or else *incomparable* with respect to some scale

68 More generally, anything called a “need” fails in fact to be necessary whenever its value consists in how *much* it is “worth” or how *much* good/harm meeting it does/averts. In such cases, talk of need is redundant (cf. Fletcher, “Needing and Necessity,” sec. 5).

or value. In WAPNS, it is not that there is some difficulty or indeterminacy in placing personal needs' and commitments' satisfactions on a scale of neededness (or contribution to well-being generally), but rather that they are what Ruth Chang terms "noncomparable" with respect to it: the basis of the desired comparison itself is unavailable.<sup>69</sup> Commitments and personal needs are incommensurable because it makes no sense even to imagine placing their satisfactions on such a dimension; the dimension itself is at fault.<sup>70</sup> Commitments' irreducible separateness, and consequent non-comparability, ensures the non-substitutability of their satisfactions by each other. Their substitutability would require a common denominator, but no such thing can be defined even in principle (or even formally, see below).

Unlike the relation between commitments and non-needs, however, commitments are not generally related to each other by further hierarchies of strong superiority (not to rule out such relations ever *also* existing). Since all entail inescapable demands, commitments are in that respect all on a par.<sup>71</sup> I discuss how else they relate in sections 6.3 and 6.4.

## 6.2. Necessity and Incomplete Lives

Commitments' irreducible plurality and noncomparability implies a sense in which a person's satisfying their commitments is "essential to" their well-being. One way of interpreting this "essential to" relation has absurd consequences. On that interpretation, satisfying every commitment is necessary in order for a person to count as having a good life. It follows that a person would simply fail to live a good life whenever even one of their personal needs went unsatisfied. This picture makes well-being out to be implausibly binary: a person either has a good life or they do not. Surely people can live quite good lives even when some of their central interests are unfulfilled (by far people's usual condition).<sup>72</sup>

WAPNS agrees. A person's life can indeed be quite good despite some of their commitments being unsatisfied. However, it matters how we understand this. As before, commitments are distinguished from non-needs by all sharing a crucial

69 Chang, "Introduction," 29. Chang designates noncomparability a "formal failure of comparability," and argues that "practical reason never confronts agents with comparisons that could formally fail" ("Introduction," 29). (For this reason, her discussion is relatively brief, and I have not seen other authors take it up.) That might be so, but it does not prevent anyone from falsely *believing* that certain things are comparable, when they are in fact non-comparable, in terms of some value.

70 Cf. Berlin, *Liberty*, 216; Wiggins, "Deliberation and Practical Reason," 362–63; Stocker, *Plural and Conflicting Values*, 177; Kekes, *The Morality of Pluralism*, 56.

71 I do not mean in Chang's technical sense of parity ("Introduction").

72 I am grateful to an anonymous reviewer for pressing me to address this concern.

characteristic: they all entail inescapable requirements, that is, personal needs. This is what gives all commitments their paramount importance, in the form of their strict priority over the dispensable attractions of non-needs. However, this is not a characteristic that commitments fulfill in lesser or greater amounts. Again, this importance they share does not designate a further end that all commitments serve, or a separate “value” that satisfied needs all produce, bear, or contribute to in varying degrees. In particular, “well-being” is not such an end, something further and separate from satisfied commitments. Rather, satisfied needs and commitments each *constitute* a central part of well-being, and separately so. There is nothing to this part of a person’s well-being *over and above* the several satisfactions of their disparate commitments. In the personal context WAPNS describes, then, well-being is a composite inseparable from its constituents, and does not come in degrees as such.<sup>73</sup>

While a person’s life can be going better or worse in respect of their commitments, then, that means nothing more than that certain of their commitments are being satisfied or not. True, we can count the number of commitments the person is satisfying. This part of a given person’s present well-being could be placed on an ordinal scale according to that number. Other things equal, it is better for a person to be higher on that scale, fulfilling more practical requirements than fewer. However, the normative pressure to be higher on the scale is not anything separate from the commitments’ several demands. The person does not gain anything their commitments have in common when placed higher on that scale; the scale and their place on it have no significance independent of the importance of meeting their commitments’ demands taken separately. Relatedly, it is not better or worse in itself to have more or fewer commitments. If a person gains a new commitment, that person does not *thereby* come to enhance their well-being in respect of their commitments when they satisfy it, compared to when they had fewer commitments.

Well-being’s fragmented structure should not be considered an unfortunate obstacle. It is necessary for registering and explaining the possibility of losses that are irrecoverable in final value. Again, I take this possibility to be a feature of many people’s self-understandings. This is not the obvious fact that it can become physically impossible to turn back time and replicate or replace certain concrete items or events. As the qualification “in final value” indicates, the phe-

73 The relation between commitments and well-being closely resembles the relation between the virtues and eudaimonia in Aristotle’s *Nicomachean Ethics* under the “inclusive end” interpretation. On that interpretation, eudaimonia is composed of fulfilling the several virtues, rather than forming a further, unitary, ultimate aim that fulfilling the virtues serves (the “dominant end” interpretation) (Ackrill, “Aristotle on *Eudaimonia*”).



nomenon is that certain things essential to a person's life cannot be recouped and canceled out in value by later gains. If the damage to a person's well-being constituted a reduction of some further end that the lost thing served, then the loss could be compensated perfectly by anything that served that further end to the same or better extent. But on the contrary, the damage is a loss of a kind of final value inherent to, and inseparable from, the particular damaged thing.

The consequences of commitments' irreducible plurality are also not all so dour. It also entails that no losses or failures can diminish those commitments that *are* satisfied. Failures do not produce a loss that can be set beside and offset the positive values of satisfactions. Even a single fragment of a person's well-being could be perfect, complete in its own way, and a source of joy, even if the person's life were in other respects damaged beyond repair.

### 6.3. *Non-substitutability, Monism, and Independence*

The assertion that commitments are irreducibly plural is not only or even primarily an ontological claim, a denial that ultimately there "really exists" only one component of well-being. That is just as well, because there is a greater challenge to it than that. A sophisticated objection to the possibility of non-substitutable interests need not rely on metaphysics. The truth of a *formally derived* structural monism I have mentioned previously would suffice to defeat WAPNS. As in Griffin's theory, this variety of structural monism is even consistent with a relatively superficial, merely enumerative pluralism: the objector can allow that the constituents of well-being are qualitatively distinct, while also showing that a unidimensional well-being scale can be formally derived nevertheless. All it needs is that outcomes can be ordered by how good they are for a person—a "betterness" ordering—where this ordering satisfies the axioms of utility theory.<sup>74</sup> It is positively helpful to address this objection, because structural monism also *requires* this to be the case.

Appealing to utility theory's axioms is attractive to those who accept that a person's *rationality* requires their ordering of *preferences* to satisfy them.<sup>75</sup> It takes substantial argument to conclude that a person's well-being aligns with their rational (and presumably adequately informed) preference, but suppose for the sake of argument that can be done.<sup>76</sup> If it can be done, and if a person has a betterness ordering satisfying the axioms, it can be shown that such a quantitative attribute as "good" exists as a formal construction. Despite being only a formal

74 At least generally. A structural monist might allow that the ordering is in some cases incomplete.

75 Following, e.g., Savage, *The Foundations of Statistics*.

76 Broome, *Weighing Goods*, ch. 6.

construction, this attribute functions as a sort of fiat currency. It denominates the values of all the practical considerations favoring an outcome (e.g., satisfying a personal need). How good an outcome is for a person can be represented as if it were the sum of amounts of good borne or produced by the individual considerations favoring the outcome.<sup>77</sup> Differential amounts of this currency that practical considerations bear or produce would represent ratios of substitution between them.

For present purposes I need not rehearse this derivation, because WAPNS does not dispute its validity. WAPNS objects to the applicability of one of the necessary assumptions—variously termed an axiom of “independence” or “separability”—in comparisons involving commitments and personal needs. Structural monism actually requires different but related independence assumptions to apply in several different dimensions of comparison, but I concentrate on the independence or otherwise of practical considerations within outcomes.<sup>78</sup> I will refer to it simply as Independence, and for efficiency, introduce and discuss it in its direct connection with (non-)substitutability and needs.

Take any two (sets of) practical considerations  $x$  and  $y$  bearing on the choiceworthiness of an outcome. An outcome’s choiceworthiness here is the extent to which it is better or worse than others for the person in question to choose. For argument’s sake suppose that the extent to which fulfilling  $x$  increases the outcome’s choiceworthiness relative to fulfilling  $y$  can be represented by the ratio  $c_x/c_y$ , where  $c_x$  and  $c_y$  are real numbers. Independence here is the assumption that  $c_x/c_y$  will be unaffected if any further (sets of) practical considerations (i.e., besides  $x$ s and  $y$ s) also come to bear on the outcome in question.<sup>79</sup> It means,

77 Broome, *Weighing Goods*, chs. 4 and 6. Again, the scale and units of these quantities are irrelevant.

78 Much discussion of independence focuses on whether rational preferences between *alternative outcomes*  $x$  and  $y$  are independent of whether any other *outcome*  $z$  is also available to choose, i.e., the “independence of irrelevant alternatives.” Like Broome (*Weighing Goods*, ch. 5) and many others, I find counterexamples to the rationality of independence in this dimension implausible (cf., famously, Allais, “The Foundations of a Positive Theory of Choice Involving Risk and a Criticism of the Postulates and Axioms of the American School”). Less commonly discussed is my focus here: whether the practical influence of considerations *within the same outcome* are independent of each other. Broome considers one way that independence might fail in this dimension—if the good of an outcome is affected by how a given sum of benefits is distributed across different persons involved—but not the more extensive, intrapersonal failure I discuss presently.

79 Since  $c_x/c_y$  may not be fixed, but rather vary systematically depending on how many  $x$ s and  $y$ s are already present: more precisely and generally, the effect of Independence is that  $c_x/c_y = f(x, y)$  is unaffected by adding any further (sets of) practical considerations. That is,  $c_x/c_y = f(x, y, z) = f(x, y)$  for all possible (sets of) additional practical considerations  $z$ .

then, that  $c_x/c_y$  comes to represent not only  $x$ 's and  $y$ 's relative bearing on the choiceworthiness of the outcome at hand. That ratio is converted into a ratio of how much each  $x$  and  $y$  are *worth*, representable by differential amounts of a currency of substitution that is valid in any circumstances of choice.<sup>80</sup> As mentioned, well-being itself functions like such a currency—i.e., structural monism is true—if and only if Independence (plus other necessary axioms) holds over a choiceworthiness ordering (now converted into a betterness ordering) that includes all of the good-making things in a person's life. But consider even just the case in which Independence (and the other axioms) always applied to the satisfactions of personal needs and their relative bearing on outcomes' choiceworthiness. Independence here would ensure the formal existence of such a thing as the unidimensionally measurable value of neededness, or commitment-living-up-to. All personal needs/commitments could be represented as possessing it in different amounts, implying rates at which each could be perfectly compensated by others. As I said in section 6.1, if this were the case, then in fact no personal "need" or "commitment" would be any such thing, since satisfying any other (or combination of others) would do equally well in sufficient number. So, Independence must fail to apply to personal needs' and commitments' practical bearing, in order for them to exist at all.

#### 6.4. Comparing Outcomes without Commensurating Practical Considerations

Independence might be considered indispensable to rational comparison for one or both of the following reasons. First, it might be considered indispensable for its own sake, as a requirement of simple consistency (utility theorists often style their axioms this way). This reason is unconvincing. It is true that the ratio  $c_x/c_y$  should consistently be insulated from wider considerations  $z$  just so long as  $z$  are irrelevant; however, whether wider considerations are sometimes relevant to  $c_x/c_y$  is precisely what is in question. And if they are, then consistently ignoring them would be *irrational*. Second, Independence might be considered indispensable to rational comparison for what it yields—namely, the formal existence of a commensurating currency. Yet rational choice does not obviously require any such thing. From the possibility of rationally comparing and ordering outcomes, it does not follow that practical considerations are already commensurated ahead of that comparison.<sup>81</sup> An ordering of choiceworthiness could emerge from deliberation without having been predetermined so, and in-

80 N.b., "worth" here does not designate subject-independent value specifically; if the considerations are subject dependent, it means "worth to the person."

81 Wiggins, "Incommensurability"; Hurley, *Natural Reasons*; Richardson, *Practical Reasoning about Final Ends*. Equivalently, from the incommensurability of  $A$  and  $B$  in value (specific-

stead “sum up deliberation effected by quite other means.”<sup>82</sup> Still, I accept that this reply will only carry conviction if we can say what such other means might be—some alternative mode of rationally ranking outcomes that does not rely on predetermined measures of how much the things involved are worth. Indeed, I suspect that the ultimate reason for structural monism’s predominance is this: a presumption that *there is no alternative* to rational choice consisting in maximizing such a commensurating value (or being representable as such). Some explicitly claim as much.<sup>83</sup>

In fact, negating Independence over personal needs and commitments is informative. It entails that they instead stand in some holistic relation of *interdependence*—meaning that personal needs’ effects on an outcome’s overall choiceworthiness do depend on which other personal needs are at stake in those circumstances. That is to say, somehow what determines whether one outcome involving personal needs should be chosen over another must be how different fulfillments of those needs fit together, not how they sum. The truth of such a holism would block the conversion of comparisons of choiceworthiness in particular situations into verdicts about some circumstance-independent quantitative value that satisfied personal needs possess or yield.<sup>84</sup> If a satisfactory account of holistic practical reason can be given, then, the non-substitutability of certain interests is vindicated.

Not enough work to this end has been done, perhaps partly because the no-alternative belief is widespread. But Henry Richardson’s account of specificationist deliberation provides a strong start.<sup>85</sup> Specification exploits the fact that norms, principles, ends, and so on are often relatively general, not tied to concrete particulars. Suppose a person goes to a restaurant just because they want something for dinner. That particular restaurant is their specific aim, but only because it *counts as* a specification of the general aim. In a real-life case,

---

ly their *noncomparability*), it does not follow that *outcomes* featuring *A* and *outcomes* featuring *B* are *incomparable* (i.e., by choiceworthiness).

82 Wiggins, “Incommensurability,” 361.

83 Bailey, “Is It Rational to Maximize?”

84 It is always possible to preserve Independence by redescribing outcomes in ways that take holistic effects into account. However, doing so risks triviality and would ensure it here (Broome, *Weighing Goods*, 107–10, 186–92; Hurley, *Natural Reasons*, 264; and Wiggins, “Incommensurability,” 360–61, 370–71).

85 Richardson, “Specifying Norms as a Way to Resolve Concrete Ethical Problems” and *Practical Reasoning about Final Ends*. Richardson builds on ideas from Dewey, “Human Nature and Conduct”; Kolnai, “Deliberation Is of Ends”; and Wiggins, “Deliberation and Practical Reason” (interpreting Aristotle). Specification is not the only mode of holistic deliberation (cf. Richardson, *Practical Reasoning about Final Ends*, ch. 7), but it is especially powerful.

other considerations will probably also favor that particular restaurant. However, as far as this general aim is concerned, any other restaurant will do—as will eating at home, even. If the restaurant is closed, they can go somewhere else and still satisfy that aim. Apparent conflicts between a person's commitments and personal needs may also often be resolvable through specification; commitments' requirements are often quite specific, but need not be maximally so. If the commitments at stake (and/or the particular needs a given commitment entails) are sufficiently nonspecific, then the apparent conflict might be resolved by specifying them in new ways that are mutually compatible. A toy example: a young person drawn to a career as a doctor might also be drawn to a career as a scientist. Both might be independently compelling for this person, non-substitutable. However, ultimately what are non-substitutable may not be those specific careers, but more general commitments to humanitarian aid and scientific inquiry. Rather than evaluating the person's choice between doctor and scientist in terms of how *much* well-being they would give the person (or the *strengths* of the reasons favoring each), choice may be guided by the possibility of satisfying both nonnegotiable commitments—how to *reconcile* rather than simply *arbitrate* between them.<sup>86</sup> This often involves rejecting a given menu of alternatives and thinking creatively—really deliberating.<sup>87</sup> Perhaps becoming an immunologist would constitute a specification of both commitments; the person could investigate the nature of human beings and their pathogens while potentially also improving many people's health. Much more needs to be said about how specification works (especially when applied in complicated scenarios and its implications for interpersonal decision making), but the example illustrates some of the proposal's essential features. First, specification's holism: the acceptability of any particular new specification of a personal need partly depends on the availability of appropriate compossible specifications of the other commitments/personal needs at stake. In structural monism, a practical consideration contributes the same increase to an outcome's total choiceworthiness irrespective of how it is specified (side effects excepted); in specificationist deliberation, any given consideration's practical influence cannot be evaluated independently of its compatibility with other considerations. Second, although specification admits a lot of flexibility, the commitments and personal needs in their general forms are not what are exchanged; *they* remain non-substitutable. Relatedly, there is no claim that appropriate specifications necessarily are available. Whether they are de-

86 Cf. Richardson, *Practical Reasoning about Final Ends*, secs. 18 and 20.

87 N.b., this is not to say that a creative outcome arrives *ex nihilo*. It remains a discovery about something—namely, about the logical and physical possibilities that the actual state of the person and the world enable.

depends on how forgiving the world is, and on how general or specific the person's commitments and needs are. So there remains the frequently realized possibility of irrecoverable losses being unavoidable—and this is realistic.

## 7. PRACTICAL RAMIFICATIONS

### 7.1. *Undercutting Aggregation at Its Foundations*

Despite being immediately situated in a personal evaluative context, WAPNS has strong indirect implications for theorizing interpersonal contexts. It bears on theories that primarily operate in interpersonal contexts, but that nevertheless critically rely on assumptions about well-being at a personal level. A clear example is classical utilitarianism, which simply extrapolates intrapersonal evaluation to interpersonal evaluation, “extend[ing] to society the principle of choice for one [person].”<sup>88</sup> In that theory, an individual's well-being is a quantity of personal utility, and social well-being is just the sum of personal utilities. Yet it is only if Independence is assumed in the personal case that well-being can be represented by anything that, like utility, is apt to be summed at all—let alone summed together with other individuals' well-beings. Personal needs halt this and any analogous transition at a basic level, because they are in the same evaluative context as—and so compete with—personal utility.<sup>89</sup> The distinctive significance of personal needs is that they too bear on outcomes in individual-level decision making. If WAPNS describes some people's well-being, and since personal needs do not aggregate and trade off, then the transition from one-person aggregation to multiperson aggregation cannot get started.

This barrier is unlike possible external constraints on interpersonal aggregation that do not themselves rebut *intrapersonal* aggregation (e.g., demands of equality or rights, and technical difficulties with merging different persons' preference orderings). This barrier also differs from that posed by positing a diversity of incommensurable “values” or “elements” or “constituents” of well-being at a high level of generality and ostensible universality. WAPNS's objection to aggregating theories is more challenging, because it confronts the derivation of interpersonal from intrapersonal aggregation on that derivation's own, intrapersonal territory.

A reasonable reply is that much of the “aggregation” discussed in recent mor-

88 Rawls, *A Theory of Justice*, 24.

89 Personal needs compete only with concepts of utility that are (or turn out to be) equated with welfare *qua* well-being, which purely formal decision-theoretic notions of utility may avoid doing. On confusing different interpretations of utility and utility theory, see Broome, “Utility”; and Bermúdez, *Decision Theory and Rationality*, 46–50.

al theory is not the aggregation of well-being as such. WAPNS's bearing on such approaches is accordingly less direct. Some things called aggregation are innocuous. Take outcomes differing only in total number of lives saved. WAPNS does not deny that saving the greater number there is better. Other things equal, in a given situation, satisfying a greater number of practical requirements than fewer is a better choice. Again, this owes to the several pressures of those requirements, not of any such thing as betterness over and above those, even as a formal construction. So, "numbers count" or aggregate in this sense (and much more needs to be said about this), but it matters crucially what we are counting. Of course, usually other things are not equal, and for this reason contemporary discussions shift to aggregating persons' *claims* (to be benefited or spared from harm by the decision maker's choice) weighted by different *strengths*.<sup>90</sup> There is no space here to discuss WAPNS's bearing on strength-weighted claims properly; however, I can say that its implications would depend on whether pairwise strength comparisons are interpreted as independent of each other or not. If they are, then an argument like that in section 6.3 can be run, and strength-weighted claims reduce to quantities of a welfare-like currency. If they are not, then strengths—now representing only circumstance-dependent contributions to choiceworthiness—are the *outputs* of some form of non-aggregating, holistic deliberation. WAPNS's critique of intrapersonal well-being aggregation lends support to such non-aggregating approaches, and may influence what form they should take.

### 7.2. On the Supposed Indispensability of Unidimensional Measurement

Another possible reply is that scales of overall well-being—and corresponding overall "levels"—are practically indispensable. If that were so, and WAPNS ruled these out, WAPNS might be less plausible. WAPNS does indeed rule out aggregating central parts of personal well-being entirely. Yet the truth here is more complicated than this thought suggests. In the first place, on closer inspection, rationales for unidimensional measurability are often doubtful. A unidimensional measure is neither the only nor obviously the best manner of representing a person's overall well-being in any evaluative context. Once more, clearly there are intermediate states between full attainment and total shortfall of well-being, yet those need not be defined as gradations on a single scale. As with the incomplete personal life (section 6.2), they are often better identifiable by noting the various qualitatively distinct respects in which a person is doing well and falling short. Measurement is not necessarily unidimensional or even scalar at all.<sup>91</sup> Indeed, when different considerations point in different directions, unidi-

90 E.g., Voorhoeve, "How Should We Aggregate Competing Claims?"

91 Cartwright and Runhardt, "Measurement," 271–76.



mensional measures obscure relevant information. Knowing the *different* ways that evaluated subjects improve or deteriorate can be crucially important—ways that fully aggregating measures brush over. Most of the pressure toward unidimensional metrics comes from needing to know what to *do* with this information—that is, determine which policies are best to choose. Yet as before, from the possibility of ranking outcomes it follows neither that the values involved can be condensed into a single value, nor that condensing them is necessary for arriving at that ranking. Other modes of decision are possible, though beyond the scope of this paper.

Concepts and theories of well-being appropriate to different contexts of evaluation are often related, and WAPNS has implications for which public-context concepts, theories, and measures will be appropriate. It provides foundational support for the growing movement toward disaggregated, multidimensional well-being measures.<sup>92</sup> These often treat discrete dimensions as incommensurable, and indeed, in some cases, severally *necessary* for well-being.<sup>93</sup> Disaggregated approaches are potentially vulnerable to critics favoring economic-welfare analysis, which is structurally monistic and rigorously motivated at a foundational level—but also undermined there by WAPNS. Potentially, too, in specific contexts appropriate dimensions to measure might often be identified as the *kinds* of things, under general descriptions, many subjects under investigation require in satisfying their personal needs.<sup>94</sup> That is, macro-level dimensions could sometimes be generalizations about which personal needs those subjects have and the necessary means to satisfying those. Nevertheless, WAPNS is not necessarily opposed to constructing fully aggregating, unidimensional measures, which may remain desirable for certain public purposes, notably, tracking well-being in a population.<sup>95</sup> What matters is *how* these are constructed and interpreted. Such indices require assigning weights to measured dimensions, which might be expected to erase personal needs' non-substitutability. It would indeed, if the weights supposedly reflected ratios of some common final value, that attainments along each dimension all possessed or realized in different quantities (section 6.1). However, that need not be so if the weights instead represent circumstantial social priorities. This is precisely how thoughtful social and policy scientists think of them: in addition to other normative and pragmatic factors,

92 Not least the capability approach to human development.

93 Cf. Fardell, "Conceptualising Capabilities and Dimensions of Advantage as Needs."

94 Cf. Sen, *Inequality Reexamined*, 109.

95 Stiglitz, Sen, and Fitoussi, *Report by the Commission on the Measurement of Economic Performance and Social Progress*.

setting weights is a political decision.<sup>96</sup> The weights' circumstantial nature also limits how accurately indices can be extrapolated over time and across relevantly differing circumstances.<sup>97</sup> It can be useful and worthwhile to retain a given index over a limited period to track changes through time, updating the weights only periodically. But it must be borne in mind that this is a pragmatic choice that may reduce the index's synchronic fidelity.<sup>98</sup> Above all, it can never simply be assumed that such a measure can be carried over for use in different practical applications. Moral and political theory might benefit from better recognizing that the scope for abstraction from specific aggregation and measurement exercises is limited. If nothing else, WAPNS's defense of non-substitutable interests helps to explain at a foundational level why such difficulty exists.

