

JOURNAL *of* ETHICS & SOCIAL PHILOSOPHY

VOLUME XXI · NUMBER 2

February 2022

ARTICLES

All Reasons Are Fundamentally for Attitudes

Conor McHugh and Jonathan Way

Practical Commitment in Normative Discourse

Pekka Väyrynen

Freedom to Roam

Matthias Brinkmann

The Indefensible Self-Defense Argument

Howard Hewitt

Is Epistocracy Irrational?

Adam F. Gibbons

DISCUSSIONS

On the Normative Connection between

Paternalism and Rights

Stephanie Sheintul

Other-Sacrificing Options: Reply to Lange

Romy Eskens

JOURNAL *of* ETHICS
& SOCIAL PHILOSOPHY

VOLUME XXI · NUMBER 2

February 2022

ARTICLES

- 151 All Reasons Are Fundamentally for Attitudes
Conor McHugh and Jonathan Way
- 175 Practical Commitment in Normative Discourse
Pekka Väyrynen
- 209 Freedom to Roam
Matthias Brinkmann
- 234 The Indefensible Self-Defense Argument
Howard Hewitt
- 251 Is Epistocracy Irrational?
Adam F. Gibbons

DISCUSSIONS

- 282 On the Normative Connection between
Paternalism and Rights
Stephanie Sheintul
- 290 Other-Sacrificing Options: Reply to Lange
Romy Eskens

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

ALL REASONS ARE FUNDAMENTALLY FOR ATTITUDES

Conor McHugh and Jonathan Way

AS RATIONAL AGENTS, we are governed by reasons. The fact that there is beer at the pub might be a reason to go there and a reason to believe you will enjoy it. Facts reported on the news might be reasons to believe there is suffering in a war zone and donate to charity. Specifically, these are *normative* reasons—considerations that rationally support responses we can give in light of them. They must be distinguished from motivating reasons: reasons for which, or considerations in light of which, we respond. Motivating reasons can be normative reasons, but they need not be: the reasons for which we think or do things may not really support those responses. Hereafter by “reasons” we mean normative reasons unless indicated.

As the above examples illustrate, there are reasons for both action and belief. But although much work on reasons focuses primarily on action and belief, there are many other responses for which there seem to be reasons. In particular, many attitudes besides belief—for example, desire, regret, admiration, contempt, fear, gratitude, and blame—seem subject to reasons. This diversity raises questions about how reasons for different responses relate to each other. Might certain such reasons be more fundamental than others? Should certain reasons and not others be treated as paradigmatic?

Often, reasons for action are at least implicitly treated as the fundamental or paradigmatic case. This can be seen, for example, in the tendency of philosophers to find it puzzling that reasons for attitudes seem not to satisfy certain conditions on reasons for action. For example, it is plausible that there is a reason to act in some way only if acting in that way is under your voluntary control. But attitudes are rarely, perhaps never, under our voluntary control. And it is plausible that, at least very often, when there is a reason to act in some way, so acting will promote or realize some value. But it seems easy to imagine cases in which there is a reason for an attitude but having that attitude will not promote or realize any value. Such observations have led some to argue that reasons for attitudes do, despite appearances, meet these conditions; they have led others

to deny that there can be reasons for attitudes.¹ But the thought that these conditions must be met by all reasons seems to assume that reasons for action are paradigmatic. And we might instead question that assumption.

Our aim here is to explore and defend an alternative approach. We suggest that reasons for attitudes are fundamental, and reasons for action are both derivative and, in certain ways, idiosyncratic. More specifically, we claim that reasons for action are derivative from reasons for intention. For most of the paper we focus on this more specific claim, but we point out that if it is defensible then it lends considerable plausibility to the general claim that *all* reasons are fundamentally for attitudes.

In section 1, we outline both our general and our more specific claim. In section 2, we offer two arguments for the more specific claim, based on two plausible (though not uncontroversial) assumptions about reasons: they are potential premises of good reasoning and they are capable of being responded to. In section 3, we consider objections. In section 4, we return to the more general claim and sketch some ways in which it is significant for theorizing about reasons.

1. TWO HYPOTHESES

We will defend the following claim:

General Hypothesis: All reasons are fundamentally reasons for attitudes.