## 8. CONCLUSION

If we are truly to promote it, we must recognize that many people's well-being is structured by personal needs. Personal need is a variant of a general concept of categorical need. On the dominant conception of categorical needs, the latter are necessarily minimal, universal, and moralized. By contrast, though likewise genuinely categorical, personal needs represent practical requirements entailed by central elements of a persons' well-being—their commitments. The theory of well-being based on this account of needs has a number of distinctive features. First, it is primarily specific to a personal evaluative context: that of an agent's decision-making in regard to all of the things that matter to them. Second, it is a theory of well-being's structure, not of which specific things form part of people's well-being or why they do so. Third, its avoidance of commitment on these scores enables it to identify important parts of well-being that are both objectively important and subject dependent in a relatively neutral way. Yet the theory's most practically significant features are the foundations it provides for the apparent non-substitutability of some people's commitments. This non-substitutability is twofold: such commitments are non-substitutable both (i) by non-needs and (ii) by each other. Both relations owe to the inescapability of commitments' requirements, that is, personal needs. Personal needs' significance lies precisely in their appearance in the personal, agential context, since it is there that the most formidable theories opposing non-substitutability are likewise initially motivated. Personal needs present a foundational challenge to any ethical, political, or economic theory that relies on intrapersonal aggregation to

96 Cf. Sen, *Development as Freedom*, 78–79; Wolff and de-Shalit, *Disadvantage*.

97 Alkire, "The Capability Approach and Well-Being Measurement for Public Policy," 618.

98 Alkire et al., *Multidimensional Poverty Measurement and Analysis*, 212.

be in principle unproblematic. This paper's practical conclusions have not been entirely negative, however. It also prepares the way for developing non-aggregating approaches to both intra- and interpersonal decision making involving needs. These are steps toward understanding the possibility of an ethical and social order that takes people's central interests seriously.<sup>99</sup>

*marlowe.fardell@gmail.com*

#### REFERENCES

- Ackrill, J.L. "Aristotle on *Eudaimonia*." In *Essays on Aristotle's Ethics*, edited by Amélie Oksenberg Rorty, 15–33. Berkeley: University of California Press, 1980.
- Alexandrova, Anna. *A Philosophy for the Science of Well-Being*. Oxford: Oxford University Press, 2017.
- Alkire, Sabina. "The Capability Approach and Well-Being Measurement for Public Policy." In *The Oxford Handbook of Well-Being and Public Policy*, edited by Matthew Adler and Marc Fleurbaey, 615–44. Oxford: Oxford University Press, 2016.
- Alkire, Sabina, James Foster, Suman Seth, Maria Emma Santos, José Manuel Roche, and Paola Ballon. *Multidimensional Poverty Measurement and Analysis*. Oxford: Oxford University Press, 2015.
- Allais, Maurice. "The Foundations of a Positive Theory of Choice involving Risk and a Criticism of the Postulates and Axioms of the American School." In *Expected Utility Hypotheses and the Allais Paradox: Contemporary Discussions of Decisions under Uncertainty with Allais' Rejoinder*, edited by Maurice Allais and Ole Hagen, 27–145. Dordrecht: Reidel, 1979.
- Anscombe, G. E. M. "Modern Moral Philosophy." *Philosophy* 33, no. 124 (January 1958): 1–19.
- Aristotle. *Nicomachean Ethics*. 2nd ed. Translated by Terrence Irwin. Indianapolis: Hackett, 1999.

99 The account developed in this paper has been significantly impacted by discussions over the years with Senthuran Bhuvanendra, Thomas Carr, Amanda Hughes, Douglas Lavin, Polly Mitchell, Véronique Munoz-Dardé, and James Wilson. It has also greatly benefited from the comments of two anonymous referees. I would also like to thank my doctoral examiners Michael Otsuka and Ingrid Robeyns for their comments on earlier incarnations of some of these ideas. Special thanks to Daniel Ranweiler for his generous comments on draft versions of this paper.

- Arneson, Richard J. "Human Flourishing versus Desire Satisfaction." *Social Philosophy and Policy* 16, no. 1 (Winter 1999): 113–42.
- Arrhenius, Gustaf. "Superiority in Value." *Philosophical Studies* 123, nos. 1–2 (March 2005): 97–114.
- Arrow, Kenneth. "Exposition of the Theory of Choice under Uncertainty." *Synthese* 16, nos. 3–4 (December 1966): 253–69.
- Bailey, James W. "Is It Rational to Maximize?" *Utilitas* 10, no. 2 (July 1998): 195–221.
- Berlin, Isaiah. *Liberty: Incorporating Four Essays on Liberty*. Oxford: Oxford University Press, 2002.
- Berlin, Isaiah, and Bernard Williams. "Pluralism and Liberalism: A Reply." *Political Studies* 42, no. 2 (June 1994): 306–9.
- Bermúdez, José. *Decision Theory and Rationality*. Oxford: Oxford University Press, 2009.
- Brandt, Richard. *A Theory of the Good and the Right*. Oxford: Clarendon Press, 1979.
- Braybrooke, David. *Meeting Needs*. Princeton: Princeton University Press, 1987.
- Brink, David O. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press, 1989.
- Brock, Gillian. "Morally Important Needs." *Philosophia* 26, nos. 1–2 (March 1998): 165–78.
- Brock, Gillian, and Soran Reader. "Needs-Centered Ethical Theory." *Journal of Value Inquiry* 36, no. 4 (December 2002): 425–34.
- Broome, John. "Utility." *Economics and Philosophy* 7, no. 1 (1991): 1–12.
- . *Weighing Goods: Equality, Uncertainty, and Time*. Oxford: Basil Blackwell, 1991.
- Campbell, Stephen M. "The Concept of Well-Being." In Fletcher, *The Routledge Handbook of Philosophy of Well-Being*, 402–13.
- Cartwright, Nancy, and Rosa Runhardt. "Measurement." In *Philosophy of Social Science: A New Introduction*, edited by Nancy Cartwright and Eleonora Montuschi, 265–87. Oxford: Oxford University Press, 2014.
- Chang, Ruth. "Introduction." In *Incommensurability, Incomparability, and Practical Reason*, edited by Ruth Chang, 1–34. Cambridge, MA: Harvard University Press, 1997.
- Crisp, Roger. *Reasons and the Good*. Oxford: Oxford University Press, 2006.
- Dewey, John. "Human Nature and Conduct." 1922. In *The Middle Works, 1899–1924*, vol. 14, edited by Jo Ann Boydston, 1–12. Carbondale, IL: Southern Illinois University Press, 1967.

- Dorsey, Dale. "Headaches, Lives, and Value." *Utilitas* 21, no. 1 (March 2009): 36–58.
- Doyal, Len, and Ian Gough. *A Theory of Human Need*. Basingstoke: Macmillan, 1991.
- Fardell, Benjamin. "Conceptualising Capabilities and Dimensions of Advantage as Needs." *Journal of Human Development and Capabilities* 21, no. 3 (2020): 263–76.
- Feldman, Fred. "Two Visions of Welfare." *Journal of Ethics* 23, no. 2 (June 2019): 99–118.
- Finnis, John. *Natural Law and Natural Rights*. Oxford: Oxford University Press, 1980.
- Fletcher, Guy. "A Fresh Start for the Objective-List Theory of Well-Being." *Utilitas* 25, no. 2 (June 2013): 206–20.
- . "Needing and Necessity." In *Oxford Studies in Normative Ethics*, vol. 8, edited by Mark Timmons, 170–92. Oxford: Oxford University Press, 2018.
- , ed. *The Routledge Handbook of Philosophy of Well-Being*. Abingdon, UK: Routledge, 2015.
- Flew, Antony. "Wants or Needs, Choices or Commands." In *Human Needs and Politics*, edited by Ross Fitzgerald, 213–38. Rushcutters Bay, Australia: Pergamon Press, 1977.
- Frankfurt, Harry. "The Importance of What We Care About." In *The Importance of What We Care About*, 80–94.
- . *The Importance of What We Care About: Philosophical Essays*. Cambridge: Cambridge University Press, 1988.
- . "Necessity and Desire." In *The Importance of What We Care About*, 104–16.
- . "Rationality and the Unthinkable." In *The Importance of What We Care About*, 177–90.
- Fried, Barbara H. *Facing Up to Scarcity: The Logic and Limits of Nonconsequentialist Thought*. Oxford: Oxford University Press, 2020.
- Gasper, Des. "Human Well-Being: Concepts and Conceptualizations." In McGillivray, *Human Well-Being*, 23–64.
- . "Understanding the Diversity of Conceptions of Well-Being and Quality of Life." *Journal of Socio-Economics* 39, no. 3 (June 2010): 351–60.
- Goodin, Robert E. "The Priority of Needs." *Philosophy and Phenomenological Research* 45, no. 4 (June 1985): 615–25.
- Griffin, James. *Value Judgement: Improving our Ethical Beliefs*. Oxford: Oxford University Press, 1996.

- . *Well-Being: Its Meaning, Measurement, and Moral Importance*. Oxford: Clarendon Press, 1986.
- Grix, Marco, and Philip McKibbin. “Needs and Well-Being.” In Fletcher, *The Routledge Handbook of Philosophy of Well-Being*, 292–306.
- Hooker, Brad. “Fairness, Needs, and Desert.” In *The Legacy of H. L. A. Hart: Legal, Political, and Moral Philosophy*, edited by Matthew Kramer, Claire Grant, Ben Colburn, and Antony Hatzistavrou, 181–99. Oxford: Oxford University Press, 2008.
- Horton, Joe. “Aggregation, Risk, and Reductio.” *Ethics* 130, no. 4 (July 2020): 514–29.
- Hurley, Susan. *Natural Reasons: Personality and Polity*. Oxford: Oxford University Press, 1989.
- Kagan, Shelly. “Me and My Life.” *Proceedings of the Aristotelian Society* 94, no. 1 (June 1994): 309–24.
- . “Rethinking Intrinsic Value.” *Journal of Ethics* 2, no. 4 (1998): 277–97.
- Kekes, John. *The Morality of Pluralism*. Princeton: Princeton University Press, 1993.
- Kolnai, Aurel. “Deliberation Is of Ends.” *Proceedings of the Aristotelian Society* 62, no. 1 (June 1961): 195–218.
- Korsgaard, Christine M. “Two Distinctions in Goodness.” *Philosophical Review* 92, no. 2 (1983): 169–95.
- Kropotkin, Peter. *The Conquest of Bread*. 1892. Rookhope, UK: Aziloth Books, 2017.
- Mason, Elinor. “Value Pluralism.” In *Stanford Encyclopedia of Philosophy* (Spring 2018). <https://plato.stanford.edu/archives/spr2018/entries/value-pluralism>.
- Max-Neef, Manfred, Antonio Elizalde, and Martín Hopenhayn. “Human Scale Development: An Option for the Future.” *Development Dialogue* 1 (1989): 7–80.
- McGillivray, Mark, ed. *Human Well-Being: Concept and Measurement*. Houndmills, UK: Palgrave Macmillan and UNU-WIDER, 2007.
- Mill, John Stuart. *Utilitarianism*. London: Parker, Son, and Bourn, 1863.
- Miller, David. *Principles of Social Justice*. Cambridge, MA: Harvard University Press, 1999.
- . *Social Justice*. Oxford: Clarendon Press, 1979.
- Munoz-Dardé, Véronique. “In the Face of Austerity: The Puzzle of Museums and Universities.” *Journal of Political Philosophy* 21, no. 2 (June 2013): 221–42.
- Nagel, Thomas. *The View from Nowhere*. Oxford: Oxford University Press, 1986.
- Norcross, Alastair. “Comparing Harms: Headaches and Human Lives.” *Philosophy and Public Affairs* 26, no. 2 (April 1997): 135–67.

- Nussbaum, Martha C. *Creating Capabilities: The Human Development Approach*. Cambridge, MA: Harvard University Press, 2011.
- O'Neill, John. "The Varieties of Intrinsic Value." *Monist* 75, no. 2 (April 1992): 119–37.
- Rabinowicz, Wlodek, and Toni Rønnow-Rasmussen. "A Distinction in Value: Intrinsic and for Its Own Sake." *Proceedings of the Aristotelian Society* 100, no. 1 (June 2000): 33–51.
- Railton, Peter. "Facts and Values." *Philosophical Topics* 14, no. 2 (Fall 1986): 5–31.
- . "Moral Realism." *Philosophical Review* 95, no. 2 (April 1986): 163–207.
- Rawls, John. "Justice as Fairness: Political Not Metaphysical." *Philosophy and Public Affairs* 14, no. 3 (Summer 1985): 223–51.
- . "A Kantian Conception of Equality." *Cambridge Review* 96 (1975): 94–99.
- . "Social Unity and Primary Goods." In *Utilitarianism and Beyond*, edited by Amartya Sen and Bernard Williams, 159–85. Cambridge: Cambridge University Press, 1982.
- . *A Theory of Justice*. Rev. ed. Oxford: Oxford University Press, 1999.
- Raz, Joseph. *The Morality of Freedom*. Oxford: Oxford University Press, 1986.
- Reader, Soran. *Needs and Moral Necessity*. Abingdon, UK: Routledge, 2007.
- Reader, Soran, and Gillian Brock. "Needs, Moral Demands and Moral Theory." *Utilitas* 16, no. 3 (November 2004): 251–66.
- Richardson, Henry S. *Practical Reasoning about Final Ends*. Cambridge: Cambridge University Press, 1994.
- . "Specifying Norms as a Way to Resolve Concrete Ethical Problems." *Philosophy and Public Affairs* 19, no. 4 (Autumn 1990): 279–310.
- Robeyns, Ingrid. *Wellbeing, Freedom and Social Justice: The Capability Approach Re-Examined*. Cambridge: Open Book Publishers, 2017.
- Rosati, Connie. "Internalism and the Good for a Person." *Ethics* 106, no. 2 (January 1996): 297–326.
- Savage, Leonard J. *The Foundations of Statistics*. New York: Wiley and Sons, 1954.
- Scanlon, T.M. "Preference and Urgency." *Journal of Philosophy* 72, no. 19 (November 1975): 655–69.
- . *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1998.
- Schönherr, Julius. "Still Lives for Headaches: A Reply to Dorsey and Voorhoeve." *Utilitas* 30, no. 2 (June 2018): 209–18.
- Sen, Amartya. *Development as Freedom*. Oxford: Oxford University Press, 1999.
- . *Inequality Reexamined*. Oxford: Oxford University Press, 1995.



- Smith, Adam. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Oxford: Clarendon Press, 1976.
- Stewart, Frances. "Basic Needs Approach." In *The Elgar Companion to Development Studies*, edited by David Clark, 14–18. Cheltenham UK: Edward Elgar, 2006.
- . *Basic Needs in Developing Countries*. Baltimore: Johns Hopkins University Press, 1985.
- Stiglitz, Joseph, Amartya Sen, and Jean-Paul Fitoussi. *Report by the Commission on the Measurement of Economic Performance and Social Progress*. Paris: Institut National de la Statistique et des Études Économiques, 2009.
- Stocker, Michael. *Plural and Conflicting Values*. Oxford: Oxford University Press, 1992.
- Thomson, Garrett. *Needs*. London: Routledge and Kegan Paul, 1987.
- Veenhoven, Ruut. "Subjective Measures of Well-Being." In McGillivray, *Human Well-Being*, 214–39.
- Voorhoeve, Alex. "How Should We Aggregate Competing Claims?" *Ethics* 125, no. 1 (2014): 64–87.
- Watson, Gary. "Volitional Necessities." In *Agency and Answerability: Selected Essays*, 88–122. Oxford: Oxford University Press, 2004.
- Wiggins, David. "Claims of Need." In *Needs, Values, Truth*, 1–57.
- . "Deliberation and Practical Reason." In *Needs, Values, Truth*, 215–37.
- . "An Idea We Cannot Do Without: What Difference Will It Make (e.g. to Moral, Political, and Environmental Philosophy) to Recognize and Put to Use a Substantial Conception of Need?" *Philosophy of Need, Royal Institute of Philosophy Supplement* 57 (2005): 25–50.
- . "Incommensurability: Four Proposals." In *Needs, Values, Truth*, 356–80.
- . *Needs, Values, Truth: Essays in the Philosophy of Value*. 3rd ed. Oxford: Oxford University Press, 1998.
- Williams, Bernard. "Conflicts in Values." In *Moral Luck*, 71–82. Cambridge: Cambridge University Press, 1981.
- . "A Critique of Utilitarianism." In *Utilitarianism: For and Against*, edited by J. J. C. Smart and Bernard Williams, 77–150. Cambridge: Cambridge University Press, 1973.
- . *Ethics and the Limits of Philosophy*. London: Fontana/Collins, 1985.
- Wolff, Jonathan, and Avner de-Shalit. *Disadvantage*. Oxford: Oxford University Press, 2007.
- Woodard, Christopher. "Classifying Theories of Welfare." *Philosophical Studies* 165, no. 3 (September 2013): 787–803.

## AN ACCOUNT OF NORMATIVE STEREOTYPING

Corey Barnes

Definitions belong to the definers, not the defined.

—Toni Morrison, *Beloved*

ADRIAN PIPER provides an innovative way to think about stereotyping and how it leads to discrimination.<sup>1</sup> Piper elucidates two kinds of discrimination—namely, first-order and higher-order political discrimination. Both rely on stereotypes that are motivated by xenophobia. The relationship between stereotyping and discrimination can be captured by what I refer to as discrimination from descriptive stereotyping. Here, stereotypical properties are taken to be possessed by and principally define individuals because of groups to which they belong. These properties are descriptive in nature; they describe what group members must be like. Discrimination results from and is thought to be justified by the perception that group members must possess certain negatively valued attributes because they belong to groups.

In this article I discuss a relationship between stereotyping and discrimination that has been rather overlooked, particularly in philosophical literature. I refer to the relationship as “discrimination from normative stereotyping.” On my account, stereotypes provide criteria for what legitimate members of some group are like, and thus which attributes group members ought to possess. Discrimination results from a failure of group members to possess these stereotypical attributes. Negative evaluations that lead to discrimination are not made insofar as group members are thought to possess disvalued attributes, but are made insofar as group members lack certain valuable, group-related attributes. Herein I take discrimination from normative stereotyping to explain the use of particular slurs—namely, race-traitor terms such as “Uncle Tom” and “nigger lover.” Targets of these slurs are discriminated against because they are perceived as failing to be legitimate group members insofar as they lack certain properties stereotypically imposed on their group.

In fleshing out an account of normative stereotyping that leads to discrimination, I begin by briefly reconstructing Piper’s view. I highlight her two kinds of

1 Piper, “Higher-Order Discrimination” and “Two Kinds of Discrimination.”

political discrimination (first-order and higher-order discrimination) and show how they derive from descriptive stereotyping. In a subsequent section I discuss normative stereotyping that leads to discrimination, and illustrate first-order and higher-order kinds of discrimination from normative stereotyping. In so doing, I show how first-order discrimination from normative stereotyping captures the use of “race traitor” and a few race-traitor terms.<sup>2</sup>

#### 1. “IS IT BECAUSE I’M BLACK?”—ON DESCRIPTIVE STEREOTYPING

Piper discusses discrimination in two senses—namely, cognitive and political. Cognitive discrimination is the capacity to distinguish and respond to different properties or objects in an appropriate way. Political discrimination is failed cognitive discrimination that is enacted upon. It is a manifest “attitude in which a particular property of a person which is irrelevant to judgments of that person’s intrinsic value or competence . . . is seen as a source of disvalue or incompetence; in general, as a source of inferiority.”<sup>3</sup>

Piper takes political discrimination to be motivated by xenophobia, which is a specific kind of pseudorationality. Pseudorationality is an attempt to make sense of anomalous data in a way that preserves one’s rational subjectivity when it is under duress. Xenophobia is an attempt to preserve one’s rational subjectivity in the face of perceptions that threaten it because of the xenophobe’s narrow conception of personhood.<sup>4</sup> It is a fear of “certain kinds of strangers, namely those who do not conform to one’s preconceptions about how persons ought to look or behave.”<sup>5</sup> Insofar as these strangers possess attributes that threaten the internal coherence of the xenophobe’s conception of personhood, they threaten the internal coherence of the xenophobe’s rational subjectivity. Instead of revising or jettisoning their inadequate conception, xenophobes rationalize or dissociate personhood away from the “others,” or simply deny the “others” personhood altogether.<sup>6</sup> Committed xenophobes—those who have a personal investment

2 I attempt to capture “thicker” or “harder” race-traitor terms. I work out a theory of “thinner” or “softer” race-traitor terms such as “oreo,” “wigger,” and “banana” in a forthcoming work.

3 Piper, “Two Kinds of Discrimination,” 193.

4 Piper understands “personhood” to be a property attributed to beings who are presumed as having consciousness, thought, rationality, and agency. She does not intend to call up issues of reidentification or characterization that accord to traditional questions of personal identity.

5 Piper, “Two Kinds of Discrimination,” 198.

6 By “rationalization” Piper means maximizing certain properties that confirm a diminished conception of personhood, while minimizing others that support personhood. By “dissociation” Piper means identifying a being without the necessary properties needed to possess

in perceiving only some humans as possessing personhood—explicitly impose stereotypes on others in a way that denies them personhood.

Now to “impose a stereotype on someone is to view her as embodying a limited set of properties falsely taken to be exclusive, definitive, and paradigmatic of a certain kind of individual.”<sup>7</sup> One fails to appreciate the complexity of group members, and instead takes certain properties to be present in and principally define them without further investigation. Stereotypes tell us what members must be like, and why some members cannot possess personhood. One need not get to know individuals; one already has knowledge that neither requires confirmation by engagement nor is ever incorrect.

The application of stereotypes to groups leads to political discrimination. For Piper there are two kinds of political discrimination. First-order political discrimination regards one’s unashamed commitment to a very narrow conception of personhood that serves some interest. The committed xenophobe unashamedly takes attributes of some person that are irrelevant to judgments of that person’s noninstrumental value or competence to be the source of inferiority, and acts on that basis. Higher-order political discrimination is an implicit commitment to a very narrow conception of personhood that, if discovered, would evoke shame in the discriminator. It is “the attitude within which a primary disvalued or valued property in turn confers disvalue or value respectively on further properties of the disvaluee or valuee respectively.”<sup>8</sup> Certain secondary attributes—say eloquence in speech—that may be valued in someone with a primary valued attribute—say whiteness—become disvalued because they are attributes of someone with a primary disvalued attribute—say Blackness. So, speaking eloquently is valuable, and is taken to be so when the person possessing the attribute is white. However, in someone who is Black, “eloquence” becomes “flowery” or “highfalutin.”

I imagine that assigning disvalue or value to higher-order attributes can regard interests and commitments. For example, if some professor wants to increase representation of some marginalized group, say Black women, and the professor has the primary valued attribute of whiteness (and particularly if male), then he is perceived as “passionate for justice” by particular persons. But when some professor wants to increase the representation of the same marginalized group, and the professor has the primary disvalued attribute of Blackness (and particularly if female), then she is perceived as “obsessed with injustice” by the

---

personhood. And by “denial” Piper has in mind suppressing recognition of the anomalous being altogether.

7 Piper, “Two Kinds of Discrimination,” 207.

8 Piper, “Two Kinds of Discrimination,” 215.

same persons.<sup>9</sup> The former is perceived as desiring to press forward to a new and progressive future, while the latter is perceived as refusing to let go or move beyond the past. Though disguised, the evaluations derive from the primary valued or disvalued attributes. In these cases the higher-order political discriminator's self-concept relies on not being a xenophobe, and thus they cannot consciously accept that the agent's primary attributes "Blackness" and "femaleness" are conferring the disvalue. Their xenophobic actions must be suppressed or denied.

Piper's first-order and (perhaps surprisingly) higher-order political discriminations involve stereotyping that follows a certain form, call it "discrimination from descriptive stereotyping." Discrimination from descriptive stereotyping accords to the following:

1. Members of group *X* have trait/attribute *Y* (which is negatively valued).
2. Person *A* is a member of group *X*.
3. Therefore, person *A* has (or rather must have) trait/attribute *Y* (which is negatively valued).
4. Discrimination results from discriminator *D*'s perception that *A* is a member of *X*, and therefore must possess the negatively valued trait/attribute by virtue of being a member of *X*.

That Piper's first-order and higher-order political discriminations involve descriptive stereotyping is implicated by her claim regarding what it means to impose a stereotype. Further, consider what stereotypes do. They equate one very limited set of properties with personhood, and exclude any other property that does not fit within that set. In so doing they exclude individuals who are not perceived as possessing all and only those properties of the limited set, attributing to them certain properties that clearly demarcate the deviance. This creates what she calls "honorific" and "derogatory" properties.<sup>10</sup> Whether honorific or derogatory, these stereotypical properties are taken to be possessed by those on whom they are imposed. Thus for Piper, stereotypes take descriptive properties to—of necessity—exist in individuals because of groups to which they belong. Discrimination results from the perception that members of groups must possess—in fact are defined by—derogatory properties.

Let us say that "whiteness" is perceived as a primary valued attribute included (among others) in the limited set of properties that would confer personhood to someone. "Blackness" is excluded. Honorific stereotypes (perhaps "intelligence," among others) are created for those who are white. Derogatory stereo-

9 Carbado and Gulati illustrate this by appeal to the perception of Black and white male law professors' pedagogical approaches. See Carbado and Gulati, *Acting White?* 36.

10 Piper, "Two Kinds of Discrimination," 207–8.

types (perhaps, “unintelligence,” among others) are created for Blacks to mark off their deviance from whiteness. These stereotypes describe what individuals must be like by virtue of being Black. Discrimination against those with derogatory stereotypes is perceived as justified on that basis.

So *M* is Black and is applying for employment as a teacher at some institution. Piper’s understanding of stereotyping and the political discrimination to which stereotyping leads can be captured as follows:

1. “Blacks” are unintelligent.
2. *M* is “Black.”
3. Therefore, *M* is (and must be) unintelligent by virtue of being Black.
4. Political discrimination (regarding *M*’s potential employment, say as a teacher) results from and is justified by *M*’s unintelligence that is owed to his being Black.

These of course follow from the first-order discriminator (xenophobe) fixing the concept “personhood” by a very limited set of properties that Blacks do not possess.

On its face, higher-order discrimination should not follow this form. Higher-order political discriminators reject the imposition of stereotypes. Further, higher-order political discriminators would be appalled and indignant upon discovery of such stereotyping, either in themselves or as practiced by others. They would consciously deny the truth of 1 and 3, and reject 4 on moral grounds. And yet, there is just something about the way in which *M* performs some action that is evidence of him being unfit in some way. Possession of an otherwise valuable property is perceived negatively, and justifies the resulting discriminatory act.

So “eloquence in speech” is seen as a valuable property in white teachers. However, the higher-order political discriminator may perceive it as “flowery” or “pretentious,” but this is explained by the suppressed perception that Blacks are unintelligent, and by *M*’s being Black. And so higher-order political discriminators appeal to the same form as first-order discriminators. The only difference is that the higher-order discriminator denies or suppresses their xenophobic reactions by attributing disvalue to higher-order properties—properties that preserve the discriminator’s self-concept.

I take Piper’s discussion of political discrimination to derive from “*is-grounded*” stereotyping. I would like to draw attention to a particular type of “*ought-grounded*” stereotyping that leads to discrimination. So instead of discrimination that is based on stereotypical properties that group members are perceived as having, discrimination follows from certain stereotypical properties that members ought to have but either do not have or do not display outwardly well enough or to the extent that they “should.”

## 2. "AM I BLACK ENOUGH FOR YOU?"—ON NORMATIVE STEREOTYPING

I take Piper's claims about both stereotyping and political discrimination to be accurate, but incomplete. Further, in terms of political discrimination, Piper's account only captures descriptive stereotyping. An irrelevant property of some person that is taken to indicate inferiority implies possession of a disvalued property. And so treatment of some person based on stereotypical properties that one ought to possess by virtue of group membership—where stereotypes associated with that group are valued but that a group member lacks—can never be political discrimination. So as to capture discrimination from both descriptive and normative stereotyping, I understand political discrimination as *a manifest attitude in which any property (either perceived to be possessed or perceived ought to be possessed) that is irrelevant to judgments of some person's intrinsic value or competence informs an action that either benefits or harms that person.*

Now political discriminators may take stereotypical properties to—of necessity—exist in and principally define some group, and in this way expect members to possess them. For these discriminators, group members are preferable and worthy of fair treatment to the extent that they perform actions in accordance with stereotypes; possession of the stereotypical properties would confer a kind of value to members. However, members' failure to be instantiations of the stereotypes frustrates discriminators' expectations or desires, and ultimately leads to discriminatory behavior. In these cases, neither primary attributes such as race or gender nor stereotypical properties such as unintelligence make members disvaluees, thereby indicating members' inferiority. And so political discrimination does not derive from descriptive properties. In these cases I understand the role that stereotypes play in discrimination somewhat differently than Piper. These are captured by a kind of *ought*-grounded stereotyping that I term "discrimination from normative stereotyping."

Though rather overlooked, and particularly so in philosophical literature, *ought*-grounded stereotyping has received some attention. Kwame Appiah seems to have coined the phrase "normative stereotyping."<sup>11</sup> For Appiah, normative stereotyping "is grounded in a social consensus about how they [members of a group] *ought* to behave in order to conform appropriately to the norms associated with membership in their group."<sup>12</sup> These stereotypes are necessary scripts for social identities.<sup>13</sup> As both important and necessary scripts, normative stereotypes are neither necessarily inaccurate nor morally problematic.

11 Appiah, "Stereotypes and the Shaping of Identity."

12 Appiah, "Stereotypes and the Shaping of Identity," 48

13 Appiah, "Stereotypes and the Shaping of Identity," 51.



However, these scripts “have to be configured in such a way as to serve as potential instruments in the construction of a dignified individuality.”<sup>14</sup> Normative stereotypes become morally problematic when they lead to inferiorizing those who bear them. When this occurs, reshaping them becomes necessary.

Feminist theorists and social psychologists have also discussed *ought*-grounded stereotyping. Feminists have primarily focused on the relationship between *ought*-grounded stereotypes and the oppression of women. These *ought*-grounded stereotypes lead women to and keep them in marriages, sexual relationships, and housewifery roles that they often despise. Further, *ought*-grounded stereotypes related to certain women are thought to have derived from a particular image of women created by men.<sup>15</sup> This image, thought to derive from the nature of women, caused women with careers to be negatively evaluated by portraying them as denying their natural femininity in a way that made them unhappy or moribund. And in this way *ought*-grounded stereotypes were tied to ethics, where nonconformity made women vicious. Still further, *ought*-grounded stereotypes have been taken to explain psychological barriers to physical tasks and intellectual pursuits, govern the ways women move and groom themselves, and the prevalence of anorexia and bulimia among women.<sup>16</sup>

Contemporary social psychologists—particularly those working on gender and workplace evaluations—discuss a particular type of *ought*-grounded stereotyping referred to as “prescriptive stereotyping.” Madeline Heilman proposes that these stereotypes “designate how women and men are but also how they should be. They function as injunctive norms, dictating what attributes and behaviors are appropriate and inappropriate for people from different groups—in this case men and women.”<sup>17</sup> Failure to act in accordance with these stereotypes often leads to negative evaluations for women and men.<sup>18</sup> Beyond the idea that there are prescriptive stereotypes for women and men, it has been theorized that

14 Appiah, “Stereotypes and the Shaping of Identity,” 49.

15 Friedan, *The Feminine Mystique*, 82–83.

16 Young, “Throwing Like a Girl”; Beauvoir, *The Second Sex*; Friedan, *The Feminine Mystique*; Bartky, “Foucault, Femininity, and the Modernization of Patriarchal Power”; Bordo, *Unbearable Weight*.

17 Heilman, “Gender Stereotypes and Workplace Bias,” 123.

18 On negative evaluations of women, see Heilman, “Gender Stereotypes and Workplace Bias”; Haddock and Zanna, “Preferring ‘Housewives’ to ‘Feminists’”; Rudman and Fairchild, “Reactions to Counterstereotypic Behavior”; Carbado and Gulati, *Acting White?* On negative evaluations of men, see Rosette, Mueller, and Lebel, “Are Male Leaders Penalized for Seeking Help?”; Moss-Racusin, Phelan, and Rudman, “When Men Break the Gender Rules”; Berdahl, “Harassment Based on Sex.”

conforming to these stereotypes influences job/career choices, education types/commitments, and even spending habits/values.<sup>19</sup>

I propose a more nuanced account of *ought*-grounded stereotyping than the above subclasses, one where the relationship between stereotyping and discrimination accords to the following form:

1. Members of group  $X$  ought to  $\phi$  or possess property  $P$  if they are legitimate members of group  $X$ .
2. Person  $A$  is a member of group  $X$ .
3. Therefore, person  $A$  ought to  $\phi$  or possess property  $P$ .
4. Discriminator  $D$  observes that person  $A$  either does not  $\phi$  or does not possess property  $P$ .
5. Discriminator  $D$  judges that person  $A$  is not a legitimate member of group  $X$ , and that they are contemptible because of it.
6. Political discrimination is the result of discriminator  $D$ 's judgment that person  $A$  fails to be a legitimate member of group  $X$  and is contemptible because of it.

In 1, "legitimate members of group  $X$ " means any particular's possession of all those properties deemed proper to it, and possession of them in a way that satisfies expectations or desires of some agent making judgments about the particular. If I purchase a watch, then I expect or desire it to perform certain tasks in certain ways given attributes that it ought to possess. My judgment that the watch fails in this way both disappoints and frustrates me. I then judge the watch to be illegitimate and of lesser, little, or no value depending on the distance between my expectations or desires and its performance.

In the above schema, however, we are discussing the legitimacy of persons. And so 5 adds a step to this more ordinary sense of judgment at some particular's illegitimacy. Failure to possess stereotypical properties in a way that satisfies "judges" licenses contempt, "which is directed toward a person that the contemnor sees as failing to meet an important standard."<sup>20</sup> In the above, con-

19 Akerlof and Kranton, "Economics and Identity" and *Identity Economics*; Fordham and Ogbu, "Black Students' School Success"; Ogbu, *Black American Students in an Affluent Suburb*. I question how much of the prescriptive stereotypes regarding women and men actually apply to their respective sets, particularly as it relates to *Black* women and men. Reading works by authors such as Sojourner Truth (*Narrative of Sojourner Truth*) and Patricia Hill Collins (*Black Feminist Thought*) regarding Black women, and W.E.B. Du Bois (*The Philadelphia Negro*) and Malcolm X ("Not Just an American Problem, but a World Problem") regarding Black men is enough to raise questions about whether many of these studies capture correct prescriptive stereotypes that apply to *all* women and men.

20 Bell, *Hard Feelings*, 33. Bell acknowledges that contempt for nonpersons is possible and may

tempt is the result of a particular kind of disappointment in and frustration with a person's failure to meet the standard of legitimacy by possession of normative stereotypes. Unlike an emotion such as resentment that is directed at targets for a wrong action, contempt is directed at targets for a state of being.<sup>21</sup> Some person is contemptible because they fail to "be" as they should, given the group to which they belong. The discriminator views the target as a failure, and as a result sees themself as superior to the target.

Finally, because we are discussing stereotyping persons in a way that leads to political discrimination, 6 adds a step to both the more ordinary judgments of some particular's illegitimacy and the contempt directed at some person's being a failure. For the discriminator, a feeling of superiority gives them a right to exercise (unjust) power over the inferior target in a way that harms the target. So, discriminator *D* judges that person *A* is not a legitimate member of group *X*; the person fails to possess all those attributes deemed proper to them in a way that satisfies discriminator *D*'s expectations or desires. Person *A*'s illegitimacy both disappoints and frustrates discriminator *D*. The disappointment and frustration motivates discriminator *D*'s judgment that person *A* is a failure. Discriminator *D* judges that person *A* is contemptible due to being a failure, and sees themself as superior to person *A*. Discriminator *D* perceives that this feeling of superiority gives them a right to exercise power over person *A* in a way that harms person *A*. The exercising of power over person *A* in a way that harms is informed by a property that is irrelevant to judgments of person *A*'s intrinsic value or competence; it is political discrimination.

My view of normative stereotyping is different from Appiah's, feminists', and social psychologists' contributions in important respects. First, I do not take a social consensus to be necessary for grounding normative stereotypes. Though they often take cues from society, and though many are widely held, normative stereotypes need not be connected to widely held or agreed-upon social identities. They can be constructed and reconstructed by individuals in ways that reflect individuals' unique psychological dispositions. An individual or group can construct stereotypes where "white," "blonde," "tall," "southern," "men" are legitimate members only if they are "conservative," "heterosexual," "gun-rights advocates," discriminating against those who possess the former attributes but

challenge her account, though it can be captured by modifications. She concedes this given our everyday usage of contempt for nonpersons. I am not convinced that our everyday usage presents such a challenge. We may speak as though we condemn objects; however, I take some agent's superiority to what pretends to be an equal to be a significant motivator for contempt (see Roberts, *Emotions*, 256). However, "pretending" requires intentionality and deception, both of which are lacking in nonpersons.

21 Bell, *Hard Feelings*, 39.

who fail to possess the latter.<sup>22</sup> This proposal explains shifting and seemingly inconsistent stereotypes that change in accordance with region, location, and circumstance. This proposal can also explain subgroup stereotypes (for example Black women or Black women of a certain skin or body type). Thus it can explain particular cases of intra-racial and intra-gender discrimination that are rather difficult to capture.

Second, much of the literature from feminists and social psychologists takes *ought*-grounded stereotypes to derive from nature. I want to be clear that discriminators need not take normative stereotypes to be natural. Though normative stereotyping takes some group to “be like *X*,” which motivates the thought that legitimate group members ought to “*X*,” “*X*” need not be thought to be natural to the group.

Third, much of the discriminatory behavior in the literature above seems to derive in part from descriptive stereotyping. In much of the literature on women and workplace discrimination, women who violate prescriptive stereotypes for the purpose of succeeding in what is considered “men’s work” are negatively evaluated and suffer penalties. However, disvalue seems to be placed on the stereotypes prescribed for women because of disvalue placed on women’s descriptive properties. Here, women are thought to have descriptive properties distinct from and inferior to men’s. Prescriptive stereotypes are then imposed on women such that they are viewed negatively when they do not conform to them. However, conforming to these stereotypes would not confer value to women. And so women are negatively evaluated whether they do or do not conform to stereotypes.<sup>23</sup> Men, on the other hand, should not conform to stereotypes prescribed for women because men should not “act like inferior beings,” and conforming to these stereotypes is tantamount to “acting like inferior beings.” Take the action of running. Men ought not “run like a girl,” because “running like a girl” is thought to be “running badly,” and men ought not “be or act as inferior.”<sup>24</sup> Even

22 The stereotypical image need not be “deep” or involve “serious” properties. It can be as simple as being “able to sing” or being “knowledgeable about a certain kind of art or sports history.”

23 Consider Immanuel Kant’s early view of women. Kant tells us that: “Laborious learning or painful grubbing, even if a woman could get very far with them, destroy the merits that are proper to her sex” (“Observation on the Feeling of the Beautiful and the Sublime,” 41). Men and women (by nature) have different virtues. Women *ought* to attempt to flourish only by conforming to women’s virtues. However, rather than esteem women’s virtues and women who conform to them, Kant later disparages women’s virtues (even though he claims that women and men have equivalent understandings—albeit oriented toward different qualities). Here: (1) men and women have different virtues; (2) women should act in accordance with their virtues; and (3) women’s virtues are of significantly less worth than men’s.

24 At the heart of “running like a girl,” I take there to be disparagement attributed to “being a girl.”

though “running like a girl” is indicative of “inferiority,” women ought to “run like a girl.” Women are negatively evaluated when they do not conform to this stereotype—particularly when they best men. However, when they do conform they are still disparaged. Why? It is indicative of “inferiority.” This seems to typify much of what is captured by “prescriptive stereotyping.” Here, discriminatory action derives in part from descriptive stereotyping because the prescriptions are grounded in disvalued descriptive properties imposed on women. Prescriptive stereotyping tends to derogate the (descriptive) content of the prescription. In this way, prescriptive stereotyping shares a property with both “*is*-grounded” and “*ought*-grounded” stereotyping. I want to be clear that in my account there is no disvalue—but rather genuine value—placed on normative stereotypes.

An example of normative stereotyping. Early during my two years at Union Theological Seminary I had ambitions of working on problems of theodicy and proofs for the existence of God that followed from an appreciation of and perhaps a bit of a fascination with the works of Leibniz, St. Anselm, and Descartes. I recall speaking with several of my then colleagues about my ambitions at a gathering during my first week at the seminary, as we were all sharing our academic interests. And I recall the response from two of my Black-male colleagues in attendance: “Who is this going to help?” The implicature was that this was not a worthwhile project for me. From the statement it was clear that because I was African American and an aspiring theologian I should have the academic aim or ambition of “helping others,” by which they meant solving a particular set of social or political problems (perhaps by proposing very specific solutions)—a fact that was later confirmed in conversation with them. However, while discussing our academic interests, this sort of criticism was not leveled against any white seminarian. My ambitions, I am sure, affected their perception of me as a potentially legitimate African American theologian. I have no doubt that—given the fervor with which this statement was uttered and confirmed—these men would have revoked my admission recommendation if they could.

Perhaps with the exception of withholding friendship, these potential political discriminators did not discriminate against me.<sup>25</sup> However, let us assume

---

It is interesting that statements like “You X like a girl/woman” are never uttered positively, even when it relates to what is considered “women’s work.” No one says “You cook like a woman” or “You sew like a woman.” Why? That would add value—in fact value greater than men’s—to what is characteristically “woman,” which undermines the disparagement attributed to womanhood.

25 Following Piper, I take the withholding of friendship under certain conditions to be both an immoral and politically discriminatory harm (“Higher-Order Discrimination”). And so I take political discrimination to regard at least two classes. First, there is political discrimination that is both illegal and immoral. Second, there is political discrimination that, while legal (and perhaps ought to be in a liberal democracy), is nonetheless immoral.

that they had certain power to affect admittance and hiring decisions, which I assume if they currently have certain positions in the academy. Let us also assume that this response to me is not so extraordinarily rare that it has only occurred in my experiences. I do not see how this political discrimination can be captured with discrimination from descriptive stereotyping. The primary attribute of being African American was not seen as a source of inferiority. These men did not commit to anything like the view that African Americans are not smart enough to work on theodicy or proofs. Further, I was not somehow inferior because theodicy or proofs of God's existence were deemed illegitimate work for theologians. The source of my inferiority was my academic interests, paired with my ethnic group, given certain expectations about how I ought to "be" as a member of the group. My actual academic interests, whatever they may have been, failed to satisfy their expectations and led them to feel disappointment and frustration. As a result, both colleagues felt contempt, which produced a feeling of moral superiority. And political discrimination would follow from this feeling. So rather than following discrimination from descriptive stereotyping, a better way to capture the potential discriminatory behavior of these men, and the actual political discrimination of persons like them, is as follows:

1. African American theologians ought to be concerned with social justice in their academic pursuits, if they are in fact legitimate African American theologians.
2. Corey Barnes is an African American theologian.
3. Therefore, Corey Barnes ought to be concerned with social justice in his academic pursuits.
4. (However) discriminators *D* and *F* both observe that Corey Barnes is not concerned with social justice in his academic pursuits (at least not to the extent that he should be, as evidenced by his particular statements).
5. Discriminators *D* and *F* judge that Corey Barnes is not a legitimate African American theologian, and that he is contemptible because of it.
6. Political discrimination is (or would have been if power was so accorded) the result of discriminators *D*'s and *F*'s judgment that Corey Barnes fails to be concerned with social justice in his academic pursuits (and thus fails to be a legitimate African American theologian), and that he is contemptible because of it.

Now failure to meet political discriminators' expectations of what African American theologians ought to be might license political discriminators to evaluate African American theologians' work of a lesser quality or value to theology and discriminate against them in hiring practices. Or, discriminators may consider

these theologians less worthy of friendship, collegiality, or a certain type of treatment as a colleague. In each case, a property that is irrelevant to judgments of the African American theologian's intrinsic value or competence would inform an action that harms them.

Perhaps contrary to intuitions, discrimination from normative stereotyping captures discrimination against individuals who are referenced by certain slurs. I take these to be examples of first-order political discrimination from normative stereotyping. Consider the term "race traitor" and the family of slurs that fall under it. The Racial Slurs Database defines "race traitor" as "a term used by whites for a white who marries a nonwhite."<sup>26</sup> The "race traitor," as defined in this way, has an attraction to the wrong kind of person; they fail to perform some action that they ought to perform—namely, marrying a white person. Someone who targets individuals with this term thinks that whites ought only to be attracted to and marry other whites if they are legitimate white persons. The term picks out a failure of this behavior to be manifest, and further signals some greater failure to "be" as the person should. Use of "traitor" seems to signal an emotion stronger than contempt—namely, (misguided) moral hatred.<sup>27</sup> The "race traitor" is taken to be a moral offense who must be overcome. And so use of "traitor" signals that discriminatory behavior against the target is not merely justified but morally required. One might think of Michal Szewczuk's statement that Prince Harry is a "race traitor," and that he should be assassinated for marrying Meghan Markle.<sup>28</sup> So:

1. Whites ought to be sexually attracted to and marry only whites, if they are in fact legitimate white persons.
2. Prince Harry is a white person.
3. Therefore, Prince Harry ought to be sexually attracted to and marry only a white.
4. Michal Szewczuk observed that Prince Harry is not sexually attracted to and has not married only a white.
5. Michal Szewczuk judged that Prince Harry is not a legitimate white person, but rather a "race traitor," and hates him because of it.
6. Political discrimination (in the form of assassination) was then recommended after Michal Szewczuk's judgment that Prince Harry fails to be sexually attracted to and marry only a white.

26 Racial Slur Database, "Races," <http://www.rsd.org/races>.

27 See Hampton, who distinguishes four kinds of hatred—namely, simple, malicious, spiteful, and moral ("Forgiveness, Resentment and Hatred").

28 Melendez, "British Neo-Nazi Who Threatened 'Race Traitor' Prince Harry Sentenced to Four Years in Prison."



“Race traitor” as defined above is inadequate. There are terms used to indicate Black (and other racial and ethnic group) “race traitors.” A better understanding of “race traitor” is *a slur word used to target any person who is perceived as having attitudes/beliefs or supporting positions that are thought to oppose the supposed attitudes/beliefs, positions, interests, advantages, or well-being of that person’s own race or ethnic group.* For those who use the term, very specific attitudes/beliefs, positions, interests, etc., are attributed to certain groups. These are properties that are supposed to define members of specific groups by virtue of their belonging to those groups. These properties are stereotypically imposed insofar as persons tend to be more complex than the properties imposed on them allow. Therefore, in some sense, these are falsely taken to be exclusive, definitive, and paradigmatic properties of groups; in a word, they are stereotypes. However, the “race traitor” would not be hated if they possessed properties that are imposed on their group. “Race traitors” are hated because they ought to possess stereotypical properties but fail to possess them. “Race traitors” are hated because they have committed a sort of racial/ethnic treason by failing to possess these properties—by failing to be legitimate members of their race or ethnic group.

“Uncle Tom”/“Aunt Jane” and the more contemporary “sellout” are “Black race traitors” who are perceived as betraying their group in some way, most often as it relates to the supposed attitudes/beliefs, positions, interests, advantages, or well-being of whites. The “Uncle Tom” and “Aunt Jane” might be perceived as being ashamed of having Black ancestry—particularly if they do not wear their hair or clothes in particular ways or if their sexual partners are not Black.<sup>29</sup> Or, they may fail to appreciate “Black culture” or honor certain historical or contemporary Black figures. “Uncle Tom” and “Aunt Jane” are especially caricatures for Black persons (most often African Americans) who are perceived as being servile to whites or other racial/ethnic groups—particularly to the detriment of themselves and other Blacks. A “Tom” or “Jane” might also fail to support policies that are thought to benefit Blacks, or may endorse policies that are thought to harm Blacks. So, one might be perceived as an “Uncle Tom” or “Aunt Jane” if one is not a progressive liberal or is critical of policies such as affirmative action, or one does not have beliefs proper to Blacks.<sup>30</sup> In certain circles, one can be called an “Uncle Tom” or “Aunt Jane” if they favor integration or assimilationism and disfavor sep-

29 The latter seems implicit in Mills’s descriptions of the “Racial Solidarity” and “Questionable Motivations” arguments that people use to explain why Black men have a moral duty to marry Black women (“Do Black Men Have a Moral Duty to Marry Black Women?”).

30 Although Christie seems unclear about the distinction between the terms “acting white” and “Uncle Tom,” he gives an account of an encounter between himself and Maxine Waters, who he claims called him an “Uncle Tom” because: “White people work for Republicans? Not African Americans!” (*Acting White*). Here, the “Uncle Tom” is one who is not liberal.

aratism or Black nationalism.<sup>31</sup> The terms might be imposed on persons who are employed with certain employers—say the police or FBI—that are seen as contributing to the oppression of Blacks. Still further, an “Uncle Tom” or “Aunt Jane” may refuse to take sides with members of the “Black race” over members of other races in circumstances wherein “allegiance to the race” is deemed to be required.<sup>32</sup> The response to all of these is: “This ‘Uncle Tom’ or ‘Aunt Jane’ isn’t down for ‘us.’” Whatever specific content motivates use of the term, “Uncle Toms” or “Aunt Janes” fail to possess appropriate properties (attitudes, interests, etc.), or fail to possess them in the right way or to the appropriate extent, given the perception of persons who target them with the terms. “Uncle Toms” and “Aunt Janes” fail to demonstrate the appropriate concern for the “Black race.” They are perceived as illegitimate and are hated for being moral failures or offenses who must be overcome. Discrimination against them becomes morally required.

The “white nigger” and “nigger lover” are “white race traitors.” Persons targeted with the terms are perceived as betraying the attitudes/beliefs, positions, interests, etc., of whites, particularly in favor of those of Blacks.<sup>33</sup> Historically, the former was a term popularized by white southerners to refer to other white persons who sided with Blacks or “Black interests” during the civil rights era, while the latter was used throughout history to characterize whites who would either involve themselves in consensual friendships/relationships with Blacks, adopt Black children, or express admiration for Black figures. Both would be applied to white persons who failed to side with whites in certain instances wherein “allegiance to the race” is deemed to be required. And currently, these terms are used to characterize white people who are critical of “whiteness” as a political ideology (those in “whiteness studies,” for example) or who are critical of “whiteness” being taken as the norm (for example, in fashion or with regard to beauty). The response is something like: “These ‘white niggers’ don’t love themselves.” And as has already been shown by appeal to Szewczuk, what makes the “white nigger” and “nigger lover” a target of hate is a perception of their failure to have the appropriate concern for white people.

Perhaps contrary to intuitions given the history of the terms, I think that one can substitute the “interests of Blacks” that allow the specific use of “nigger” for

---

(I doubt that Waters’s statement would be different if Christie was a libertarian or constitutional party member.)

- 31 “Black Nationalism, also known as black separatism, is a complex set of beliefs emphasizing the need for the cultural, political, and economic separation of African Americans from white society” (Appiah and Gates, *Africana*, 80).
- 32 For a genealogy of typical traits of and politics behind “sellout,” see Kennedy, *Sellout*.
- 33 I understand this term in contrast to those like J. L. A. Garcia, who seem to read race-traitor terms such as “nigger lover” as descriptive stereotypes (“The Heart of Racism”).