A reason for a non-attitude is *fundamentally* a reason for an attitude if what it is to be a reason for the non-attitude is to be a reason for an attitude. As we might also say, such reasons are “nothing over and above” reasons for attitudes. It is a contested question how best to understand such locutions—for instance, whether in terms of identity, constitution, grounding, or essence. We will just assume they are in good order.

According to the General Hypothesis, if there are reasons for action then they are fundamentally reasons for some attitude. How could this be? Well, reasons for action seem closely related to reasons for intention. Very often, at least, a reason for one seems to be a reason for the other. So a natural further claim for a proponent of the General Hypothesis to make is this:

1 For the first reaction, with respect to reasons for belief, see Foley, *The Theory of Epistemic Rationality*; Ryan, “Doxastic Compatibilism and the Ethics of Belief”; Steglich-Peterson, “How to Be a Teleologist about Epistemic Reasons”; Steup, “Doxastic Freedom”; and Weatherson, “Deontology and Descartes’s Demon.” For the second, see especially Maguire, “The Value-Based Theory of Reasons” and “There Are No Reasons for Affective Attitudes”; but cf. also Gluer and Wikforss, “Against Content Normativity”; and Papineau, “Normativity and Judgement.”

Special Hypothesis: Reasons for action are fundamentally reasons for intention.²

Are there other sorts of reason that challenge the General Hypothesis? Well, in ordinary talk we are happy to say that you have reason to be at the pub now (since it is past 6 p.m.), or reason to be stronger (since you are running a marathon soon), or reason to be able to speak French (since you are in Montreal). But being at the pub, being stronger, and being able to speak French are neither attitudes nor actions.

However, it is very natural to reinterpret these claims in terms of reasons for action or attitudes. You have a reason to be in the pub insofar as you have a reason to get there or to want to be there. You have a reason to be stronger, or to be able to speak French, insofar as you have reasons to desire, and to bring it about, that you are stronger or able to speak French. This reinterpretation is natural because reasons are for responding to. We cannot respond to a reason to be somewhere or to have an ability except by acting and having attitudes. So it is natural to think that what underlies talk of such reasons are reasons to act and have attitudes.

This suggests that we should be cautious about taking ordinary talk of reasons at face value. This opens the door for the Special Hypothesis. Nonetheless, it is less natural to think that you have reasons to act only insofar as you have reasons to intend—indeed, things might seem to be the other way around.³ We thus take reasons for action to pose the most significant challenge to the General Hypothesis. In the rest of this paper, we will make a case for the Special Hypothesis. Assuming that what we have just said is right, the main obstacle to acceptance of the General Hypothesis will thereby have been removed. Before presenting that case, we offer some clarifications to make the Special Hypothesis more precise.

The Special Hypothesis holds that reasons to act are fundamentally reasons to intend. There are different ways this could be so, but on the version of the Special Hypothesis we will work with:

Special Hypothesis, Simple Version: For p (a fact) to be a reason to ϕ (an act type) is for p to be a reason to intend to ϕ . For p to be a reason to *not*- ϕ is for p to be a reason to intend to *not*- ϕ .⁴

- 2 Gibbard (*Wise Choices, Apt Feelings*, 38–39), Hedden (*Reasons without Persons*, ch. 6), Portmore (*Commonsense Consequentialism*, 63), Scanlon (*What We Owe to Each Other*, 21), and Smith (“The Ideal of Orthonomous Agency, or the How and Why of Buck-Passing,” 60) endorse claims like this. Gibbard, Portmore, Scanlon, and Smith do not defend them at length. Hedden’s view, which concerns what he calls the “rational ought,” depends on a kind of mentalism that we would reject as applied to reasons.
- 3 Dancy says that “it seems just obvious” that reasons to intend derive from reasons to act (*Practical Shape*, 4).
- 4 Versions of the Special Hypothesis vary, for instance, on the relation between the act type

Talk of “actions” can refer to act types or token actions. Act types are types of things—events or perhaps processes—that can be done intentionally (including, for our purposes, omissions). Token actions are instances of these types. Not all act types are things that *must* be done intentionally: you might bring your book to work unintentionally because you left it in your bag, or press the button unintentionally while reaching for your coffee. A reason to act supports a certain act type, such as bringing the book to work or pressing the button, rather than any token action. Compliance with the reason requires that some token of the relevant type occur.