“interests of X,” without changes to the individual being perceived as a “white race traitor,” and thus as one who has committed a sort of racial treason. “Nigger lover,” would become “derogatory-term-for-nonwhite-group-X lover,” (“spic lover,” for instance), and the individual targeted by the term would be hated in exactly the same way as the “nigger lover.” Like the “Uncle Tom,” whatever specific content licenses use of the term, targeted white persons fail to possess appropriate properties, or fail to possess them in the right way or to the appropriate extent, given the perception of persons who use the terms.<sup>34</sup>

It may appear that descriptive stereotyping captures the use of these slurs. I deny this. Obviously, certain properties describe targeted individuals—servility to race X, attraction to and friendships with members of race Y, etc. Certainly there must be some historical properties (content) for these terms. Some particular—some individual or term—must be perceived or thought of as possessing properties (content) in order for it to make sense. However, as it relates to these terms, the disvalue that connects stereotypes to discrimination does not follow from specific properties per se. Discrimination derives from the perception that certain groups should possess certain properties and that they are valuable to the extent that they possess them. People targeted by these terms fail to possess these. So the “white nigger” in the South during the civil rights era is killed alongside the “nigger,” but not because of the properties that they are perceived as having—an affinity for Black humanity and a perception that Blacks are equal to whites. Rather, it is the properties that they lack—the vision of white supremacy, a heightened concern or love for white people over all other groups, and certain psychological, economic, or social interests relating to whites—that make them the target of the term, and therefore make their death morally required. The “white nigger” could just as easily be the “white chink” or the “white kike,” and the reaction will be the same: “So-and-so is a ‘race traitor,’ and deserves to be

34 The white race traitor is a bit trickier than the Black race traitor with respect to normative versus descriptive stereotyping. This is so because it is more easily grounded in a hatred or disdain of other races than the Black race traitor. Those who use white race-traitor terms more often take other racial/ethnic groups to be subhuman or less. For example, when a Black political discriminator uses the term “Uncle Tom,” it is less likely that they perceive whites as subhuman or less human than Blacks. Thus, when some Black person is an “Uncle Tom” or “Aunt Jane” for a political discriminator, it is less clear that what makes the Black person contemptible are the actual properties that they have. Now, Szewczuk may think that Prince Harry—being a white person—has some duty to continue the “pure” white race by marrying a white woman. As a result, Szewczuk may not think that marrying a nonwhite person is contemptible *simpliciter*. However, being a neo-Nazi, Szewczuk probably thinks that Black people just are contemptible, and that attraction to them is a property that makes white people deserving of contempt. The latter is less often the case when Black, Jewish, or Asian political discriminators use race-traitor terms.

assassinated.” Currently, I would imagine that, to many supporters of what is now called “white nationalism,” white persons who either do not support it or deplore it as yet another kind of white supremacy are white race traitors. Why? They fail to support a certain vision that they ought to support by virtue of being white.

Further, it is commonly thought that the use of “Uncle Tom” indicates an awareness that the target has either an overvalued conception of whiteness or a belief that whites are superior to Blacks. However, there have been “race men” who struggled for Black equality, engaged in cultural practices typically understood to be Black, supported certain policies perceived as beneficial to Black people, etc., but who were disparaged and delegitimated by the term, and thereby discriminated against for merely engaging in relationships with women who were not Black.<sup>35</sup> There need be no perception that targets believed that whites generally, or white and non-Black women specifically, are superior to Blacks. Failing to engage in a relationship with women whom they ought motivated use of the term. These men committed a kind of racial treason by merely loving members of non-Black racial groups. Additionally, it seems flatly inaccurate to think that every Black conservative, every person who is employed as a police officer, or every Black person who opposes policies like affirmative action, and who have been the target of the term, overvalue a conception of whiteness or believe that white people are superior to Black people. They merely supported policies or positions, or held jobs, that were perceived as “not Black.” Now these terms all have a very particular history, and thus are more often used in certain situations. However, the point of using them seems to be the same—namely, to pick out individuals who are illegitimate members of some race or ethnic group because they fail to satisfy some stereotypical image of the race or ethnic group.

The above slurs (“Uncle Tom” and “nigger lover”) are different from ones like “nigger,” “kike,” and “chink” in at least two ways. First, they tend to be imposed by in-group members, while the latter tend to be imposed by members of out-groups. Second, the latter—however they are defined—seem to impose descriptive, disvalued, stereotypical properties on their targets.<sup>36</sup> The “nigger” is “lazy,” “ignorant,” etc. The “kike” is “avaricious,” “deceitful,” etc. The “chink” is “untrustworthy,” “shifty,” etc. Discrimination from these follows descriptive stereotyping. “Race-traitor” terms—however they are defined—attempt to impose certain descriptive stereotypical properties on targets by imposition on groups to which targets belong. However, hatred directed at targets derives from their failure to possess these properties. Discrimination from these follows normative stereo-

35 See Kennedy, *Sellout*, 64n.

36 For discussion of work on the relationship between slurs and stereotypes, see Jeshion, “Slurs and Stereotypes.”

typing. What all of these slurs have in common is that they either recommend or require political discrimination merely by use.<sup>37</sup>

When political discriminators use race-traitor terms such as “Uncle Tom” or “nigger lover,” they perform three acts. First, discriminators signal that some target is an illegitimate individual of some group. They have failed to possess all those properties deemed proper to them, and in a way that satisfy expectations of some agent making judgments about them.

Second, rather than merely signaling that race traitors are contemptible, discriminators direct an attitude of moral hatred at the target for being illegitimate. The standard that the target fails to meet is a serious moral standard for which mere contempt cannot account. Following Jean Hampton, I understand moral hatred as “an aversion to someone who has identified himself [or herself] with an immoral cause or practice, prompted by moral indignation and accompanied by the wish to triumph over him [or her] and his [or her] cause or practice in the name of some fundamental moral principle or objective, mostly notably justice.”<sup>38</sup> Users are committed to the view that the race traitor is a treacherous being who has committed to some perverse cause over which the discriminator—being morally superior—has an obligation to triumph.

Finally, because the target is a treacherous being who must be triumphed over, discriminators are signaling to others that poor treatment against the target is more than merely justified, but required. The requirement of poor treatment is often more austere than political discrimination that is considered justified in more common cases of normative stereotyping. Michal Szewczuk recommended assassination for Prince Harry’s racial treachery. Martin Delany was shot at after having been branded a race traitor.<sup>39</sup> Fannie Lou Hamer, when forcefully declaring that Uncle Toms must be stopped, exclaimed: “I don’t believe in killing, but a good whipping behind the bushes wouldn’t hurt them.”<sup>40</sup> So, the political discriminator—in using race-traitor terms—expresses something like: “Because you fail to possess certain properties given the group to which you belong, you are illegitimate. More than just being a failure, you are a moral offense. I hate you! And because you must be triumphed over, I demand you be treated poorly (or significantly worse than legitimate members).”

It is obvious that discrimination from normative stereotyping need not in-

37 Hom has proposed that slurs possess thick, negative, truth-conditional content (“A Puzzle about Pejoratives”). And so, slurs prescribe behavior based on an evaluation from a fact. This account is referred to as normative descriptivism (Cappelen and Dever, *Bad Language*, 93).

38 Hampton, “Forgiveness, Resentment and Hatred,” 61.

39 Kennedy, *Sellout*, 38.

40 Kennedy, *Sellout*, 49.

clude a term that captures the political discriminator's hatred for illegitimate targets. Further, a political discriminator need not use "rate traitor" or race-traitor terms to direct moral hatred at some target. In the preceding, the discussion merely regarded those who commit to the use of these terms.

I understand cases wherein race-traitor terms are used as specific cases of first-order political discrimination from normative stereotyping. These discriminators, in using these terms and thus recommending the aforementioned acts, have an unashamed commitment to particular stereotypical images of groups. They use these terms to target persons for whom they have moral hatred. Still, there are cases of higher-order political discrimination from normative stereotyping.

Recall that, for Piper, higher-order political discrimination is "the attitude within which a primary disvalued or valued property in turn confers disvalue or value respectively on further properties of the disvaluee or valuee respectively." With normative stereotyping, certain things that a person—call her or him the disvaluee—either does or fails to do are judged through the light of a failure to possess all those attributes deemed proper to a discriminator's stereotypical image of how group members ought to be. So, let us say that a person has a stereotypical image of female academics such that a woman is a legitimate female academic to the extent that she works in feminism. A woman who works on figures such as Hegel or Kant, but who does not work in feminism, might have her work on Hegel or Kant negatively evaluated in light of her failure to produce feminist scholarship. The discriminator judges that this academic fails to meet an important standard—namely, the standard of legitimacy as a female academic. The discriminator, being disappointed by and frustrated with the academic's failure, judges that the academic is contemptible. This judgment produces a feeling of superiority in the discriminator that justifies the exercise of (unjust) power over the "inferior" target in a way that harms her.

Recall also that a marked distinction between the first-order and higher-order political discriminator is that the latter rejects the imposition of stereotypes, and would reject that they impose the stereotypes on others. A part of these discriminators' self-concept is tied to being this kind of person. And so, higher-order political discriminators deceive themselves in some way so as to be blind to both their stereotyping and the discrimination deriving from it. The higher-order political discriminator who judges the academic's work in accordance with her failure to be a legitimate female academic denies or suppresses their normative stereotyping by attributing disvalue to higher-order properties—the academic's failure as a Hegelian or Kantian. The contempt is taken to be licensed by a failure to produce good Hegelian or Kantian scholarship, and not a failure to be a legitimate female academic. This deception preserves the discriminator's self-concept.

A discriminator's appeal to the stereotypical image of the "female academic" may cause them to evaluate a woman's non-feminist academic work in light of her failing (or succeeding) to satisfy the stereotypical image. A non-feminist academic is deemed a bad Hegelian or Kantian because she fails to produce feminist work. She may be so deemed even if her work on Hegel or Kant should/could be unrelated to work on feminism.<sup>41</sup> And this may lead to an unwillingness to hire or promote the academic, or failure to review her work on Hegel or Kant fairly.

Now normative stereotyping and the political discrimination that follows from it are both less recognizable and much more tolerated when recognized than descriptive stereotyping and the political discrimination that follows from it. Normative stereotyping is less recognizable because it receives far less attention than descriptive stereotyping. We have become more sensitive to acts of racism, sexism, etc. And these acts—when linked to beliefs—are connected to stereotypical traits that group members are thought to possess on the basis of belonging to groups. When people are treated in ways that harm them because they are thought to possess stereotypical disvalued group traits, we recognize the treatment as both discriminatory and immoral. We recognize the treatment in this way because it derives from racist or sexist beliefs, and we take these beliefs to be more than simply epistemically wrong, but immoral. This is not the case with normative stereotyping, which does not impose stereotypical disvalued group traits on members. Thus, it is not connected to racism, sexism, etc. When people are treated in ways that harm them on the basis of failing to possess proper traits—that is, traits that they *ought* to possess—we often fail to categorize the treatment as discriminatory because it is not easily connected to racist or sexist beliefs. There is much more difficulty naming this treatment as discriminatory and tracing it to a harmful kind of stereotyping.

Normative stereotyping and the political discrimination that follows from it are also much more tolerated or respectable even when recognized. This is for four reasons. First, when we recognize the imposition of normative stereotypes, we less often think that it will motivate discriminatory practices against those who fail to possess the stereotypes. We tend to take this the imposition to be "kooky," "odd," "old-timey," or "folksy" behavior that is unserious. Second, when normative stereotyping is recognized, it tends not to be recognized as unjust in the way that descriptive stereotyping is. In a sense, because the stereotyping is not commonly recognized as unjust, actions that follow from it are not taken to be discriminatory. And so, for these two reasons, normative stereotyping tends

41 Further, one's being a good Hegelian or Kantian might depend on producing work in feminism that is in vogue. One may have to be the "right kind of feminist" in order to have other attributes legitimated.



not to be accorded the kind of seriousness that it should be accorded. Third, even when we recognize the stereotype as normative and that certain treatments follow from it, we are less likely to acknowledge the treatment as immoral. There tends to be buy-in for certain race-traitor terms and ideas surrounding them—particularly those that derive from groups that have been the targets of out-group hate. Use of the terms and promulgation of the ideas surrounding them legitimate treatment in a way that makes the target—and not the discriminator—responsible for the treatment. Fourth, with out-group normative stereotyping, there is often an endorsement of some stereotypical image by the discriminator who does not fit the stereotype themselves. As a result of the discriminator's acceptance of and preference for a stereotypical image, there is often an unwillingness to call the discriminator racist, sexist, etc. For non-Black persons who prefer some stereotypical image of, say Black women, it might appear to others that they are “culturally pluralist.” Persons who make statements such as: “I need to channel my inner Black woman!” are very often taken to “embrace difference.”<sup>42</sup> This shields them in a particular way. And so vocalizing contempt for and acting in a discriminatory way toward non-stereotypical targets are less objectionable to people around them. It can be voiced in public without as much interrogation and criticism, so long as the stereotypical image is in vogue.

Consider the following statements:

1. White man says: “You know those damned Black women are committed to ‘*telling it like it is*’—as they call it. Pitiful. So you know what you’d get from C” (who is Black).

Here, the speaker promotes “telling it like it is” (speaking an uncomfortable truth in a blunt, unvarnished, and perhaps indelicate way) as a disvalued descriptive stereotype. It will be recognized as such, and will most likely be connected to sexist and racist beliefs about Black women. Any act that harms C on the basis of a perception that she possesses this trait will immediately be recognized as discriminatory and immoral, and will neither be tolerated nor respected because of its connection to sexism and racism. Note the difference between 1 and the following:

2. White woman/Black man says: “You know Black women *tell it like it is*. They give it to you straight, and that’s a good thing. But C [who is Black] doesn’t. I don’t know what’s wrong with her.”
3. Black woman says: “You know we Black women *tell it like it is*. C doesn’t

42 This kind of expectation that masks as “embracing difference” explains Als’s statement: “The sad fact is that in order to cross over, most black actors of [Anthony] Mackie’s generation must ‘act black’ before they’re allowed to act human” (“Underhanded”).

though. You know she's not really one of us" (she's not "Black woman" enough).

Though these promote a (normative) stereotype, they may fail to be recognized as such. And even when recognized as promoting a stereotype, they will often fail to be granted the kind of seriousness that would motivate discrimination because they will not be connected with sexism and racism. The statement may be construed as the "odd" behavior of an overly "down" (committed) but good-hearted liberal white woman, or the "folksy" beliefs/"talk" of Blacks. Statement 2 presents dialogue that endorses the stereotyped image by a discriminator who does not fit the stereotype herself. The endorsement of the stereotype lessens the likelihood that the discriminator can be named in a particular way that presents the treatment as problematic. With 3, acts that harm *C* on the basis her failure to possess this valued attribute may fail to be recognized as immoral. If there is group buy-in for "telling it like it is," and particularly if it is raised to the level of a group virtue, *C*—not the discriminator(s)—will be thought to bear the responsibility for any harmful treatment. And so when treatment is recognized as deriving from imposed stereotypes, it may be tolerated or even respected.

### 3. "BLACK AND BLUE"—CONCLUSION

I would like to close with a few remarks on the moral considerations of normative stereotyping and the discrimination that follows from it. For Piper: "Instances of first-order discrimination are familiar targets of moral condemnation because they disvalue individuals for having attributes perceived as primary disvalued attributes that are not in actuality sources of disvalue."<sup>43</sup> Individuals' race, sexuality, gender, class, etc., are attributes possessed by individuals that confer disvalue to them. In actuality, however, these properties are not signals of inferiority. This makes both the descriptive stereotyping and the discrimination that follows unwarranted and morally condemnable. In short, discrimination from descriptive stereotyping violates basic notions of fairness. One is judged to lack competence or ability in some way unrelated to a job, skill, responsibility, etc., that is under discussion, and this is unfair.

In addition to a notion of fairness, I take normative stereotyping and the discrimination that follows from it to promote an unjust restriction on autonomy. There is value to individuals (particularly with different emotional dispositions, tastes, and perspectives) cultivating themselves in ways that are expressive of certain properties that they choose. Individuals' own choices regarding mating, aca-

43 Piper, "Higher-Order Discrimination," 286.

demic pursuits, music or culture, political leanings, etc., are important to an individual's flourishing. These choices (*ceteris paribus*) ought to be respected because individuals' ability to fashion their own lives in a way that captures their conception of what it means to flourish ought to be respected. Judging persons to be legitimate only insofar as they satisfy some stereotypical image of a group that political discriminators have either set up or to which they appeal, and then discriminating against individuals who do not satisfy the image, unjustly restricts autonomy.

Now, a more complete work is necessary to provide justification for this account.<sup>44</sup> And further, one might think that there are justificatory reasons to defend the use of and ideas behind race-traitor terms such as "Uncle Tom"/"Aunt Jane" or "sellout"—even if there are no such reasons to defend other race-traitor terms ("nigger lover" and "white nigger"). There may be moral asymmetries that give members of certain groups obligations to be legitimate members of some group. One may take appeals to legitimate group members and race-traitor terms to be necessary for liberation, justice, group survival, honor, or self-respect. Still further, one might take there to be good reasons to support group virtues like "telling it like it is," such that group members who do not possess these group virtues are vicious. One might ground these virtues in historical contexts whereby the group virtue has become necessary. So, say that "telling it like it is" derives from and is integrally connected to "speaking truth to power." And say that "speaking truth to power" has become a necessary virtue for an oppressed people seeking liberation. One might think that there is good reason to believe that members of the oppressed group are virtuous to the extent that they possess the trait, and vicious—deserving contempt or moral hatred—to the extent that they do not. Though I cannot treat these cases here, I ultimately think that these are problematic because they are not weighty enough to override individual autonomy. Agents' ability to fashion their own lives ought to be respected, which is to say that individuals ought to be able to define themselves. Normative stereotyping and the discrimination that follows from it harm those who do not fit the discriminator's definitions.

*University of San Diego*  
coreybarnes@sandiego.edu

44 In a forthcoming work I seek to treat moral considerations of normative stereotyping and discrimination that derives from it. In so doing I cede possible moral asymmetries for different groups that would license normatively stereotyping members of oppressed groups. Therein I philosophically engage a number of arguments that would seek to justify race-traitor terms such as "Uncle Tom"/"Aunt Jane" and "sellout."

## REFERENCES

- Akerlof, George A., and Rachel E. Kranton. "Economics and Identity." *Quarterly Journal of Economics* 115, no. 3 (August 2000): 715–53.
- . *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*. Princeton: Princeton University Press, 2010.
- Als, Hilton. "Underhanded: Martin McDonagh's Slap in the Face." *The New Yorker*, March 15, 2010. <https://www.newyorker.com/magazine/2010/03/15/underhanded>.
- Appiah, Kwame Anthony. *The Ethics of Identity*. Princeton: Princeton University Press, 2005.
- . "Stereotypes and the Shaping of Identity." *California Law Review* 88, no. 1 (January 2000): 41–53.
- Appiah, Kwame Anthony, and Henry Louis Gates Jr. *Africana: The Encyclopedia of the African and African American Experience*. Philadelphia: Running Press, 2003.
- Bartky, Sandra Lee. "Foucault, Femininity, and the Modernization of Patriarchal Power." In *Writing on the Body: Female Embodiment and Feminist Theory*, edited by Katie Conboy, Nadia Medina, and Sarah Stanbury, 129–54. New York: Columbia University Press, 2006.
- Bell, Macalester. *Hard Feelings: The Moral Psychology of Contempt*. New York: Oxford University Press, 2013.
- Berdahl, Jennifer L. "Harassment Based on Sex: Protecting Social Status in the Context of Gender Hierarchy." *Academy of Management Review* 32, no. 2 (April 2007): 641–58.
- Bordo, Susan. *Unbearable Weight: Feminism, Western Culture, and the Body*. Berkeley: University of California Press, 1993.
- Cappelen, Herman, and Josh Dever. *Bad Language: Contemporary Introductions to Philosophy of Language*. Oxford: Oxford University Press, 2019.
- Carbado, Devon W., and Mitu Gulati. *Acting White? Rethinking Race in Post-Racial America*. New York: Oxford University Press, 2013.
- Christie, Ron. *Acting White: The Curious History of a Racial Slur*. New York: Oxford University Press, 2013.
- Collins, Patricia Hill. *Black Feminist Thought*. 2nd ed. New York: Routledge, 2009.
- De Beauvoir, Simone. *The Second Sex*. Translated by Constance Borde and Sheila Malovany-Chevallier. New York: Vintage Books, 2009.
- Du Bois, W. E. B. *The Philadelphia Negro: A Social Study*. Philadelphia: University of Pennsylvania Press, 1995.

- Fordham, Signithia, and John U. Ogbu. "Black Students' School Success: Coping with the 'Burden of "Acting White."'" *Urban Review* 18, no. 3 (September 1986): 176–206.
- Friedan, Betty. *The Feminine Mystique*. New York: W.W. Norton and Company, 2001.
- Garcia, J. L. A. "The Heart of Racism." In *Race and Racism*, edited by Bernard Boxill. 257–96. New York: Oxford University Press, 1996.
- Haddock, Geoffrey, and Mark P. Zanna. "Preferring 'Housewives' to 'Feminists': Categorization and the Favorability of Attitudes toward Women." *Psychology of Women Quarterly* 18, no 1 (March 1994): 25–52.
- Hampton, Jean. "Forgiveness, Resentment and Hatred." In *Forgiveness and Mercy*, edited by Jeffrie G. Murphy and Jean Hampton, 35–87. New York: Cambridge University Press, 1998.
- Heilman, Madeline E. "Description and Prescription: How Gender Stereotypes Prevent Women's Ascent up the Organizational Ladder." *Journal of Social Issues* 57, no. 4 (2001): 657–74.
- . "Gender Stereotypes and Workplace Bias." *Research in Organizational Behavior* 32 (2012): 113–35.
- Hom, Christopher. "Pejoratives." *Philosophy Compass* 5, no. 2 (February 2010): 164–85.
- . "A Puzzle about Pejoratives." *Philosophical Studies* 159, no. 3 (July 2012): 383–405.
- . "The Semantics of Racial Epithets." *Journal of Philosophy* 105, no. 8 (August 2008): 416–40.
- Jeshion, Robin. "Slurs and Stereotypes." *Analytic Philosophy* 54, no. 3 (September 2013): 314–29.
- Kant, Immanuel. "Observation on the Feeling of the Beautiful and the Sublime." In *Anthropology, History, and Education: The Cambridge Edition of the Works of Immanuel Kant*, edited by Robert B. Loudon and Günter Zöllner, 18–62. New York: Cambridge University Press, 2007.
- Kennedy, Randall. *Sellout: The Politics of Racial Betrayal*. New York: Vintage Books, 2008.
- Melendez, Pilar. "British Neo-Nazi Who Threatened 'Race Traitor' Prince Harry Sentenced to Four Years in Prison." *Daily Beast*, June 18, 2019. <https://www.thedailybeast.com/michal-szewczuk-british-neo-nazi-who-threatened-race-traitor-prince-harry-sentenced-to-4-years-in-prison>.
- Mills, Charles W. "Do Black Men Have a Moral Duty to Marry Black Women?" *Journal of Social Philosophy* 25, no. s1 (June 1994): 131–53.
- Moss-Racusin, Corinne A., Julie E. Phelan, and Laurie A. Rudman. "When Men

- Break the Gender Rules: Status Incongruity and Backlash against Modest Men." *Psychology of Men and Masculinity* 11, no. 2 (2010): 140–51.
- Ogbu, John U. *Black American Students in an Affluent Suburb: A Study of Academic Disengagement*. Mahwah: Lawrence Erlbaum Associates, 2003.
- Piper, Adrian M. S. "Higher-Order Discrimination." In *Identity, Character, and Morality: Essays in Moral Psychology*, edited by Owen Flanagan and Amelie Oksenberg Rorty, 285–309. Cambridge, MA: MIT Press, 1990.
- . "Pseudorationality." In *Perspectives on Self-Deception*, edited by Brian P. McLaughlin and Amelie Oksenberg Rorty, 297–323. Berkeley: University of California Press, 1998.
- . "Two Kinds of Discrimination." In *Race and Racism*, edited by Bernard Boxill, 193–237. New York: Oxford University Press, 1996.
- Post, Robert C. *Prejudicial Appearances: The Logic of American Antidiscrimination Law*. Durham, NC: Duke University Press, 2002.
- Roberts, Robert C. *Emotions: An Essay in Aid of Moral Psychology*. New York: Cambridge University Press, 2003.
- Rosette, Ashleigh Shelby, Jennifer S. Mueller, and R. David Lebel. "Are Male Leaders Penalized for Seeking Help? The Influence of Gender and Asking Behaviors on Competence Perceptions." *Leadership Quarterly* 26, no. 5 (October 2015): 749–62.
- Rudman, Laurie A., and Kimberly Fairchild. "Reactions to Counterstereotypic Behavior: The Role of Backlash in Cultural Stereotype Maintenance." *Journal of Personality and Social Psychology* 87, no. 2 (2004): 157–76.
- Truth, Sojourner. *Narrative of Sojourner Truth: A Bondswoman of Olden Time, with a History of Her Labors and Correspondence Drawn from Her Book of Life*. Edited by Nell Irvin Painter. New York: Penguin Books, 1998.
- X, Malcolm. "Not Just an American Problem, but a World Problem." In *February 1965: The Final Speeches*, edited by Steve Clark, 147–76. New York: Pathfinder, 1992.
- Young, Iris Marion. "Throwing Like a Girl: A Phenomenology of Feminine Body Comportment Motility and Spatiality." *Human Studies* 3, no. 1 (April 1980): 137–56.

## WHAT IS THE BAD-DIFFERENCE VIEW OF DISABILITY?

Thomas Crawley

THE BAD-DIFFERENCE VIEW (BDV) of disability postulates a negative connection between disability and well-being.<sup>1</sup> It says, roughly, that disability makes one worse off. The Mere-Difference View (MDV) of disability says, roughly, that it does not. In recent work, Barnes—an MDV proponent—offers a detailed exposition of the MDV.<sup>2</sup> No BDV proponent has done the same. While many make it clear that they endorse a BDV, they do not carefully articulate their view.<sup>3</sup> And various views might constitute a BDV, so it is important to determine which is best and most likely to be endorsed by BDV proponents.

In this paper, I clarify the nature of the best version of the BDV by discussing two issues—instrumentality and probability—that must be settled for a full characterization to be properly developed. Modifying and expanding upon work from Campbell and Stramondo, I argue that the BDV's best interpretation is probabilistic and compares the overall value of disability and non-disability.<sup>4</sup> It is, roughly, the view that a person is *likely to be, all things considered*, worse off with a disability than without. Thus, Barnes—who criticizes the view that disability *by itself, intrinsically or automatically* makes one worse off—does not challenge the BDV's best interpretation.<sup>5</sup> She attacks a version unlikely to be endorsed by BDV proponents and misses an opportunity to challenge the most plausible and relevant version. As such, one can be persuaded by Barnes's arguments and still hold a plausible version of the BDV, and the best version remains unchallenged.

- 1 The BDV has also been called the "Standard View" (Amundson, "Disability, Ideology, and Quality of Life," 103; Campbell and Stramondo, "The Complicated Relationship of Disability and Well-Being," 151) and the "Received View" (Schramme, "Disability (Not) as a Harmful Condition").
- 2 Barnes, *The Minority Body*.
- 3 E.g., McMahan, "Causing Disabled People to Exist and Causing People to Be Disabled"; Shakespeare, *Disability Rights and Wrongs Revisited*; Singer, "Ethics and Disability."
- 4 Campbell and Stramondo, "The Complicated Relationship of Disability and Well-Being."
- 5 Barnes, "Disability, Minority, and Difference" and *The Minority Body*.



Developing a proper understanding of the BDV (and MDV) is important. This debate is new and underdeveloped and it is important that it gets off on the right footing with clear and precise understandings of the views. Moreover, the BDV is often endorsed but rarely defended, and most discuss it primarily to criticize it.<sup>6</sup> Articulating the strongest account from a charitable perspective is something missing from the literature that will be of interest to many philosophers. In addition, this debate arose in response to the real concerns of disabled people who wanted to change conceptions of disability.<sup>7</sup> They too have an interest in the views being properly characterized. Finally, questions about disability's relationship to well-being are plausibly relevant to many contemporary normative debates and real-world, applied issues, e.g., about what justice requires for disabled people and the (im)permissibility of prenatal selection for or against disability. Understanding the natures of the views may well be a vital step toward advancing these other important debates.<sup>8</sup>

In section 1, I present six candidate interpretations of the BDV. In section 2, I discuss how I understand disability and well-being. In section 3, I present two important desiderata for identifying the best version of the BDV and rule out five candidate understandings because they fail to adequately satisfy these desiderata. This leaves the probabilistic, all-things-considered interpretation, which I argue best satisfies the desiderata and is thus the best version. In section 4, I show that Barnes characterizes the BDV uncharitably. Thus, she attacks a version unlikely to be endorsed by many, and she misses the opportunity to challenge the most plausible and relevant version of the BDV. Finally, in section 5, I respond to an objection.

### 1. CANDIDATE INTERPRETATIONS

Roughly speaking, BDV proponents believe that disability makes one worse off.<sup>9</sup> Singer holds that "other things equal, it is better not to be disabled."<sup>10</sup> And

- 6 See, e.g., Barnes, *The Minority Body*; Campbell and Stramondo, "The Complicated Relationship of Disability and Well-Being."
- 7 No terminology is uncontroversial here. Some prefer "people with disabilities." I choose "disabled people" primarily because it reflects how we describe other minority groups, e.g., "Black people" (Barnes, *The Minority Body*, 6).
- 8 Although see Schroeder for skepticism of the importance of the connection between disability and well-being to selection debates ("Well-Being, Opportunity, and Selecting for Disability").
- 9 See, e.g., McMahan, "Causing Disabled People to Exist and Causing People to Be Disabled"; Shakespeare, *Disability Rights and Wrongs Revisited*; Singer, "Ethics and Disability."
- 10 Singer, "Ethics and Disability," 113.

Shakespeare says that “disability makes it harder to have a good life.”<sup>11</sup> MDV proponents disagree.<sup>12</sup> They think that disability does not make one worse (or better) off. For example, Barnes says that “having a disability is something that makes you different, but not . . . worse off.”<sup>13</sup>

In my attempt to explicate the best interpretation of the BDV, my starting point is work by Campbell and Stramondo in which they present three candidate interpretations:

1. Disability is non-instrumentally bad for well-being.
2. Disability is instrumentally bad for well-being.
3. Disability is comparatively bad for well-being.<sup>14</sup>

If disability is bad for well-being, it must be so either instrumentally, non-instrumentally, or both. Something is non-instrumentally bad if it is bad *in itself* or *independently of its causal consequences*.<sup>15</sup> Something is instrumentally bad if it is bad *in virtue of its causal consequences*. That is, if it prevents non-instrumental goods or causes non-instrumental bads. This causation can be direct or indirect. Pleasure and pain are typically thought to be non-instrumentally good and bad, respectively. On this view, things that prevent pleasure (e.g., lack of disposable income) or cause pain (e.g., stubbing a toe) are instrumentally bad.

According to 1, disability is bad for well-being non-instrumentally. That is, disability itself contains more bad than good.<sup>16</sup> (I use “*x* contains *y*” and “*x* involves *y*” to mean that “*y* is a feature of *x* itself.” This contrasts with things that *x* causes.) If pleasure and pain were the only non-instrumental values, 1 would be saying that disability contains more pain than pleasure; 1 makes no claims about disability’s instrumental value.

Another way of saying that something is non-instrumentally bad for well-being is that it is non-instrumentally *harmful*. Harm has been understood comparatively and non-comparatively.<sup>17</sup> On non-comparative accounts, disability is non-instrumentally harmful (roughly) iff (when discounting its effects) it puts

11 Shakespeare, *Disability Rights and Wrongs Revisited*, 103.

12 See, e.g., Amundson, “Disability, Ideology, and Quality of Life”; Barnes, “Disability, Minority, and Difference” and *The Minority Body*.

13 Barnes, *The Minority Body*, 78.

14 Campbell and Stramondo, “The Complicated Relationship of Disability and Well-Being.”

15 Non-instrumental value is sometimes called “intrinsic value.” I choose “non-instrumental value” to avoid giving the impression that something is valuable in itself only in virtue of its intrinsic properties (Korsgaard, “Two Distinctions in Goodness”). For more discussion of the nature of non-instrumental value, see O’Neill, “The Varieties of Intrinsic Value.”

16 Campbell and Stramondo, “The Complicated Relationship of Disability and Well-Being,” 154.

17 Bradley, “Doing Away with Harm.”

one in a non-instrumentally bad state—where “bad” is understood in the absolute sense.<sup>18</sup> My well-being is bad in an absolute sense if it falls below some threshold of neutrality such that I have a life not worth living. So,  $\mathbf{1}$  understood non-comparatively says that, discounting disability’s effects, disabled lives are not worth living. But BDV proponents do not seem to be saying this. First, because it is utterly implausible. And second, because when understood non-comparatively,  $\mathbf{1}$  is consistent with disabled lives being non-instrumentally *better* than non-disabled lives. That is,  $\mathbf{1}$  is true on the non-comparative understanding if disability is *slightly* (non-comparatively, non-instrumentally) bad but non-disability is *extremely* (non-comparatively, non-instrumentally) bad. But this situation hardly exemplifies a BDV.

Comparative accounts of harm say (roughly) that disability is non-instrumentally harmful iff (when discounting disability’s effects) I am worse off with a disability than I would have been without.<sup>19</sup> This is a claim about *worseness* rather than absolute badness, so it would be false if disability is non-instrumentally bad but non-disability is non-instrumentally worse. This seems like the right result. Moreover, it is at least somewhat plausible that disability is non-instrumentally, comparatively bad for well-being. These considerations speak in favor of understanding  $\mathbf{1}$  as utilizing a comparative account of harm. In any case, the debate is more interesting if understood as comparative.<sup>20</sup>

Interpretation 2 focuses on disability’s instrumental value: its causal consequences. It says that disability is instrumentally bad, that it *causes* more badness than goodness.<sup>21</sup> Interpretation 2 is silent on disability’s non-instrumental value.

Interpretation 2 is also best understood as comparative. Instrumental value concerns causation, which is itself a comparative notion.<sup>22</sup> Something is instrumentally good if it causes me to be better off *than I would have been without it*. To determine whether disability is instrumentally bad, then, we need to compare the effects of disability with those of non-disability.

Interpretation 3 considers instrumental *and* non-instrumental value together, or disability’s *all-things-considered* value. It makes the comparative claim that a person’s well-being is overall worse with a disability than it would have been without.<sup>23</sup> Note that a disability *D* might be non-instrumentally good but com-

18 Bradley, “Doing Away with Harm,” 398–401.

19 Bradley, “Doing Away with Harm,” 396; cf. Brown, “Is Disability a Neutral Condition?”

20 For discussion of comparative and non-comparative evaluations within the disability context, see Schramme, “Disability (Not) as a Harmful Condition,” 86–89.

21 Campbell and Stramondo, “The Complicated Relationship of Disability and Well-Being,” 155.

22 Lewis, “Causation.”

23 Campbell and Stramondo, “The Complicated Relationship of Disability and Well-Being,” 155.

paratively bad, if *D* is instrumentally bad to a sufficient extent. Likewise, *D* might be instrumentally good but comparatively bad. Both 1 and 2 must be true to guarantee 3.

Campbell and Stramondo do not discuss whether they understand 1 as comparative.<sup>24</sup> Another reason for doing so is that—because 2 and 3 are best understood as comparative—understanding 1 as non-comparative would mean that their taxonomy does not provide a neat way of carving up potential BDV understandings. Rather, it would seem to run together comparative and non-comparative value with *pro tanto* and all-things-considered value. Relatedly, it is unclear why 3 is described as the view that disability is “comparatively bad” given the comparative natures of 1 and 2. As such, I will henceforth refer to views that consider disability’s instrumental and non-instrumental value together as considering its “all-things-considered” or “overall” value.

Notice that all the candidate interpretations are compatible with many, most, or all disabled people having overall good lives. They make no claims about the absolute well-being levels.

Also note, crucially, that, following Barnes, I take all understandings to be discounting negative effects arising from unjust discrimination against disabled people, or “disablism.”<sup>25</sup> Clearly, many of disability’s negative effects—including, e.g., being treated as incompetent and discriminated against in the job market—would not occur in a non-disablist world. This has been long established in disability scholarship and is accepted by thinkers on both sides.<sup>26</sup> I take all understandings of the BDV to hold that disability would be bad for well-being *even in a non-disablist world*. Note, however, that theorizing about a non-disablist world does not imply that the only relevant disadvantages are intrinsic features of disability. Social practices might instrumentally disadvantage disabled people without doing so *unfairly*.<sup>27</sup>

Comparing the well-being of disabled and non-disabled people raises issues about identity. If disabilities are identity determining, we cannot compare *a particular person’s* well-being with a disability and without.<sup>28</sup> I will not discuss whether disabilities are identity determining. Both MDV and BDV proponents

24 Perhaps their labeling of 3 suggests that they do not.

25 Barnes, *The Minority Body*.

26 See, e.g., Oliver, *Understanding Disability*; Barnes, *The Minority Body*; and Singer, “Response to Mark Kuczewski,” 56. For statistics that suggest the disablist nature of our world, see Office for Disability Issues, “Public Perceptions of Disabled People.”

27 Amundson, “Disability, Ideology, and Quality of Life,” 114; Howard and Aas, “On Valuing Impairment,” 1129.

28 Campbell and Stramondo, “The Complicated Relationship of Disability and Well-Being,” 161.

must overcome this problem. One solution might be to insist that if disability is identity determining we can still compare the well-being of metaphysically different but relevantly similar people. And, plausibly, numerical identity need not be preserved for comparisons between individuals to bear moral relevance.<sup>29</sup>

So, Campbell and Stramondo present three candidate interpretations of the BDV, two that consider *pro tanto* value and one that considers overall value. However, once we have determined whether the BDV is interested in instrumental, non-instrumental, or overall value, at least one more important question must be answered before we can fully characterize it. BDV proponents might think that disability is *always* bad for well-being, or they might take a probabilistic approach.<sup>30</sup> Given this, there are at least six plausible candidate interpretations. I, therefore, propose to modify Campbell and Stramondo's taxonomy thusly (here I suppress clauses common to all views, which I reintroduce below):

- 1\*. Disability is *always* non-instrumentally bad for well-being.
- 2\*. Disability is *typically* non-instrumentally bad for well-being.
- 3\*. Disability is *always* instrumentally bad for well-being.
- 4\*. Disability is *typically* instrumentally bad for well-being.
- 5\*. Disability is *always* all-things-considered bad for well-being.
- 6\*. Disability is *typically* all-things-considered bad for well-being.

In section 3, I argue that 6\* is the best interpretation. Before that, I will clarify key terms.

## 2. UNDERSTANDING DISABILITY AND WELL-BEING

For the purposes of this paper, I understand disability using Barnes's Solidarity Account, which says that a person *S* is physically disabled in a context *C* iff (i) *S* is in some bodily state *x*, and (ii) the rules for making judgments about solidarity employed by the disability rights movement classify *x* in *C* as among the physical conditions that they are seeking to promote justice for.<sup>31</sup> Following Barnes and others in the debate, I restrict my discussion to physical (including sensory) disabilities.<sup>32</sup> This significantly reduces the heterogeneity of the group

29 Kahane and Savulescu, "The Welfarist Account of Disability," 37.

30 Campbell and Stramondo mention that probabilistic versions are possible, but they do not consider probabilistic versions of all candidate understandings ("The Complicated Relationship of Disability and Well-Being," 168).

31 Barnes, *The Minority Body*, 46.

32 Barnes, *The Minority Body*, 2–3; and see, e.g., Putnam, "Disability, Democratic Equality, and Public Policy," 306.

I am reasoning about and avoids me having to answer difficult questions about how to evaluate the well-being of people with nonphysical disabilities.<sup>33</sup> I am hopeful that my arguments can be extended to nonphysical disabilities, but I do not make or defend that claim here.

Why the Solidarity Account? Its principal attraction is that it allows meaningful investigation into the connection between disability and well-being. On Welfarist Accounts, such as Kahane and Savulescu's, to be disabled just is (roughly) to have a condition that is detrimental to well-being.<sup>34</sup> These accounts thus preclude meaningful investigation into whether the MDV or BDV is true. Similarly, the MDV is trivially true on strong versions of the Social Model—which understand disability as the unjust oppression of people with certain bodily features (known as *impairments*) and imply that disability disadvantages result *entirely* from this oppression.<sup>35</sup> On this account, it is trivially true that, in non-disablist worlds, disability would not be detrimental to well-being.

Strong versions of the Medical Model—which understand disability as an intrinsic feature of malfunctioning bodies (an impairment) and imply that the limitations associated with disability result *entirely* from impairments—are also widely rejected.<sup>36</sup> The second attraction of the Solidarity Account is that it leaves open the possibility that disability is both social and medical in nature. This is best, because it allows questions to be asked about how both intrinsic features of bodies (impairments) and social factors (oppression) affect disabled people's well-being, the answers to which should not be entailed by our understanding of disability.

A final attraction of the Solidarity Account is dialectical. It is favored by Barnes, whose views I criticize here.<sup>37</sup> As such, it is best for me to argue on her terms.

Note that I am not arguing that the Solidarity Account is the best understanding of disability per se, only that it is a good one to use in this context. As such, I will not discuss objections to it.<sup>38</sup> I also do not think it matters much whether the reader endorses the Solidarity Account, as my discussion does not rely on it. I focus on paradigm cases of disability, so my discussion will be relevant to anyone who endorses an account that (i) implies that paradigm cases of disability are disabilities, and (ii) does not preclude meaningful investigation into relevant issues.

33 Barnes, *The Minority Body*, 2–3.

34 Kahane and Savulescu, "The Welfarist Account of Disability."

35 Shakespeare and Watson, "The Social Model of Disability," 10.

36 Wasserman, Asch, Blustein, and Putnam, "Disability."

37 Barnes, *The Minority Body*.

38 See, e.g., Lim, "Disabilities Are Also Legitimately Medically Interesting Constraints on Legitimate Interests," 982–87.

Many accounts meet these conditions, including the Nordic Relational Model, Shakespeare's Interactionist Account, and the World Health Organization's understanding.<sup>39</sup> This is because one goal of most accounts of disability is to cohere with our pre-theoretic intuitions about which conditions count as disabilities. They aim to imply that paradigm disabilities are disabilities and that paradigm non-disabilities are not.<sup>40</sup> Thus, the extensions of the concepts produced by most accounts of disability overlap considerably; they each identify very similar sets of conditions, even if they understand disability differently. And where they do come apart, this will tend to be on borderline or indeterminate cases. So, if I stick to discussing paradigm cases of disability (and do not perform conceptual analysis on the concept of *disability* offered by the Solidarity Account), then my discussion should be able to cut across others that utilize accounts that also satisfy i and ii. By the same logic, the views of those whose accounts of disability satisfy i and ii will be relevant to my discussion, even if they do not endorse the Solidarity Account, as they will be discussing the right set of conditions. The conditions are key here, as opposed to the theoretical nature of disability.

I understand well-being as the non-instrumental value a life has for the person whose life it is. Well-being is the kind of value we have in mind when thinking about harms and benefits for a person. To harm someone is to negatively impact their well-being, and to benefit someone is to positively impact it.<sup>41</sup> I will not endorse any theory of well-being; my discussion aims to be neutral between the major philosophical theories. These are Hedonism, the Desire Theory, and the Objective List Theory.<sup>42</sup> Hedonism claims that well-being consists in experiencing the largest net sum of pleasure minus pain.<sup>43</sup> According to the Desire Theory, well-being consists in getting what one non-instrumentally desires and ill-being consists in having one's non-instrumental desires frustrated.<sup>44</sup> And Objective List Theories hold that well-being consists in the attainment of objectively valuable things.<sup>45</sup> Where necessary, I will discuss the implications that endorsing different theories of well-being might have on my arguments.

39 Gustavsson, "The Role of Theory in Disability Research"; Shakespeare, *Disability Rights and Wrongs Revisited*, 74–84; World Health Organization, "International Classification of Functioning, Disability and Health."

40 This is certainly an aim of Barnes's account (*The Minority Body*, 10–11).

41 For more discussion on the nature of well-being, see Crisp, "Well-Being."

42 Parfit, *Reasons and Persons*, 491–503.

43 Parfit, *Reasons and Persons*, 493.

44 Heathwood, "Desire Satisfactionism and Hedonism," 541.

45 Parfit, *Reasons and Persons*, 499.



## 3. THE BEST INTERPRETATION

In this section, I argue that 6\* is the best interpretation of the BDV. How do we identify the BDV's best interpretation? Given that we are attempting to characterize a view endorsed by various thinkers, our first desideratum is *charity*. Charity requires that we assume that BDV proponents are reasonable and rational. So, it requires that the BDV is interpreted in its most plausible or rational form. This means that (if possible) we must interpret the BDV as being *prima facie* internally coherent and consistent with facts not in dispute. Charity also requires (*inter alia*) that the BDV be interpreted, as far as is possible, as consistent with what BDV proponents actually say.

Charity must be balanced with another desideratum: *appropriate normative relevance*. Thinkers engaging in this debate typically understand the BDV to have normative implications. It is often taken to imply things about the (im)permissibility of selecting for or against disability via reproductive technologies.<sup>46</sup> And some thinkers take it to imply certain things about distributive justice—specifically the fair distribution of health-care provision and social support.<sup>47</sup> For example, quality adjusted life years (QALYS) are commonly taken to be an appropriate metric for health-based utility, which is used to determine the just division of health-care resources.<sup>48</sup> The idea is that we should distribute resources such that they produce the most utility. QALYS take the additional number of years a health-care allocation will produce, together with the quality of those years, to determine its utility. Clearly, then, the quality of disabled people's lives—that is, whether the BDV is true—affects the number of QALYS produced by allocations that result in additional years with a disability, and thus which health-care allocations are viewed as just.<sup>49</sup> The best interpretation of the BDV must make sense of the moves being made in the debate. That is, it must adequately explain why the BDV is (at least *prima facie*) relevant to these normative issues. If it cannot, then it has changed the subject and cannot be properly called “the BDV.” (This second desideratum might be thought of as an element of charity, given that it seems charitable to assume that BDV proponents are not mistaken in taking their view to be relevant to these things.)

46 For examples on both sides of this debate, see McMahan, “Causing Disabled People to Exist and Causing People to Be Disabled”; Barker and Wilson, “Well-Being, Disability, and Choosing Children.”

47 E.g., Singer, McKie, Kuhse, and Richardson, “Double Jeopardy and the Use of QALYS in Health Care Allocation.”

48 Franklin, “Calibrating QALYS to Respect Equality of Persons,” 65.

49 Cf. Bickenbach, “Disability and Health Care Rationing.”

The charity desideratum does not imply that the BDV must turn out, after more work, to be internally coherent and consistent with the facts. After we identify the best interpretation and scrutinize it, it might turn out to be false, incoherent, or unsupportable. But it would be bad philosophical practice to *start out* by characterizing the BDV as such, if it is possible not to. Nor does the appropriate relevance desideratum require that the BDV be interpreted such that it has any of the specific normative implications sometimes argued for by BDV proponents (e.g., that disabled people have less claim to scarce medical resources). Even if the BDV is true, whether these claims are also true is an open question. The point is that the BDV should be interpreted such that it appears at least *prima facie* relevant to the normative issues, understood generally, to which people take it to be relevant: it should be *prima facie* relevant to selection debates, rather than implying anything specific about the normative status of selecting for a given disability.

In sum, the two desiderata I will use for identifying the best interpretation of the BDV are charity and appropriate normative relevance. These desiderata must be balanced: loss in one might be acceptable if necessary for gain in the other. The best interpretation will be the one that best balances charity and relevance.

I will now evaluate 1\*–6\* with these desiderata in mind, arguing that 6\* is the best interpretation. In section 3.1, I reject non-probabilistic interpretations as uncharitable, and in section 3.2, I reject *pro tanto* interpretations as not as normatively relevant as overall ones. Thus, some candidates are criticized on multiple grounds.

### 3.1. Rejecting Non-probabilistic Interpretations

Interpretations 1\*, 3\*, and 5\* are all non-probabilistic; they each state that disability is *always* bad for well-being in some respect. First, I will argue that 1\*—the view that disability is always *non-instrumentally* bad—is an uncharitable interpretation.

Whether one thinks that disability is always non-instrumentally bad will depend on their theory of well-being. Hedonism says that painful mental states are the only non-instrumental bads and pleasurable mental states are the only non-instrumental goods. The hedonist will thus endorse 1\* only if they think that disability always contains (as opposed to causes) less net pleasure than non-disability.<sup>50</sup> But there appears to be no good motivation for this belief.

What people enjoy depends on their unique psychologies. As such, disability itself might be experienced as net pleasurable, painful, or neutral, depending on a person's psychology. Some disabilities are intrinsically painful (e.g., rheuma-

<sup>50</sup> I take net pleasure to be pleasurable mental states minus painful mental states, adjusted for intensity.

toid arthritis).<sup>51</sup> However, many paradigm disabilities—blindness, paraplegia, etc.—need not involve pain. Plausibly, these disabilities are sometimes experienced as net pleasurable; this seems especially likely for disabilities that involve a completely different way of experiencing the world. Some blind people might take pleasure in their unique sensory experience and some (totally) deaf people might take pleasure in the silence involved in their condition, or other elements of it—this seems to be part of the thought behind the Deaf Gain movement. Even more plausibly, non-painful disabilities are at least sometimes experienced as net neutral. This would happen whenever they contain no pleasurable or painful mental states or when these states balance. A good candidate for a non-instrumentally neutral disability on Hedonism is achondroplasia, or short-limbed dwarfism, which seems to contain no mental states at all.

Likewise, non-disability might be experienced as net pleasurable, painful, or neutral, depending on a person's psychology. Perhaps non-disability is rarely experienced as painful. But it appears difficult to deny that at least some people experience non-disability as non-instrumentally neutral. Many people take pleasure in non-disability's *instrumental* benefits—perhaps they enjoy mountain climbing or watching sunsets—but it seems likely that at least some people do not experience non-disability *itself* as pleasurable, such that it would improve their well-being even if it had no instrumental benefits. If *anyone* experiences non-disability as non-instrumentally neutral (or bad), then 1\* is false on Hedonism on the (very plausible) assumption that disability is at least sometimes experienced as neutral or net pleasurable.<sup>52</sup> Thus, it seems implausibly strong to say that disabilities *always* involve less net pleasure than non-disability.