Intention is a type of attitude: a persisting psychological state with a content or object. It is closely related to action, in at least the following respects. First, intentions take act types, or propositions about them, as their objects. Second, action is the *telos* of intention. To intend is to aim to act, and to execute an intention is to act intentionally. A future-directed intention to ϕ tends to persist until it becomes present directed. A present-directed intention to ϕ manifests itself in presently trying to ϕ , which, if successful, counts as, or helps constitute, intentional ϕ -ing. Some philosophers hold, further, that intentionally ϕ -ing requires intending to ϕ , or at least some relevant intention.⁵ We will not assume this.⁶

It is sometimes claimed that forming an intention can itself be an action—a mental act of deciding. Similarly, it is sometimes claimed that beliefs can be formed through acts of judgment. Arguably, however, decisions and judgments are not actions in the sense presently at issue, since they cannot be performed intentionally.⁷ If so, the Special Hypothesis does not apply to them. Given the intimate relations between decision and intention on the one hand, and between judgment and belief on the other, we take it that reasons for decisions and judgments can instead be treated as reasons for attitudes, and so as falling under the General Hypothesis. We will not explore this matter further here; instead we focus on reasons for act types in the sense defined. Note that this includes some mental act types, such as imagining a hippopotamus or calculating the square

for which there is a reason and the content of the intention for which that reason is a reason.

- 5 The first of these claims is the “Simple View”; the second is the “Single Phenomenon View.” See Bratman, *Intentions, Plans, and Practical Reason*.
- 6 We hope these remarks on action and intention are relatively uncontroversial; they fit with, but do not presuppose, the influential “planning theory” of intention (Bratman, *Intentions, Plans, and Practical Reason*). For general discussion, see Setiya, “Intention”; and Wilson and Shpall, “Action.”
- 7 That is, you cannot intentionally decide on a particular course of action or judge a particular proposition to be true. Many authors who regard judging and deciding as acts accept this point, e.g., Hieronymi, “Controlling Attitudes”; and McHugh, “Judging as a Non-Voluntary Action.”

root of 441. The Special Hypothesis applies to reasons for such mental act types, though our examples will mainly be of physical act types.

Reasons and intentions are indexed to times, twice over. You might now have reason to go to the pub now, or now have reason to go to the pub later.⁸ Similarly, you might now intend to go to the pub now, or later. How should the Special Hypothesis accommodate this? We shall take it to say that a reason (at t_1) to (ϕ at t_2) is a reason (at t_1) to intend (at t_1) to (ϕ at t_2). We will not generally need to take care over this, though issues about time will be relevant in section 3.3.

The Special Hypothesis takes reasons to act to be nothing over and above reasons to intend. It is plausible that, as such, you are *criticizable* with respect to reasons to act only insofar as you are criticizable with respect to reasons to intend. Compare: you are criticizable for not being at the pub only insofar as you are criticizable for not going there.⁹ We grant, for the sake of argument, that the Special Hypothesis has this implication. To illustrate, compare Jack and Jill, each of whom would benefit from ϕ -ing, has no reason not to ϕ , and knows all this. Jack fails to even intend to ϕ . Jill forms and maintains the intention to ϕ , and tries her best to ϕ , but fails. Here, only Jack is rationally criticizable for his intentions. The Special Hypothesis thus implies that only Jack is rationally criticizable for his actions. This is a plausible verdict. We will call Jack irrational, silly, perverse, and so on. We might call Jill clumsy or inept, but insofar as she tried her best we will not call her irrational or the like.¹⁰

As characterized, the Special Hypothesis implies not only that all reasons to act are reasons to intend but also that all reasons to intend are reasons to act. This might be disputed because of examples such as Kavka's "toxin puzzle," where there is a large reward for merely intending to drink an unpleasant but otherwise harmless toxin.¹¹ However, the literature divides on such examples. Some hold that such considerations are not genuine reasons to intend, but rather reasons to want to intend or bring it about that you intend. Others hold that they are genuine but "wrong-kind" reasons to intend. For simplicity, we will assume the first view here. Those who prefer the second can amend the Special Hypothesis to say that reasons to act are fundamentally *right-kind* reasons to intend.¹²

8 For discussion see Kiesewetter, *The Normativity of Rationality*, 215ff.

9 Cf. Broome, *Rationality through Reasoning*. We do not deny that reasons that are derivative in weaker senses might ground criticism. For instance, you can be criticized for failing to act on an instrumental reason, even if you end up achieving the end from which this reason derives.