The desire theorist will endorse 1\* only if they believe that disability *always* involves less aggregate desire satisfaction than non-disability.<sup>53</sup> But there appears to be no good motivation for this belief either. What people desire varies considerably. Thus, for some people, disability might involve large amounts of desire satisfaction, if they, e.g., non-instrumentally desire to live in a world of calming silence or to have an amputated limb, as those with body integrity identity disorder often do.<sup>54</sup> Moreover, it seems too strong to say that disability always frustrates desires. Perhaps disability would often frustrate the desires of non-disabled people. But people often adapt their preferences to suit their

51 See the National Center on Health, Physical Activity and Disability's rheumatoid arthritis "Primer on Pain," <https://www.nchpad.org/281/1789/primer~on~pain>.

52 Cf. Schramme, "Disability (Not) as a Harmful Condition," 82.

53 I take aggregate desire satisfaction to be non-instrumental desire satisfaction minus non-instrumental desire frustration, adjusted for desire intensity.

54 Bayne and Levy, "Amputees by Choice."

capabilities—so-called adaptive preference—which implies that many disabled people might not possess desires not to have a disability or to do things that their disability prevents them from doing.<sup>55</sup> As such, the Desire Theory does not appear to license the view that disability is *always* non-instrumentally worse for well-being than non-disability.<sup>56</sup>

Besides, it is simply not in the spirit of the Desire Theory to endorse non-probabilistic claims like  $1^*$ . The point of the Desire Theory is that well-being is dependent on pro-attitudes, which vary. The nature of different disabilities also varies considerably. Desire theorists are thus unlikely to think that all (or even the vast majority of) the disparate conditions labeled “disabilities” (from deafness to spina bifida to muscular dystrophy) are uniformly non-instrumentally bad for well-being. This would fly in the face of a primary thought behind the Desire Theory.

Perhaps  $1^*$  is more plausible on *idealized versions* of the Desire Theory—on which well-being consists in getting what an idealized version of you would want.<sup>57</sup> Perhaps, although disability does not always involve less aggregate satisfaction of actual desires, it always involves less aggregate satisfaction of idealized desires. Whether this is true depends on how we understand idealization. Without getting too far into the options, an important point is that the Desire Theory is a subjective theory of well-being—which is to say that, on the Desire Theory, the particular things that are good for you are not always good for me, and vice versa. If our method of idealization maintains this feature, then it is still unlikely that disabilities will involve less aggregate desire satisfaction for every idealized agent, because even idealized agents will have idiosyncratic desires. If, however, our chosen method of idealization implies that all idealized agents possess the same set of desires then, on this version of the Desire Theory,  $1^*$  might be true. But I would contend that this theory is not a Desire Theory at all. It would, I think, be an objective view in subjective clothing. If so, then what I have to say about Objective List Theories is relevant.

It is more plausible that objective list theorists would endorse  $1^*$  than hedonists or desire theorists because, on Objective List Theories, what well-being consists in is the same for everyone and not dependent on experiences or attitudes.<sup>58</sup> On this view, it might be that disability always involves more objective bad and/or less objective good than non-disability. But it is entirely unclear that

55 I am not assuming that adapted preferences are undesirable or suboptimal; cf. Barnes, “Disability and Adaptive Preference.”

56 Cf. Schramme, “Disability (Not) as a Harmful Condition,” 84.

57 Heathwood, “Subjective Theories of Well-Being,” 212.

58 That is, unless the list includes pleasure or getting what one wants. I will ignore this issue as

this is true. No objective list (that I know of) includes non-disability. And the presence or absence of many putative objective goods—knowledge, achievement, friendship, etc.—is not contained within disability or non-disability themselves. One might think that pain is objectively bad, and that disability always involves more pain than non-disability, but I have rejected this. The other way that  $i^*$  might be true on an Objective List Theory (that I will consider) involves appeal to ill health. Some Objective List Theories maintain that health is non-instrumentally good, and one might think that disability entails ill health, whereas non-disability does not itself contain any objective bads, so disability is always non-instrumentally worse than non-disability.<sup>59</sup> This thought is more promising, as it is plausible, on some conceptions of health and disability, that disabled people are necessarily unhealthy.

I will not define “health” here, as this is beyond my scope, but two things are worth noting. First, disabled people commonly reject the identification of ill health and disability.<sup>60</sup> Second, on some accounts of health—such as Carel’s, on which to be healthy is, roughly, to feel in harmony with one’s body—disabled people need not be unhealthy.<sup>61</sup> However, one might endorse an account on which disability does entail ill health. On that view,  $i^*$  could be true.

So,  $i^*$  is compatible with certain Objective List Theories. Notice, though, that for  $i^*$  to be compatible with a person’s views they must believe: (i) that an Objective List Theory is true; (ii) that health is an objective good; (iii) that disabled people are always unhealthy; (iv) that disabilities *never* involve enough counterbalancing non-instrumental good(s) to make them non-instrumentally better than non-disability; and (v) that non-disability *never* involves enough non-instrumental bad(s) to make it non-instrumentally worse than disability.<sup>62</sup>

Interpretation  $i^*$  would therefore only be endorsed by the (presumably small) set of people who satisfy i–v. Thus it is an uncharitable understanding of the BDV. All of i–v are contentious. So, characterizing the BDV as  $i^*$  increases the contentiousness of the presuppositions of BDV proponents. It is more charitable, other things equal, to characterize the BDV as having less contentious presuppo-

---

I have argued that neither consideration of pleasure nor desire satisfaction provides reasonable grounds for endorsing  $i^*$ .

59 E.g., Finnis, *Natural Law and Natural Rights*, 86.

60 Wendell, “Unhealthy Disabled.”

61 Carel, *Illness*.

62 Some theories of health might imply that ill health is non-instrumentally bad on subjective theories of well-being. Nordenfelt defines ill health as a state in which your body frustrates your vital goals, which are defined relative to preferences (*On the Nature of Health*). On this account, ill health might be non-instrumentally bad on Desire Theories. However, on these accounts, it is also unlikely that disability entails ill health.

sitions, such that it relies on none of *i-v*.<sup>63</sup> In particular, it is more charitable to characterize it as being consistent with a larger proportion of the major theories of well-being. And it is possible to do this, as I will argue below. As such, *i\** is not a charitable understanding of the BDV.

Understanding the BDV to be *i\** is also an uncharitable for another reason. Recall, an important element of charity is that we must, as far as possible, interpret views as being consistent with what their proponents say, and many BDV proponents—although they do not precisely cash out their view—say things inconsistent with *i\**. Singer says:

*I don't hold that anyone with a disability "will be necessarily disadvantaged."* That would be an absurd claim. In unusual circumstances—for example, when all able-bodied people are conscripted to fight in a dangerous war—having a disability may be an advantage. I would argue only that, other things being equal, being able to walk, to move one's arms, to hear, to see, to recognize other people and communicate with them, are advantages.<sup>64</sup>

Here, Singer elaborates on his view that “other things equal, it is better not to be disabled.”<sup>65</sup> His elaboration indicates that his version of the BDV does not consider only non-instrumental value; he considers disability's instrumental value too, such as that of preventing conscription. It also shows that Singer's version of the BDV is probabilistic: “in unusual cases . . . disability may be an advantage.” It is certain, then, that Singer's version of the BDV is not *i\**.

McMahan says that

even if the abilities whose absence is constitutive of disability are good only instrumentally. . . . The lack of an ability that is instrumentally valuable to those who have it is, in general, an obstacle to the achievement of the full range of goods characteristic of human life. . . . I believe, moreover, that the value of certain abilities . . . is only partly instrumental. The possession and exercise of certain . . . capacities is intrinsically good.<sup>66</sup>

Here, McMahan also does not appear interested in only disability's non-instru-

63 If someone thought that disability always involves some objective bad other than ill health, then their reasoning would be slightly different but no less controversial. And I do not know what this objective bad might be.

64 Singer, “Ethics and Disability,” 130.

65 Singer, “Ethics and Disability,” 133.

66 McMahan, “Causing Disabled People to Exist and Causing People to Be Disabled,” 96. McMahan does not explicitly endorse the BDV here, but he strongly suggests that he does earlier in the paper (96), and I take this to be a clarification of his view.

mental value. He says that disabled people lack instrumentally valuable abilities because of their disability, which is equivalent to saying that disability is instrumentally bad. If instrumental value is relevant to his discussion, this suggests that McMahan's version of the BDV also is not 1\*.

In sum, 1\* is an uncharitable interpretation of the BDV because it is incompatible with various major theories of well-being and inconsistent with the expressed views of BDV proponents.

Having rejected 1\*, I will now briefly reject the other non-probabilistic interpretations, for similar reasons. According to 3\*, disability is always instrumentally bad. Now, disability often is instrumentally bad.<sup>67</sup> This is, first, because it sometimes causes non-instrumental bads. Many bads caused by disability are presumably due to disablism. However, others are not; for instance, the pain and discomfort caused by some disabilities would exist in virtually all social environments.<sup>68</sup> People also sometimes feel distressed because their disability prevents them from doing certain things, especially if they acquire a disability and must adjust to their new condition, which might include "transition costs" that often accompany coming to terms with being unable to engage in activities important to them or to pursue their goals that involved the use of abilities they no longer have.<sup>69</sup> Disability can be instrumentally bad, second, by preventing goods. For example, while deafness is likely not bad non-instrumentally, it can prevent one from communicating effectively, which might prevent one from attaining as much friendship as one could have if not deaf; disabilities that shorten life span (e.g., cystic fibrosis) can prevent one from attaining goods (e.g., pleasure, achievement) that they would have if they lived longer, and so on.<sup>70</sup>

However, disability also can be instrumentally good, as the first-person testimony of disabled people often points out.<sup>71</sup> For example, disability might cause

67 Campbell and Stramondo, "The Complicated Relationship of Disability and Well-Being."

68 Shakespeare, *Disability Rights and Wrongs Revisited*, 75.

69 Barnes, *The Minority Body*, 148; McMahan, "Causing Disabled People to Exist and Causing People to Be Disabled," 95.

70 Campbell and Stramondo, "The Complicated Relationship of Disability and Well-Being," 161. It is difficult to give uncontroversial examples of instrumental and non-instrumental value. Whether disability is understood as non-instrumentally or instrumentally bad depends on our theory of well-being and on whether we think certain things (e.g., pain, shortened life span) are intrinsic features or causal consequences of disability. For example, even if a shortened life is an intrinsic feature of cystic fibrosis, the hedonist will take this to be an instrumental bad, as it is not a painful mental state but plausibly causes one. I will not try to settle these issues here. Throughout, I try to give plausible examples of instrumental and non-instrumental value that will be acceptable to many, but unfortunately not all. However, the general points apply whether or not one agrees with the characterization of values in the examples.

71 See, for example, Barnes, *The Minority Body*, 119–43; Schramme, "Disability (Not) as a



a person to have a more positive attitude toward life and enjoy it more; it might help one find friends or mature as a person; it might prevent someone from being conscripted, or allow access to certain disability-specific goods like access to the disabled community, being able to work peacefully in loud environments, or being permitted to avoid long queues.<sup>72</sup>

The point is that disability can be instrumentally good and bad, and there is no evidence that its instrumental badness *always* outweighs its instrumental goodness. In fact, it is implausible that it does. Just one case where disability causes more goodness than badness would falsify 3\*, and it seems probable that there are many such cases, given the number of disabled people and the multitude of potential instrumental benefits of disability. Given this, the view that it is *always* instrumentally bad is implausible.

This speaks to a general problem with non-probabilistic versions of the BDV: they make extremely strong claims that are implausible and easily refuted. A *single case* where a disabled life is *not* instrumentally, non-instrumentally, or overall worse than a relevantly similar non-disabled life would be all that is needed to falsify non-probabilistic versions of the BDV, and it seems likely that there are many such cases, given that disability is not always non-instrumentally bad (at least on many theories of well-being) and can have instrumental benefits.

Moreover, any reasonable person would make room in their theory for the *mere possibility* that someone could do better with a disability by, for example, avoiding conscription. And we have textual evidence that BDV proponents do exactly this. Recall, Singer made clear that disability can sometimes be an advantage.<sup>73</sup> Likewise, Glover says that “disabilities . . . often (though not always) mean that people have less good lives.”<sup>74</sup> And Andric and Wundisch say, “of course, it is not true in *all* cases that persons are better off if they [are non-disabled] . . . , however, we claim that this will *typically* be the case.”<sup>75</sup> This shows that these prominent BDV proponents do not endorse non-probabilistic versions of the BDV.

In sum, then, non-probabilistic versions of the BDV are implausibly strong and inconsistent with the expressed views of many BDV proponents. This provides reason to reject 1\*, 3\*, and 5\* as uncharitable.

---

Harmful Condition,” 72; Eyre, “A Few Awesome Things about Being Disabled”; and Steer, “I’m in a Wheelchair.”

72 Barnes, *The Minority Body*, 116; Schroeder, “Well-Being, Opportunity, and Selecting for Disability,” 15; Campbell and Stramondo, “The Complicated Relationship of Disability and Well-Being,” 158.

73 Singer, “Ethics and Disability,” 130.

74 Glover, *Choosing Children*, 1.

75 Andric and Wundisch, “Is It Bad to Be Disabled?” 16.

### 3.2. *Rejecting Pro Tanto Interpretations*

*Pro tanto* interpretations of the BDV say that disability is bad for well-being *in one way*, either instrumentally or non-instrumentally. These contrast with all-things-considered interpretations. In this section, I reject *pro tanto* versions of the BDV.

Suppose that a recluse *A* acquires a painful disability *D* that causes her to become an active member of the disabled community. *A*'s disability, although non-instrumentally bad, is instrumentally good as it brings her friendship and pleasure. Suppose further that *B* acquires a disability *D\** that involves no non-instrumental bad but causes him to become depressed due to changes in his self-conception: *B*'s non-instrumentally neutral disability is instrumentally bad. Suppose finally that people who have *D* reliably experience instrumental benefits and that people who have *D\** reliably experience instrumental harms, in ways similar to *A* and *B*. Now, it would be misleading to say that *D\** is better for well-being than *D* merely because *D\** is *non-instrumentally* better. This claim—although true on one reading of “better”—might lead someone to think that priority should be given to allocating resources to fund treatments for *D* before *D\**, that it is worse to select for *D* than *D\**, etc. But these normative conclusions might be erroneous, because focusing on one kind of value can obscure important things.

The point is that the BDV would not be as relevant to normative issues if it considered only *pro tanto* value. Doing so can provide a misleading picture, as certain non-instrumentally bad disabilities (e.g., rheumatoid arthritis) might correlate with instrumental goodness or neutrality, while certain non-instrumentally neutral disabilities (e.g., quadriplegia) might correlate with instrumental badness. More generally, something can be non-instrumentally bad but, all things considered, good, or instrumentally good but, all things considered, bad.

There appears to be no good reason for the BDV to focus on *pro tanto* value when considering overall value offers a more comprehensive picture of disability's effect on well-being. Consider the issues to which the BDV is taken to be relevant. If a BDV proponent wanted to use the BDV to argue that selecting for disability is impermissible (as McMahan seems to), then surely both instrumental and non-instrumental value are relevant.<sup>76</sup> That a condition is non-instrumentally or instrumentally valuable, disvaluable, or neutral is not all that matters for whether we ought to select for or against it. If some non-instrumentally neutral disability *F* were reliably correlated with large amounts of instrumental badness, then this appears to provide a defeasible reason to select against *F*. Likewise, if some mildly non-instrumentally bad disability *G* were reliably correlated with

76 McMahan, “Causing Disabled People to Exist and Causing People to Be Disabled.”

large amounts of instrumental goodness, then there would be no good reason to select against *G*—there may even be good reason to select *for G*. What is relevant to the selection debate, and other relevant debates, is overall value, not instrumental or non-instrumental value *alone*. Focusing on *pro tanto* value thus hinders the BDV's relevance to appropriate normative issues.

In sum, considering overall value provides a more full picture of disability's effect on well-being that is more relevant to appropriate normative issues, and views that consider only one kind of value do not license the normative conclusions argued for by BDV proponents. Therefore, all candidate interpretations that merely consider *pro tanto* value should be rejected. So, we should reject 1\*, 2\*, 3\*, and 4\* for this reason, as well as those mentioned above.

### 3.3. Typically All-Things-Considered Bad

I have argued that 1\*, 3\*, and 5\* are uncharitable interpretations and that all of 1\*–4\* should be rejected, as *pro tanto* interpretations hinder the BDV's appropriate normative relevance. One candidate remains:

6\*. Disability is typically all-things-considered bad for well-being.

Elaborating and reintroducing clauses suppressed above:

BDV: Discounting the effects of disablism, a person is (*ceteris paribus*) typically (likely to be) overall worse off with a disability (in virtue of their disability) over the course of their life than they would have been without.

This view says that, discounting disablism, a disabled person is (other things equal) likely to have lower lifetime well-being (in virtue of their disability) than they would have done if they were non-disabled. It follows that, discounting disablism, a non-disabled person is (other things equal) likely to have higher lifetime well-being than they would have done (in virtue of their disability) were they disabled.

Regarding the desiderata, 6\* is a charitable interpretation because disability is *sometimes* non-instrumentally bad and *sometimes* instrumentally bad, and it is *prima facie* plausible that these bads often, *but not always*, make disability worse for well-being than non-disability. Moreover, 6\* is not committed to any of i–v. It might be true on any major theory of well-being; it does not imply that disability *always* involves or causes non-instrumental bads; it is compatible with some cases of disability being non-instrumentally or instrumentally good or neutral; and it allows that disabled lives can sometimes be better than relevantly similar non-disabled lives. In short, 6\* makes no blanket claims about the quality of disabled people's lives, so is compatible with the heterogeneous nature of

disability and the complicated nature of the relationship between disability and well-being. It is also consistent with the quotations from BDV advocates, above. As such,  $\delta^*$  is a more charitable interpretation than all non-probabilistic interpretations. Moreover, considering overall value increases the BDV's relevance to appropriate normative issues, as this gives a more comprehensive picture of disability's impact on well-being. Given this,  $\delta^*$  is more normatively relevant than all non-comparative views.

That concludes my argument that  $\delta^*$  is the best interpretation of the BDV, as far as instrumentality and probability go. In light of this, the MDV should be understood as:

MDV: Discounting the effects of disablism, a person is not (*ceteris paribus*) typically (likely to be) overall worse off or better off with a disability (in virtue of their disability) over the course of their life than they would have been without.

To discount disablism is to consider only harms (and benefits) of disability (and non-disability) that are *not* caused by unjust discrimination against disabled people. One way of doing so is to consider well-being levels in the nearest non-disablist worlds.<sup>77</sup> Thus, to responsibly endorse either view, we must form an idea of what the nearest non-disablist worlds are like. A great help in this enterprise would be to specify what counts as disablism. But doing so this is not easy, as disablism cannot be understood as anything that lowers the well-being of disabled people, for this would make the MDV trivially true. Determining what disablism is would allow us to have a reasonable idea of what lives would be like in the closest non-disablist worlds. We would need to consider this information carefully to decide whether disabled people are typically overall worse off in these worlds (more on this below).

It is worth briefly noting the differences between my arguments and Campbell and Stramondo's.<sup>78</sup> They end up suggesting that disability seems to be typically overall bad for well-being. However, their discussion differs from mine in important respects. First, they discuss whether candidate versions of the BDV are *true*, whereas I aim to identify the best version without considering its truth. Second, I have improved upon their taxonomy in various ways (see section 1), which has led me to understanding all candidates as comparative, improving their labeling, and providing the first consideration of probabilistic and non-probabilistic versions of each candidate. Third, they do not discuss the normative relevance of the candidate interpretations. And fourth, they evaluate

77 Cf. Barnes, *The Minority Body*.

78 Campbell and Stramondo, "The Complicated Relationship of Disability and Well-Being."

whether the views are true *in the actual world*, whereas my discussion discounts disablism. This is significant, as nonideal social conditions are one of their primary reasons for suggesting that disability is typically overall bad in the actual world. Thus, it is unclear whether Campbell and Stramondo would think that  $\delta^*$ —which discounts disablism—is true.

I choose to discount disablism, first, because this is more faithful to the origins of this debate. It originated in the views of disability rights activists and those who endorse the Social Model; and both groups hold that disability would not be bad for well-being *absent prejudice*.<sup>79</sup> Second, because it is obvious that disabled lives are typically worse in disablist worlds. The more interesting question is whether disablism is *all* that makes them worse. Those who want to reject  $\delta^*$  will have to argue as much. Thus, the view that Campbell and Stramondo suggest is true makes a significantly weaker and less interesting claim than  $\delta^*$ , which I have identified as the BDV's best interpretation.

#### 4. BARNES'S ATTACK ON THE BDV

I have argued that  $\delta^*$  is the BDV's best interpretation. In this section, I discuss an important implication of my argument. I first show that Barnes understands the BDV as something like  $\iota^*$ .<sup>80</sup> Then, I draw on my arguments in section 3 to show that, in doing so, she attacks an uncharitable interpretation unlikely to be endorsed by many BDV proponents.

Two considerations suggest that Barnes's characterization of the BDV is akin to  $\iota^*$ . The first is her language. In various places, she characterizes the BDV as the view that disability *by itself*, *intrinsically*, or *automatically* makes one worse off.<sup>81</sup> Now, the word "automatically" straightforwardly suggests a non-probabilistic characterization—if something automatically accompanies disability, then it occurs in every case of disability. And her understanding the BDV as the view that disability "by itself" or "intrinsically" makes one worse off suggests that Barnes's interpretation focuses on disability's non-instrumental (or intrinsic) value and ignores its instrumental value.

The second, more important, consideration is her argument for the MDV (and against the BDV). Her version of the MDV is the Value-Neutral Model, which she defends by arguing that disability is "neutral simpliciter." This is to say that disability is *not* bad or good *simpliciter*. Barnes thinks that the MDV is the conjunction of the denials of the BDV and the Good-Difference View (GDV), so we

79 E.g., Oliver, *Understanding Disability*.

80 Barnes, "Disability, Minority, and Difference" and *The Minority Body*.

81 See, e.g., Barnes, "Disability, Minority, and Difference," 338, and *The Minority Body*, 6, 55.

can infer that she takes the BDV to be the view that disability is bad *simpliciter* and the GDV to be the view that disability is good *simpliciter*.<sup>82</sup> Examination of Barnes's thoughts on what it is to be bad *simpliciter* will thus illuminate how she understands the BDV.

According to Barnes, some feature *F* is bad *simpliciter* iff for any person *P* who has *F*, *P* has a lower level of well-being in virtue of having *F* than they would have had if they lacked *F*.<sup>83</sup> Notice that for *F* to be bad *simpliciter* is for it to be bad for *any person who has it*, and something is only bad for any person who has it if it is bad non-probabilistically. This indicates that Barnes characterizes the BDV as the non-probabilistic view that disability is always bad for well-being.

Barnes elaborates on what it means to be bad *simpliciter*: "if something is bad *simpliciter*, your life goes worse *in virtue of it specifically*, even if its overall causal effects ... make you better off."<sup>84</sup> Barnes does not unpick what she means by "in virtue of it specifically." However, it is illuminating that she contrasts effects in virtue of *F* specifically with *F*'s causal (or instrumental) effects. Value must either be instrumental or non-instrumental, so if the *simpliciter* value of *F* discounts its instrumental value, then this suggests that it is constituted entirely by its non-instrumental value. It seems, then, that what Barnes means by "*F* is bad *simpliciter*" is that *F* is *non-instrumentally* bad for well-being.

Putting these points together, we can see that Barnes's characterization of the BDV (that disability is bad *simpliciter*) seems akin to *1\** and that her argument that disability is neutral *simpliciter* (that the MDV is true) seems to be an argument that disability is not *always non-instrumentally* bad (or good) for well-being (minus the effects of disablism).<sup>85</sup> Less strongly, her argument for the MDV is only effective against versions of the BDV that claim that disability is always non-instrumentally bad, as arguing that disability is neutral *simpliciter* would not falsify probabilistic versions or those that focus on instrumental or overall value.

If my interpretation of Barnes is correct, then Barnes characterizes the BDV as something akin to *1\**: the view that disability is always non-instrumentally bad for well-being. But as I argued in section 3, *1\** is a bad understanding of the BDV. First, because *1\** focuses on non-instrumental value, which hinders its normative relevance. Second, because *1\** makes an implausibly strong claim. In fact, Barnes acknowledges that the way she is using "bad *simpliciter*" is very strong.<sup>86</sup>

82 Barnes, *The Minority Body*, 69.

83 Barnes, *The Minority Body*, 86.

84 Barnes, *The Minority Body*, 87.

85 This is curious, as she mentions that there are various potential understandings of the BDV, only some of which focus on non-instrumental value (*The Minority Body*, 54–77).

86 Barnes, *The Minority Body*, 87; cf. Hawkins, review of *The Minority Body*, 465.

And, third, because many BDV proponents make clear that their views are not properly expressed by 1\*.

If my arguments are correct, then, in rejecting a version of the BDV akin to 1\*, Barnes does not attack the BDV's best interpretation.<sup>87</sup> She attacks an uncharitable understanding that is unlikely to be endorsed by many BDV proponents, and she misses the opportunity to challenge the most plausible and relevant version, which is 6\*. The crucial upshots of this are (i) that one can be persuaded by Barnes's arguments on this topic and still hold a plausible version of the BDV, and (ii) that the best version of the BDV seems to remain unchallenged.<sup>88</sup>

Now, it is possible that Barnes's arguments are designed to reject Welfarist Accounts on which disability is, *by definition*, bad for well-being.<sup>89</sup> However, if this were her aim, it seems likely that she would have made this clear rather than suggesting that her aim is the broader one of defending the MDV, and that she would have devoted significant time to rejecting Welfarist Accounts, rather than merely stating that they are not useful in the relevant context.<sup>90</sup> So, it is hard to see a good reason for thinking that her arguments are designed to reject Welfarist Accounts. It is also possible that her arguments are designed to reject folk conceptions of disability, which might be akin to 1\*. But, insofar as Barnes's arguments are designed to challenge the views of other *philosophers* (who do not endorse Welfarist Accounts), it seems that she misses the target by attacking an uncharitable interpretation of the BDV that is unlikely endorsed by many.

##### 5. BARNES'S OBJECTION

I have argued that Barnes misses the target by attacking 1\*. Barnes in fact acknowledges that the BDV may be probabilistic:

Perhaps what we commonly think ... is not that having a disability *will* make a person worse off ... but rather that having a disability *will likely* make a person worse off.<sup>91</sup>

87 Barnes, "Disability, Minority, and Difference" and *The Minority Body*.

88 Barnes coined the term "BDV," so there is a sense in which the BDV just is the view that she attacks. However, I think that any view that postulates a negative connection between disability and well-being could be appropriately described as a BDV, and Barnes seems to agree (see esp. *The Minority Body*, ch. 2). On this understanding, Barnes might not attack the best interpretation.

89 E.g., Kahane and Savulescu, "The Welfarist Account of Disability."

90 Barnes, *The Minority Body*, 12.

91 Barnes, "Disability, Minority, and Difference," 343.



However, she objects that we should not interpret the BDV as a probabilistic variant like  $\delta^*$  because there is no appropriate way of making non-question-begging judgments about whether  $\delta^*$  is true. She believes that the relevant notion of probability in  $\delta^*$  is *objective chance for a given person* of doing worse with a disability than without. Objective chance is distinguished from subjective chance. One can say that a fair coin has the objective chance of 0.5 of landing heads, whereas subjective chance is concerned with what agents are justified in believing given their evidence.<sup>92</sup> Barnes argues that subjective chance—determined by looking at averages across populations—merely tells us about average well-being and nothing about whether a given individual is objectively likely to be worse off with a disability. She thinks that to make judgments about objective chances we need to make question-begging assumptions about the non-instrumental value of disability:

The relevant notion is objective chance *for x* at high quality of life. . . . It's not enough to simply determine the average quality of life of persons in similar circumstances with similar disability and compare it to persons in similar circumstances without disability. . . . We cannot tell what will happen to a particular person just by calculating averages. So unless we *assume* that disability is somehow intrinsically negative (which begs the question) we cannot make inferences about that person's chances at an overall high quality of life based solely on the presence of a disability.<sup>93</sup>

I will now reply to Barnes. It is difficult to understand why she thinks that the BDV must be interpreted in terms of objective chance. Perhaps the thought is that the badness of disability must be a property of the person themselves. But it is not clear why she would insist on this. Also note that some are skeptical that objective chances exist, because objective chances (other than 0 or 1) appear incompatible with determinism.<sup>94</sup> Given this, it seems too quick to stipulate that objective chance is the relevant notion.

Putting aside these worries and allowing (for argument's sake) that objective chance is the relevant notion, I still do not think it follows that we cannot make reasonable judgments about  $\delta^*$  without begging the question. Consider the implications of Barnes's reply. She suggests that we cannot make reasonable judgments about a particular individual's objective chance of  $x$  being good or bad for her based on information about averages. But we often (perhaps always) do not have epistemic access to objective chances. If Barnes is right, this would

92 Hájek, "Interpretations of Probability."

93 Barnes, "Disability, Minority, and Difference," 343.

94 Bradley, "Are Objective Chances Compatible with Determinism?"

license radical skepticism about what we can reasonably believe. Multiple judgments based on expected utility would be unreasonable. We would (almost?) never be able to form reasonable beliefs about what is likely to be good or bad for an individual. But it seems clear that we can, and often do, form such reasonable beliefs, and we *at least sometimes* form them based on information about averages. And this is true even if we allow that objective chance is the relevant notion. When we do not have epistemic access to objective chances, we can use subjective chances to form reasonable beliefs *about objective chances*; we can use subjective chances to form reasonable beliefs about what is objectively likely to be good or bad for an individual.

Suppose that Joe has a deadly disease and is offered an operation. We know that 85 percent of relevantly similar individuals (when considering things like age, sex, and medical history) who have the disease and receive the operation make a full recovery, while for the other 15 percent the operation has no effect. When deciding whether the operation is likely to be good for Joe, suppose that the relevant probability is the objective chance of it being good for him. But we do not (and perhaps cannot) know what this objective chance is. Is it true, then, that we cannot make *any* reasonable judgments about how likely the operation is to be good for Joe? I do not believe so. We can use subjective chances to form reasonable beliefs about the objective chance of Joe's operation being successful. Through our knowledge that eighty-five percent of people within the relevant population make a full recovery, we can form the reasonable belief that the objective chance of Joe's operation being successful is roughly 0.85. One way of forming reasonable beliefs about the objective chances of the success of Joe's operation is by using information about averages.

Many other cases are similar. We use averages across populations to determine that children, in general, do better on their exams if they study. Based on this, we can reasonably believe that *our child* has a better chance of doing well on her exams if she studies, and encourage her to study, even though it is possible that she is in the minority for whom studying would be detrimental. In this case, subjective chance plays an important role in our belief-forming process. It would certainly appear odd to maintain that we cannot form reasonable beliefs by this method about how good studying is likely to be for our child because subjective chances merely tell us about averages. If so, then almost all our beliefs about probabilities would be unreasonable, and I think most would agree that many such beliefs are reasonable.

If subjective chances can play this belief-forming role in everyday cases, then there seems to be no reason why they cannot play an analogous role in the disability case. That is, there seems to be no reason why we cannot form reasonable

(non-question-begging) beliefs based on averages across populations about the objective chance of an individual doing as well with a disability as without.

I have argued that forming reasonable judgments about an individual's objective chances based on information about averages need not be problematic, as in the Joe case. However, there is a disanalogy between the Joe case and the disability case. Both require taking information about averages and extrapolating this to individual chances. However, forming judgments in the disability case requires working out average well-being levels in the closest non-disablist worlds. And perhaps it is *this idealization* that Barnes thinks requires begging the question. Or perhaps, as she alludes to in later work, Barnes thinks it is "close to impossible" to make inferences about well-being levels in non-disablist worlds.<sup>95</sup>

I will now argue, briefly and finally, that making at least *reasonably robust* inferences about well-being levels in non-disablist worlds does not require question begging and that there is no good reason to think that it is impossible. Why might Barnes think making the relevant inferences requires assuming that disability is non-instrumentally bad? Perhaps she thinks that everything that instrumentally (and disproportionately) disadvantages disabled people is due to disablism. If so, all instrumental harms of disability would not occur in non-disablist worlds and we *would* have to make assumptions about the non-instrumental value of disability to infer the well-being of disabled people in these worlds. However, this assumption seems unwarranted. Things might disproportionately disadvantage disabled people without being disablist.<sup>96</sup> For example, it might turn out that the just division of resources disadvantages disabled people, or that reduction of valuable options is a non-disablist harm of disability.<sup>97</sup> Of course, an argument would be required to show that these things are non-disablist instrumental harms of disability, but it appears possible to provide such arguments without making *any* assumptions about disability's non-instrumental value.

Admittedly, making reasonable judgments about average well-being levels in non-disablist worlds is difficult, but Barnes has offered no compelling reason to accept her extremely strong claim that it is impossible. Here is one sketch of a suggestion about how we *might* do it. First, we could identify what counts as disablism. Then, we might start from the neutral (and charitable to the MDV) assumption that disability and non-disability are equally non-instrumentally valuable. This would shift focus onto instrumental value, and seems a reasonable starting point given that debates about non-instrumental value often ap-

95 Barnes, *The Minority Body*, 99–100.

96 Amundson, "Disability, Ideology, and Quality of Life," 114; Brown, "Is Disability a Neutral Condition?" 195; Howard and Aas, "On Valuing Impairment," 1129.

97 Singer, "Response to Mark Kuczewski," 56; Crawley, "Disability, Options and Well-Being."

pear intractable. Next, we could identify the instrumental harms and benefits of disability and non-disability in our world and consider whether these would be present in the closest non-disablist worlds. This requires determining whether these harms or benefits are disablist. Finally, we would need to consider whether disability or non-disability would have any additional instrumental harms or benefits in non-disablist worlds.

These are admittedly difficult tasks, but it seems that careful consideration of this information would provide a reasonable idea of whether disability makes one likely to be worse off in non-disablist worlds. This process is certainly tricky—and the resulting judgment would, of course, be fallible—but it is hard to see why it would be *impossible*. In fact, this process is made easier because  $\delta^*$  postulates a difference between the average well-being levels of two large groups. As such, we need not identify absolute or even average well-being levels, nor do we have to worry about differences between particular pairs of individuals or nonstandard cases. We need only make judgments about *significant and systematic differences* in the well-being of disabled and non-disabled people in the closest non-disablist worlds, which, I contend, does seem possible.

In sum, Barnes objects that to form reasonable beliefs about  $\delta^*$  we must make question-begging assumptions about the non-instrumental value of disability. This is false. We can make reasonable (although fallible) judgments about  $\delta^*$  by looking at averages across populations, as we do in other contexts. The fact that we must consider well-being levels in idealized worlds makes forming the relevant judgments trickier, but Barnes offers no conclusive reason why doing so is impossible or requires question begging. Therefore, Barnes's objection fails to show that probabilistic views like  $\delta^*$  are bad interpretations of the BDV.

## 6. CONCLUSION

In this paper, I have clarified the nature of the best version of the BDV. I argued in section 3 that the BDV's best interpretation is  $\delta^*$ : it is the view that a person is *likely to be, all things considered*, worse off with a disability than without. Next, in section 4, I argued that, in characterizing the BDV as akin to  $\iota^*$ , Barnes attacks an uncharitable understanding of the BDV that is unlikely to be endorsed by many BDV proponents, and she misses the opportunity to challenge the most plausible and relevant version, which is  $\delta^*$ . As such, one can be persuaded by Barnes's

arguments and still hold a plausible version of the BDV, and the best version remains unchallenged.<sup>98</sup>

*University of Nottingham*  
tom.crawley94@yahoo.com

#### REFERENCES

- Amundson, Ron. "Disability, Ideology, and Quality of Life: A Bias in Biomedical Ethics." In *Quality of Life and Human Difference: Genetic Testing, Health Care, and Disability*, edited by David Wasserman, Jerome Bickenbach, and Robert Wachbroit, 101–24. Cambridge: Cambridge University Press, 2005.
- Andric, Vuko, and Joachim Wundisch. "Is It Bad to Be Disabled? Adjudicating between the Mere-Difference and the Bad-Difference Views of Disability." *Journal of Ethics and Social Philosophy* 9, no. 3 (December 2015): 1–16.
- Barker, Matthew J., and Robert A. Wilson. "Well-Being, Disability, and Choosing Children." *Mind* 128, no. 510 (April 2019): 305–28.
- Barnes, Elizabeth. "Disability and Adaptive Preference." *Philosophical Perspectives* 23, no. 1 (December 2009): 1–22.
- . "Disability, Minority, and Difference." *Journal of Applied Philosophy* 26, no. 4 (November 2009): 337–55.
- . *The Minority Body: A Theory of Disability*. Oxford: Oxford University Press, 2016.
- Bayne, Tim, and Neil Levy. "Amputees by Choice: Body Integrity Identity Disorder and the Ethics of Amputation." *Journal of Applied Philosophy* 22, no. 1 (2005): 75–86.
- Bickenbach, Jerome. "Disability and Health Care Rationing." *Stanford Encyclopedia of Philosophy* (Spring 2016). <https://plato.stanford.edu/entries/disability-care-rationing>.
- Bradley, Ben. "Doing Away with Harm." *Philosophy and Phenomenological Research* 85, no. 2 (September 2012): 390–412.
- Bradley, Seamus. "Are Objective Chances Compatible with Determinism?" *Philosophy Compass* 12, no. 8 (August 2017): 1–11.

98 This work was supported by the Arts and Humanities Research Council's Midlands3Cities Doctoral Training Partnership (grant number AH/L50385X/1). I would like to extend special thanks to Chris Woodard and Iain Law for offering invaluable advice on this project. I would also like to thank two reviewers from *JESP* for their thorough and thoughtful comments and those who took part in helpful discussions at presentations of this work in Nottingham, Reading, Brighton, and Porto.

- Brown, Jeffrey M. "Is Disability a Neutral Condition?" *Journal of Social Philosophy* 47, no. 2 (Summer 2016): 188–210.
- Campbell, Stephen M., and Joseph A. Stramondo. "The Complicated Relationship of Disability and Well-Being." *Kennedy Institute of Ethics Journal* 27, no. 2 (June 2017): 151–84.
- Carel, Havi. *Illness: The Cry of the Flesh*. Dublin: Acumen, 2008.
- Crawley, Thomas. "Disability, Options and Well-Being." *Utilitas* 32, no. 3 (September 2020): 316–34.
- Crisp, Roger. "Well-Being." *Stanford Encyclopedia of Philosophy* (Fall 2017) <https://plato.stanford.edu/entries/well-being>.
- Eyre, Sarah. "A Few Awesome Things about Being Disabled." xoJane, May 17, 2012. <https://www.xojane.com/healthy/disabilities-unexpected-silver-linings>.
- Finnis, John. *Natural Law and Natural Rights*. Oxford: Oxford University Press, 1979.
- Franklin, Donald. "Calibrating QALYs to Respect Equality of Persons." *Utilitas* 29, no. 1 (March 2017): 65–87.
- Glover, Jonathan. *Choosing Children: Genes, Disability, and Design*. Oxford: Oxford University Press, 2006.
- Gustavsson, Anders. "The Role of Theory in Disability Research: Springboard or Strait-jacket?" *Scandinavian Journal of Disability Research* 6, no. 1 (2004): 55–70.
- Hájek, Alan. "Interpretations of Probability." *Stanford Encyclopedia of Philosophy* (Fall 2019). <https://plato.stanford.edu/entries/probability-interpret>.
- Hawkins, Jennifer. Review of *The Minority Body: A Theory of Disability*, by Elizabeth Barnes. *Ethics* 128, no. 2 (January 2018): 462–67.
- Heathwood, Chris. "Desire Satisfactionism and Hedonism." *Philosophical Studies* 128, no. 3 (April 2006): 539–63.
- . "Subjective Theories of Well-Being." In *The Cambridge Companion to Utilitarianism*, edited by Ben Eggleston and Dale Miller, 199–219. Cambridge: Cambridge University Press, 2014.
- Howard, Dana, and Sean Aas. "On Valuing Impairment." *Philosophical Studies* 175, no. 5 (May 2018): 1113–33.
- Kahane, Guy, and Julian Savulescu. "The Welfarist Account of Disability." In *Disability and Disadvantage*, edited by Kimberley Brownlee and Adam Cureton, 14–53. Oxford: Oxford University Press, 2009.
- Korsgaard, Christine M. "Two Distinctions in Goodness." *Philosophical Review* 92, no. 2 (April 1983): 169–95.
- Lewis, David K. "Causation." *Journal of Philosophy* 70, no. 17 (October 1973): 556–67.

- Lim, Chong-Ming. "Disabilities Are Also Legitimately Medically Interesting Constraints on Legitimate Interests." *Mind* 127, no. 508 (October 2018): 977–1002.
- McMahan, Jeff. "Causing Disabled People to Exist and Causing People to Be Disabled." *Ethics* 116, no. 1 (October 2005): 77–99.
- Nordenfelt, Lennart. *On the Nature of Health: An Action-Theoretic Approach*. Dordrecht: Springer, 1995.
- Office for Disability Issues. "Public Perceptions of Disabled People: Evidence from the British Attitudes Survey 2009." [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/325989/ppdp.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/325989/ppdp.pdf).
- Oliver, Mike. *Understanding Disability: From Theory to Practice*. Basingstoke: Macmillan, 1996.
- O'Neill, John. "The Varieties of Intrinsic Value." *Monist* 75, no. 2 (April 1992): 119–37.
- Parfit, Derek. *Reasons and Persons*. Oxford: Oxford University Press, 1984.
- Putnam, Daniel. "Disability, Democratic Equality, and Public Policy." In *The Routledge Handbook of Ethics and Public Policy*, edited by Annabelle Lever and Andre Poama, 306–18. London: Routledge, 2018.
- Schramme, Thomas. "Disability (Not) as a Harmful Condition: The Received View Challenged." In *Disability and the Good Human Life*, edited by Jerome E. Bickenbach, Franziska Felder, and Barbara Schmitz, 72–92. Cambridge: Cambridge University Press, 2013.
- Schroeder, Andrew. "Well-Being, Opportunity, and Selecting for Disability." *Journal of Ethics and Social Philosophy* 14, no. 1 (October 2018): 1–27.
- Shakespeare, Tom. *Disability Rights and Wrongs Revisited*. London: Routledge, 2013.
- Shakespeare, Tom, and Nicholas Watson. "The Social Model of Disability: An Outdated Ideology?" *Research in Social Science and Disability* 2 (2002): 9–28.
- Singer, Peter. "Ethics and Disability: A Response to Koch." *Journal of Disability Policy Studies* 16, no. 2 (Fall 2005): 130–33.
- . "Response to Mark Kuczewski." *American Journal of Bioethics* 1, no. 3 (2001): 55–56.
- Singer, Peter, John McKie, Helga Kuhse, and Jeff Richardson. "Double Jeopardy and the Use of QALYs in Health Care Allocation." *Journal of Medical Ethics* 21, no. 3 (June 1995): 144–50.
- Steer, Emma. "I'm in a Wheelchair: Here's 5 Truly Awesome Things about It." *Cracked*, July 16, 2017. <https://www.cracked.com/blog/5-ways-my-disability-actually-improved-my-life>.



- Wasserman, David, Adrienne Asch, Jeffrey Blustein, and Daniel Putnam. "Disability: Definitions, Models, Experience." *Stanford Encyclopedia of Philosophy* (Summer 2016). <https://plato.stanford.edu/archives/sum2016/entries/disability>.
- Wendell, Susan. "Unhealthy Disabled: Treating Chronic Illness as Disabilities." *Hypatia* 16, no. 4 (August 2001): 17–33.
- World Health Organization. "International Classification of Functioning, Disability and Health," May 2001. <https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health>.

## THE GOAL PROBLEM IN THE “NOW WHAT” PROBLEM

*Xinkan Zhao*

SUPPOSE MORAL ERROR THEORY is true. The natural question that comes next is what we practically should do with this already existing system of morality, a question often dubbed the “now what” problem.<sup>1</sup> For those error theorists who go on to deny even instrumental reasons, this very question of what we *should* do next is to be further paraphrased in metaphysically innocuous terms, but many error theorists wish to hold a non-error theory regarding instrumental reasons, and for them, the “now what” problem can be taken at face value. For the sake of argument, let us further suppose that the latter group is correct and that we can take the question literally.

An array of proposals has been put forward.<sup>2</sup> Roughly categorized, these proposals fall under three types. First, conservatism suggests that we keep morality as before; second, abolitionism suggests that we discard morality altogether; third, substitutionism suggests that we keep the shell of moral discourse but supply a different, non-erroneous semantics for it.<sup>3</sup> Note that substitutionism as such is itself a group of views. According to different substitute semantics, moral terms could be expressing certain conative attitudes or ascribing certain natural properties, among others. In this paper, I do not intend to adjudicate on this dispute; instead, I wish to point out the problematic assumption largely taken for granted by most, if not all, of the proposals. I identify it as the goal problem in the “now what” problem. Simply put, theorists have been too casual in identifying the agents’ set of goals that generates the instrumental reasons to adopt their proposals. In what follows, I will present arguments against the background of

1 The label comes from Lutz, “The ‘Now What’ Problem for Error Theory.”

2 For a careful, up-to-date survey, see Jaquet, “Sorting Out Solutions to the Now-What Problem.”

3 For typical examples of the three types, respectively, see Olson, *Moral Error Theory*; Garner, “Abolishing Morality”; and Lutz, “The ‘Now What’ Problem for Error Theory.” I also intend to count fictionalism (such as Joyce, *The Myth of Morality*) as a species of substitutionism. But for a nuanced taxonomy that treats content fictionalism and force fictionalism differently, see Jaquet, “Sorting Out Solutions to the Now-What Problem.”

Lutz, since his is one of the most recent systematic critiques and defenses on this problem (and indeed also where the useful label of the problem comes from), but the idea generalizes to other proposals that bear relevant similarity.<sup>4</sup>

#### 1. THE ARGUMENT FROM INSTRUMENTAL REASONS

According to the version of substitutionism that Lutz favors, we are to replace the extant moral semantics with a different one that assign properties about the speaker's attitudes to moral terms as semantic values.<sup>5</sup> For example, "Murder is wrong" would no longer ascribe to murder the property of moral wrongness, which moral error theorists take to be queer; instead, it would ascribe to murder the property of being disapproved of by the speaker. This new semantics may incur other problems, which I will turn to shortly, and it may even fail to count as *moral* semantics, but it is rid of uniform falsity. As long as the speaker does disapprove of murder, "Murder is wrong" expresses a truth. *To the extent that we care about truth*, we should, or at least have *pro tanto* reason to, favor substitutionism over, say, conservatism.

Lutz's argument is clearly in the form of instrumental reasoning:

1. We as normal agents have a normal set of goals (such as to hold only true beliefs).<sup>6</sup>
2. Anyone with this normal set of goals instrumentally should adopt substitutionism (of a certain version—I will drop this qualification for now).
3. Therefore, we should adopt substitutionism.

Call this specific, *actual*, normal set of goals *S*. Lutz is not very explicit about what members are in *S*, besides offering two examples: to hold only true beliefs and to get along with friends.<sup>7</sup> To better see the goal problem, we need a slightly more detailed list of the members in *S*.

Consider 2. Following Lutz, we may agree that the goal to only believe truths is in *S*, but that is certainly not sufficient for the adoption of substitutionism. For the case to be made, *S* needs also to be such that the agent who has it is willing to tolerate a certain level of discourse disorder and a certain level of insincerity, both of which are necessitated consequences of substitutionism.

4 Lutz, "The 'Now What' Problem for Error Theory."

5 Lutz, "The 'Now What' Problem for Error Theory."

6 Note that, here and throughout, the set needs to be understood as *ordered*, since the same goals that are assigned different priorities should count as forming different sets.

7 Lutz, "The 'Now What' Problem for Error Theory," 353.

Since the semantics has changed, not all original discourse patterns can be preserved. Some parts work as before, others not. Consider again the discourse involving right and wrong after the substitution. When I say “Stealing is wrong,” I am in fact saying that I disapprove of stealing, and from this we may infer, as before, that I have a reason (at least of an instrumental kind) not to steal. Similarly, just as before, my discovery that my neighbor approves of lying does not warrant my assenting to “Lying is right,” since now my assenting to “Lying is right” is only warranted by *my* approval of lying. However, some discourse patterns are disrupted by the substitution. Suppose I approve of donating. Now, according to the new semantics, my approval would automatically render my utterance “Donating is right” true, or in other words, “Donating is right because I approve of it” would express a trivial truth. Similarly, suppose I believe that lying is wrong. I would now *not* be making a legitimate move if I infer from this that my neighbor has a reason not to lie, because my disapproval of lying certainly should not affect reasons for my neighbor to lie if he decides not to care about my (dis)approvals. The moral discourse is to a certain extent in disorder.

Moreover, as Lutz has pointed out, substitutionism may leave the agents who adopt this approach at an insincere position, since they knowingly talk to people with a semantics that the interlocutors do not know they have adopted. In a sense, substitutionists are being deceptive.<sup>8</sup>

For Lutz, these problems should not prevent us from adopting substitutionism, because the new semantics and the old have a substantial degree of overlap, so that, by and large, our moral discourse should function as smoothly as before. From this, we can say something more about the set of goals *S*. Roughly put, *S* needs to be such a set of goals that any agent with it wants to believe in and assert truths only and has a certain level of tolerance toward discourse disorder and conversational insincerity, and that there are no members in the set that may override these features. Perhaps yet more details are needed in order for *S* to suffice for substitutionism, but suppose this rough characterization will do, so that 2 is rendered true.

## 2. THE GOAL PROBLEM

When we have specified what *S* is like in the way above, 2 comes out true, but now we have the goal problem: it is no longer clear that 1 is true. That is, *do* we, as normal agents, have *S* as our set of goals? There is very good *prima facie* reason to think not, and this can be most clearly shown by counterexamples. We can easily imagine scenarios where agents have different goals that generate instrumental

8 Lutz, “The ‘Now What’ Problem for Error Theory,” 366.

reasons for them to accept "now what" solutions other than substitutionism, but may count as perfectly normal agents nonetheless.

Consider George, someone who shares a deep concern for truth and finds erroneous moral discourse unpalatable, just like the substitutionist. However, he is a direct and honest person and has little to no tolerance for twisting language like substitutionism requires, and finds totally unacceptable thoughts like donation is right because he approves of it. "If I approve of it, I approve of it. Just forget about right and wrong!" says George. His honesty also prevents him from using words like "is wrong" to privately mean that he disapproves of something when he talks to a friend; instead, he spits out (what he takes to be) the truth when he talks to his friends in the most direct, honest language, which he regards as the best way of treating a friend. This means that George has a different set of goals than S, but is George a "normal agent"? Does he count as one of "us"? I think he most certainly does.

For another counterexample, consider Peter, someone who also shares a deep concern for truth, but not unconditionally. Peter forms beliefs according to truth-oriented epistemic norms only if it does not make him suffer from major negative feelings. He believes in an afterlife because that alleviates his existential angst; he believes stone and sand have minds capable of human understanding because that makes him feel less lonely; he also believes in values and rules because that provides him with something he can cling on to for navigating himself through people with very different personalities and commitments. He is also familiar with error-theoretic arguments that moral properties are queer or even impossible, but he remains unmoved by this. Morality serves him well, and disbelieving it incurs too big a price for him, so he decides to continue belief in morality. Is Peter one of us normal agents? It might seem not at first glance, but we should be much more inclined toward a positive answer if we consider the number of non-philosophical believers who are devout but ask for no minimally plausible arguments for their beliefs. Indeed, if normality is defined in terms of population percentage, it is we argument-hunting philosophers who are abnormal.

We can have further counterexamples with yet different sets of goals. For example, Dockstader identifies a therapeutic need of agents, and argues for what he calls "reactionary moral fictionalism," according to which we should shy away from moral discourse as much as possible, and assume a fictionalist stance when hiding is no longer an option.<sup>9</sup>

9 Dockstader, "Reactionary Moral Fictionalism." Novel as the idea is, I am not sure that Dockstader's development is successful, since it seems to me that the therapeutic need is incompatible with the occasional insincerity this approach requires. I leave the assessment of the argument to the readers' discretion.

But I think the message is already clear: 1 is most likely false. Many agents, whom we are inclined to regard as perfectly normal, simply do not seem to have *S* as their set of goals, and since 2 shows only that agents with *S* have instrumental reasons to accept substitutionism, it does not follow that *we* as normal agents should accept substitutionism. George may have good reasons to adopt abolitionism, and Peter conservationism, both being members of “us,” and for those of us with yet different sets of goals, yet other options may be more suitable than substitutionism.

This line of comment generalizes beyond Lutz’s case. Theorists who have proposed different routes forward after moral error theory have produced many interesting and compelling arguments to the effect that we should accept a certain proposal *if* we have certain set of goals. But this falls short of validation of the proposal *per se* since, in addition to the conditional, they would still need arguments to the confirmation of the antecedent, namely, that we do in fact have the set of goals.

But what exactly our set of goals is in fact like is a strictly empirical thesis that can only be determined through serious empirical investigation. We philosophers’ armchair pondering can be fatally misleading in this respect for at least three reasons. First, when we engage in armchair theorizing, the potential agents we can think of can easily fall prey to selective bias. We surround ourselves with philosophers and may therefore naturally assume that all “normal” agents are rational and very willing to follow arguments, but that could be an inaccurate representation of the actual world. Second, the real goals of agents are not always obvious. Testimony and apparent behavior may well be misleading, since under the habitual mean lines and aggressive postures there could be very kind intentions, and vice versa: cruel, selfish goals could hide behind the guise of warm smiles and friendliness. Again, philosophers should claim no expertise in this field. Third, many goals are interrelated in a way unbeknownst to the agents having them, and even unbeknownst to experts prior to substantial long-term studies, so that our actual overall goal may still elude us even if we know the goals separately with certainty. Imagine someone wants to be the most powerful person in the community, but further imagine, as a matter of fact unknown to him, being in a powerful position would bring a huge amount of stress, frustrating his other goal of living a happy life, where happiness is defined in terms of subjective feelings. When this happens, we might say he really does *not* have the goal of living a happy life, or that his goal is in fact overridden, despite his own self-conception to the contrary. Once more, such interrelatedness of possible goals is discoverable only through empirical investigation.

Perhaps it will turn out that we do have the set of goals like the previously

characterized  $S$ , and then we do have good reasons to adopt substitutionism. But insofar as characters like George and Peter also look like normal agents, we have at least very good *prima facie* reason to question what exactly our set of goals is in fact like. It may turn out to support abolitionism, conservationism, or what have you. Indeed, it could also turn out that there is no such set shared by all, and the “now what” problem framed in terms of instrumental reasons and a collective “we” simply evaporates. But no matter which is the case, there is no armchair solution to be drawn for the “now what” problem because of the empirical nature of the goal problem.

Of course, the philosophers who have proposed different solutions may choose to retreat to some conditional solution, in the form of  $\mathcal{Z}$ , that *if* we have such-and-such set of goals, then we should adopt this or that solution. True, but we should also note that a play-safe strategy of this kind may deprive the solutions of a substantial amount of theoretical interest. As we have seen in the discussion of Lutz’s proposal, the set of goals may yield a rather long antecedent of the conditional solution. When we unpack the set, the solution will eventually be something like “If we care about  $x, y, z, \dots$  and if these items are prioritized in a certain way, then we should adopt...” The truth of the proposal comes at some cost of its nontriviality, and I am not sure whether this is a price that philosophers in the debate are willing to pay.

### 3. CONCLUSION

In this brief note, I argued that the philosophers who propose solutions to the “now what” problem typically face a goal problem. The problem has its root in the argument they back up their proposal with, which is one of instrumental reason, consisting of two premises. First, we as normal agents have a certain set of goals. Second, agents with this set of goals instrumentally should accept their proposal. I have argued that when we specify the set of goals with sufficient detail so that the second premise comes out true, the first premise will most likely come out false. These philosophers could retreat to a conditional solution, but that comes with the cost of the solution being less nontrivial; instead, they may try to establish the truth of the first premise, but that requires sufficient empirical investigation for which no armchair speculation will suffice.

Peking University  
zhaoxinkan@pku.edu.cn



## REFERENCES

- Dockstader, Jason. "Reactionary Moral Fictionalism." *Philosophia* 48, no. 2 (April 2020): 519–34.
- Garner, Richard. "Abolishing Morality." *Ethical Theory and Moral Practice* 10, no. 5 (November 2007): 499–513.
- Jaquet, François. "Sorting Out Solutions to the Now-What Problem." *Journal of Ethics and Social Philosophy* 17, no. 3 (June 2020): 239–58.
- Joyce, Richard. *The Myth of Morality*. Cambridge: Cambridge University Press, 2001.
- Lutz, Matt. "The 'Now What' Problem for Error Theory." *Philosophical Studies* 171, no. 2 (November 2014): 351–71.
- Olson, Jonas. *Moral Error Theory: History, Critique, Defence*. Oxford: Oxford University Press, 2014.

## NORMATIVE UNCERTAINTY WITHOUT UNJUSTIFIED VALUE COMPARISON

A RESPONSE TO CARR

*Ron Aboodi*

**N**ORMATIVE UNCERTAINTY is uncertainty about *normative* questions, such as whether it is permissible to eat fish (or whether the suffering of fish is as intrinsically bad as human suffering), in contrast to *empirical* questions such as whether fish feel pain (or whether their suffering can be as intense as human suffering). What strategy would it be rational to use for making practical decisions under (purely) normative uncertainty?<sup>1</sup> Jennifer Rose Carr's 2020 paper, "Normative Uncertainty without Theories," defends the strategy of choosing the option with the highest expected value against worries concerning *intertheoretic value comparison* (hereafter "IVC").<sup>2</sup> To illustrate these worries in the way that I find most persuasive, consider the following example:

Extani is uncertain whether it is OK to eat fish because he is uncertain whether to adopt W.D. Ross's normative theory or Peter Singer's, justifiably assuming that each would support a different answer.

Maximizing expected value requires cardinal comparisons of the units of value across the relevant theories.<sup>3</sup> Extani's case does not allow such comparisons based on what Christian Tarsney calls "intertheoretic agreement."<sup>4</sup> Neither could structural methods for commensuration be applied here (if anywhere) without unjustified arbitrariness.<sup>5</sup> If there is a way to justify IVC in Extani's case,

- 1 For the *moral* importance of coping with such uncertainty, see Aboodi, "One Thought Too Few."
- 2 Carr, "Normative Uncertainty without Theories," 755. I use the term "value" in this context interchangeably with "utility" as Carr uses this term, and with "choice-worthiness."
- 3 Carr, "Normative Uncertainty without Theories," 750.
- 4 Tarsney, "Intertheoretic Value Comparison." Note that Carr identifies a problem with Tarsney's approach ("Normative Uncertainty without Theories," 759–60).
- 5 This applies to methods proposed in Lockhart, *Moral Uncertainty and Its Consequences*; Ross,

despite the known problems, it has not been published yet (to the best of my knowledge).<sup>6</sup>

Carr neither defends nor wants to rely on the justifiability of IVC. What makes Carr's proposal unique is her claim that one could maximize expected value under normative uncertainty *without* IVC. According to Carr, her proposal avoids IVC simply because it avoids *theories*.<sup>7</sup> Carr envisions agents who distribute their credence not among normative theories, but among *hypotheses about the objective values of the options at hand*. These are the values assigned by "the utility function determined by whichever moral theory is in fact correct," hereafter *the ideal normative function*.<sup>8</sup> While this function may be unique only up to positive affine transformation, the ratios of the differences between its assigned values must agree with the objective normative truth.<sup>9</sup> I suggest precisifying the ideal normative function as follows: a function whose expected-value maximization under *nonnormative* uncertainty generates decisions that are in line with the *correct normative views*.<sup>10</sup> But my argument could work with some other precisifications as well.

I agree with Carr that the problem of IVC does not stand in the way once an agent rationally distributes her credence only among candidate specifications of the ideal normative function that share the same unit of value. However, this leaves us with what I will call the *justificatory problem of IVC*: how to reach such a credal distribution justifiably. I diagnose the difficulty of doing so in section 1. Carr does not provide any illustration of justifiably reaching a credal distribution of the type she requires, nor an argument for this potential justifiability. This raises the worry that no such credal distribution could ever be justified, which implies that Carr's proposal cannot be (justifiably) implemented. My first aim is to show how some such credal distributions could be justified, thereby allaying this worry. Some of Carr's formulations create the impression that the solvability of the justificatory problem of IVC turns on whether or not the agent's uncertainty is about *theories* as such; but I will argue that this claim would be false. More-

---

"Rejecting Ethical Deflationism," 764–65; and Sepielli, "What to Do When You Don't Know What to Do."

6 Concerning the relevant problems, see Sepielli, "Moral Uncertainty and the Principle of Equity among Moral Theories"; Gustafsson and Torpman, "In Defence of My Favourite Theory," 163–64; Hedden "Does MITE Make Right?" 112; and Carr, "Normative Uncertainty without Theories," 752–53.

7 Carr, "Normative Uncertainty without Theories," 755.

8 Carr, "Normative Uncertainty without Theories," 754.

9 Carr, "Normative Uncertainty without Theories," 756–57.

10 I will not focus on addressing metaphysical worries concerning IVC in this paper, but my precisification of the ideal normative function above constitutes a beginning of an answer to some of them.

over, I will identify other features of the agent's epistemic state that are more relevant for determining whether she can justifiably reach a credal distribution among candidate specifications of the ideal normative function that share the same unit of value. Helping to illuminate the conditions for this justifiability is my second aim.

In section 1, first, I use a variation on Extani's case to illustrate how the justificatory problem of IVC can persist even when the agent focuses directly on functions rather than theories. It does not seem that Carr's proposal can be justifiably implemented in such a case. Second, generalizing from this case, I suggest that the justificatory problem of IVC (normally) constitutes a serious obstacle to maximizing expected value justifiably whenever the defining features of the normative hypotheses with which the agent starts out do not refer to the same unit of value.

Third, aided by this diagnosis, I identify a type of normative uncertainty wherein the agent may justifiably conceive of all the relevant normative hypotheses as referring to the same unit of value. This would allow the agent to maximize expected value along Carr's lines (allaying the implementation worry). I focus on normative uncertainty that stems from indecisive normative intuitions. Despite being one of the most common types of normative uncertainty, it has not been sufficiently examined in the relevant literature. I will illustrate such normative uncertainty in section 2, and argue that Carr's proposal can be justifiably implemented in particular instantiations of it.

#### 1. THE JUSTIFICATORY PROBLEM OF INTERTHEORETIC VALUE COMPARISON

To see that the justificatory problem of IVC does not dissipate merely by avoiding theories, consider the following variation on Extani's case:

Extani\* is uncertain whether it is permissible to eat fish because he is uncertain whether to side with W.D. Ross or Peter Singer, just like Extani. But Extani\* is not thinking of their *theories* as such. Extani\* finds a new blog where Singer argues that, for practical questions of this type, his view should be implemented by maximizing the expected value of function  $F_1$ . Additionally—in an exciting historical breakthrough—Extani\* digs up an old manuscript by Ross that endorses maximizing the expected value of  $F_2$ . He is thus uncertain which function will generate the right verdict on whether to eat fish.

Had Extani\*'s uncertainty led him to distribute his credence among  $F_1$  and  $F_2$  and maximize the expected value (without normalization), this would have

been unjustified, because there is no reason to think that  $F_1$  and  $F_2$  share the same unit of value. In order to reach a justified decision, Extani\* must find a justified “exchange rate” between the unit that  $F_1$  uses and the unit that  $F_2$  uses. (Carr neither provides guidance for Extani\* on how to do so, nor a reason to think that doing so is possible in Extani\*’s case.)

I propose the following diagnosis of the difficulty of solving the justificatory problem of IVC: if the normative hypotheses with which the agent starts out (whether they constitute theories, functions, or other types of normative hypotheses) do not refer to the same unit of value, it is hard to imagine how the agent could have access to facts that determine the right “exchange rate” among their units of value (if any such facts exist). This seems typical when the source of the normative uncertainty is trusting experts or peer disagreement.

In light of this diagnosis, it makes sense to turn our attention to situations wherein the facts that determine the right IVCs are directly accessible from the agent’s perspective. Such is the case when the agent constructs all the relevant normative hypotheses on her own, and does so by reference to the same unit of value, so that the right intertheoretic comparisons are trivially derived from the *defining features* of the hypotheses. This seems natural in some epistemic states wherein the source of normative uncertainty is the agent’s *indecisive normative intuitions*. (I use this term in a broad sense, covering any uncertainty or inconsistency at the level of intuitive normative judgments, or normative “seemings.”) In some such epistemic states, the agent may justifiably construct each normative hypothesis—by reference to the same unit of value—on the basis of a different, internally consistent subset of her own intuitive normative judgments. In the following section I illustrate such a case.

I will rely on two preliminary assumptions. First, the relevant justifiability ultimately turns on the agent’s relevant *evidence* (in the broad sense that covers all the elements in the agent’s epistemic state that may have a role in justifying her credence). Second, normative intuitions can be part of the relevant evidence, providing at least an initial, defeasible justificatory force.<sup>11</sup>

## 2. JUSTIFIABLE VALUE COMPARISONS UNDER INDECISIVE NORMATIVE INTUITIONS

Consider Inti, whose uncertainty about whether it is permissible to eat fish

<sup>11</sup> The justificatory force of intuitions is endorsed by proponents of the “reflective equilibrium” method (such as Daniels, “Wide Reflective Equilibrium and Theory Acceptance in Ethics”), as well as intuitionists (such as Huemer, *Ethical Intuitionism*). I consider my relevant assumption above as weaker than each of these views.

stems from indecisive normative intuitions. On the one hand, it seems to her that the normative significance of animal suffering is negligible when compared to human matters. On the other hand, this intuition seems suspicious to Inti because it implies some sort of problematic speciesism.

Fortunately, Inti is having lunch with “Carr\*”:

1. Inti: This fish looks yummy. But I don’t know whether to eat it, due to the indecisive intuitions I told you about.
2. Carr\*: I wonder whether my research could come in handy here. Can we give it a shot?
3. Inti: Sure!
4. Carr\*: OK, let’s start by assuming that there’s nothing wrong with your intuitive judgment that animal suffering is negligible when compared to human matters. Could you evaluate your options under this assumption first? You have two options: *Eat* and *Avoid*.
5. Inti: Well, I’d assign a higher value to *Eat* than to *Avoid*, given our assumption. And I can say that *Eat* would be better than normatively neutral choices (such as my choosing to lean slightly to the left now), and *Avoid* worse than such neutral choices.
6. Carr\*: Let’s denote the value of *Eat* given our assumption by 1, and the value of normatively neutral actions by 0. How would you rate *Avoid* using this unit of value, under the given assumption?
7. Inti: Hmm, using this unit *Avoid* gets a  $-1$ .
8. Carr\*: OK great. Hold on to this unit of value as we take the next step: now assume that your intuition that animal suffering is negligible when compared to human matters should be rejected. How would you evaluate your options on this assumption, using the same unit?
9. Inti: If this intuition should be rejected, then—according to my remaining intuitions—there’s no difference between human and animal suffering. The totality of my evidence decisively supports treating the badness of the frustration of not eating what I crave as remaining stable across my normative hypotheses. So, to answer your question: *Avoid* still gets a  $-1$  but *Eat* would get a  $-6$ .
10. Carr\*: I see! Now please specify your credal distribution among these hypothetical functions, alongside any other function whose expected value maximization under nonnormative uncertainty would possibly cohere with the right normative view, from your epistemic perspective. (Any such alternative hypothetical function must be constructed using the same unit of value and cannot be *equivalent* to any of the

other hypothetical functions, in the sense of having the same practical implications in every possible situation of nonnormative uncertainty.)

11. Inti: I would distribute my credence equally and exclusively among the two functions that I mentioned. This is, to the best of my estimates, what my evidence calls for.
12. Carr\*: You know what this implies, don't you?
13. Inti: Yes. Take this tempting fish away from me!!!

The justifiability of each of Inti's deliberative steps ultimately depends on her intuitions (and the rest of her evidence).<sup>12</sup> It does not matter, for my argument, how (un)common epistemic states with intuitions that render such a deliberative route justifiable are. The worry I need to address (in order to defend the implementability of Carr's proposal in such cases) is that *no* specification of Inti's case could render her deliberation justifiable.

Inti's deliberation up to step 5 seems pretty safe from this worry. There is nothing wrong (at least under some specifications of Inti's case) with evaluating options in ordinal terms under a particular normative assumption. Following Carr, I will assume here that there is nothing wrong, in principle, with the cardinal normative evaluation of options, and this allays the discussed worry concerning step 7. Concerning step 9, remember that—unlike Extani\*—what determines the right “exchange rate” between the units of Inti's hypothetical functions is their defining features, which Inti can access directly because they are her own constructs. As long as Inti held fixed the same unit of value in her mind from step 6 onward, and constructed all the hypothetical functions as referring to this unit, step 9 might be justified. The burden of proof lies with anyone who would claim that there is an *inherent* problem here. And the same applies to step 11. Indeed, on some specifications of the case, Inti's evidence would have called for higher credence in one of the functions, or having positive credence in some additional hypothetical functions (I avoided such complications in step 11 for simplicity). But surely there is some set of intuitions that would justify Inti's actual answer.<sup>13</sup> Lastly, Inti's conclusion relies on the mathematical fact that avoiding

- 12 This applies also to the agreement between the hypotheses in step 9 (an agreement I do not see as an essential feature of this case). By contrast, Tarsney does not seem committed to the dependence of the relevant significance of such agreement between theories on the agent's evidence (“Intertheoretic Value Comparison”). I believe that by adopting this commitment and modifying Tarsney's approach along the lines of my proposal, the problem that Carr identifies can be avoided (“Normative Uncertainty without Theories,” 759–60).
- 13 I am ignoring views that necessitate imprecise credences in such cases. They deserve a separate discussion.



the fish has a higher expected value  $[(0.5 \times -1) + (0.5 \times -1) = -1]$  than eating it  $[(0.5 \times 1) + (0.5 \times -6) = -2.5]$ , given Inti's credal distribution.

If there is any problem with maximizing expected value in such a case, it is not the (justificatory) problem of IVC. Inti's value comparisons across the normative hypotheses are trivially justified because she justifiably constructs them as referring to the same unit of value. Inti constructs each hypothesis on the basis of one internally consistent subset of her relevant intuitive judgments. I am not trying to argue that merely starting out with indecisive normative intuitions guarantees the justifiability of such construction, or of distributing credence solely among normative hypotheses that share the same unit of value. But the fact that the source of Inti's normative uncertainty is her indecisive normative intuitions renders the justifiability of her credal distribution much more plausible than that of any credal distribution that would allow Extani\* to implement Carr's proposal.

I stress that the mere fact that Inti does not have *theories* (as such) in mind does not play an essential role here. To see this, consider the following variation on Inti's case:

Inti\*'s case is identical to Inti's, except that she conceives of each of her hypothetical functions as representing a normative theory or a *family* of normative theories that share common features.

By comparing values across hypothetical functions, Inti\* would be *ipso facto* comparing values across the associated (families of) normative theories. And these IVCs would be trivially justified, just like Inti's.

### 3. CONCLUSION

Regardless of whether the agent's normative uncertainty is about *theories* as such, the justificatory problem of IVC threatens the justifiability of maximizing expected value under normative uncertainty. However, I have argued that this justificatory problem can be solved in some cases of indecisive normative intuitions, wherein the agent constructs the relevant normative hypotheses as referring to the same unit of value.

On the one hand, my argument helps Carr by allaying the worry that her proposal could never be implemented justifiably. On the other hand, my argument raises the suspicion that implementing Carr's proposal requires the same type of epistemic state in which explicit IVC (value comparison across normative theories as such) can be justified (as in Inti\*'s case), which threatens some of the significance of Carr's proposal. (One natural way to remove this suspicion

would be to show that Carr's proposal could be implemented in cases wherein explicit IVC is *necessarily* unjustified. Carr does not show this.)

But even if this suspicion turns out to be right, we should still acknowledge the fecundity of Carr's argument. First, normatively uncertain agents who do not have normative *theories* in mind deserve our attention too. Second, Carr highlights a stage that is *necessary* for any justifiable procedure of maximizing expected value under normative uncertainty: reaching a justifiable credal distribution among candidate normative functions that share the same unit of value. Proponents of (explicit) IVC must accept the necessity of this stage and defend the possibility of reaching it. Third, Carr's argument can help *opponents* of IVC realize that when this stage is reached, value comparison across the functions is unproblematic.

Future research should investigate whether the gap between proponents and opponents of IVC may be bridged further by attending to the relevant differences that I identified between Extani\*'s and Inti\*'s types of normative uncertainty. Perhaps the proponents are correct only with respect to some cases wherein the uncertainty stems from the agent's indecisive normative intuitions, or wherein she constructs the normative hypotheses by reference to the same unit of value.

In sum, the combination of Carr's paper and mine helps to illuminate the conditions for maximizing expected value under normative uncertainty without unjustified value comparison.<sup>14</sup>

University of Toronto  
ron.aboodi@mail.utoronto.ca

#### REFERENCES

- Aboodi, Ron. "One Thought Too Few: Where *De Dicto* Moral Motivation is Necessary." *Ethical Theory and Moral Practice* 20, no. 2 (April 2017): 223–37.
- Carr, Jennifer. R. "Normative Uncertainty without Theories." *Australasian Journal of Philosophy* 98, no. 4 (2020): 747–62.
- Daniels, Norman. "Wide Reflective Equilibrium and Theory Acceptance in Ethics." *Journal of Philosophy* 76, no. 5 (May 1979): 256–82.
- Gustafsson, Johan E., and Olle Torpman. "In Defence of My Favourite Theory." *Pacific Philosophical Quarterly* 95, no. 2 (June 2014): 159–74

<sup>14</sup> This work was conducted as part of a research project of Nikkonna (registered Amuta). I thank each person who helped, including Zeev Goldschmidt, Nicholas Makins, Ittay Nissan-Rozen, Amit Pinsker, Andrew Sepielli, and an anonymous referee for this journal.

- Hedden, Brian. "Does MITE Make Right? Decision-Making under Normative Uncertainty." In *Oxford Studies in Metaethics*, vol. 11, edited by Russ Shafer-Landau, 102–28. Oxford: Oxford University Press, 2016.
- Huemer, Michael. *Ethical Intuitionism*. London: Palgrave Macmillan, 2005.
- Lockhart, Ted. *Moral Uncertainty and Its Consequences*. Oxford: Oxford University Press, 2000.
- Ross, Jacob. "Rejecting Ethical Deflationism." *Ethics* 116, no. 4 (July 2006): 742–68.
- Sepielli, Andrew. "Moral Uncertainty and the Principle of Equity among Moral Theories." *Philosophy and Phenomenological Research* 86, no. 3 (May 2013): 580–89.
- . "What to Do When You Don't Know What to Do." In *Oxford Studies in Metaethics*, vol. 4, edited by Russ Shafer-Landau, 5–28. Oxford: Oxford University Press, 2009.
- Tarsney, Christian. "Intertheoretic Value Comparison: A Modest Proposal." *Journal of Moral Philosophy* 15, no. 3 (June 2018): 324–44.