10 One might think that in other cases we can be rationally criticizable irrespective of our intentions. We consider this below (section 3.3).

11 Kavka, "The Toxin Puzzle."

12 For the first view, see, e.g., Gibbard, *Wise Choices, Apt Feelings*; McHugh and Way, "Fitting-

The Special Hypothesis contrasts with views that take reasons to act and reasons to intend to be independent and views that take reasons to intend to be, or at least to derive from, reasons to act. We take the latter to be the more common and plausible alternative. On the most popular version of this view, reasons to intend are instrumental: there is reason to intend to ϕ just when and because there is reason to ϕ and intending to ϕ promotes ϕ -ing.¹³ Since this seems the most natural story about how reasons to intend might derive from reasons to act, we will take it to be the main alternative to our account. Note that this view implies that not all reasons to act are reasons to intend. In a case in which intending to ϕ does not promote ϕ -ing—for instance because you are going to ϕ regardless of whether you intend to—a reason to ϕ is not a reason to intend to ϕ . We return to this below.

Another alternative to the Special Hypothesis takes reasons for *intentional action* to be prior to both reasons for action (which, recall, need not be intentional) and reasons to intend. For instance, perhaps to be a reason to ϕ is to be a reason to intentionally ϕ . For simplicity, and because we do not know of any defenses of this view, we shall mostly ignore it.¹⁴ However, our arguments for the Special Hypothesis might be developed to support this view, given certain further assumptions (see especially section 2.2). It would also face some of the same objections as the Special Hypothesis, and can make some of the same replies. So the paper's arguments should be of interest to those considering such a view.

2. ARGUMENTS FOR THE SPECIAL HYPOTHESIS

The example of Jack and Jill provides some initial motivation for the Special Hypothesis. In this section we offer two further arguments for it.

Our arguments will make two assumptions about reasons. First, we assume that it is a central truth about reasons that they are potential premises of good reasoning. Second, we assume that reasons are subject to a *response constraint*:

ness First"; Parfit, *On What Matters*; and Shah "How Action Governs Intention." For the second, see, e.g., Hieronymi, "The Wrong Kind of Reason"; Howard, "The Fundamentality of Fit"; and Schroeder, *Slaves of the Passions*.

The Special Hypothesis may require further amendments too. For example, if you think "reason to ϕ " implies "can ϕ " but "reason to intend to ϕ " does not, you might add the condition that the agent can perform the action. For simplicity we ignore this possibility.

¹³ Cf. Heuer, "Reasons to Intend"; Pink, *The Psychology of Freedom*; and Kolodny, "The Myth of Practical Consistency."

¹⁴ However, the view may appeal to defenders of the view that intentional action is prior to intention. See, e.g., Thompson, *Life and Action*; McDowell, "Some Remarks on Intention in Action"; and, for discussion, Setiya, "Intention."

normative reasons must be capable of being responded to—of being our motivating reasons. These assumptions are popular and plausible.¹⁵ Both are motivated by the attractive idea that the point of reasons is to guide us. The way in which reasons most explicitly guide us is through reasoning: thus we should expect reasons to be potential premises of good reasoning. And a reason you cannot respond to is not one that can guide you. While neither assumption is uncontentious, we cannot defend them further here; instead we aim to show that if they are granted, they naturally lead to the Special Hypothesis.

2.1. *The Response Constraint*

Our first argument has two parts. We first argue that the response constraint supports the claim that any reason to ϕ is also a reason to intend to ϕ . We then show that this claim supports the Special Hypothesis.

The response constraint tells us that a reason for acting is a reason that can be acted on: if p is a (normative) reason for you to ϕ , you can ϕ for the (motivating) reason that p .¹⁶ But acting for a reason, we suggest, entails acting from an intention. In particular, you act for a reason only if you execute an intention held for (based on) that reason.¹⁷ To illustrate, suppose you intend to take a book to work with you, for the reason that the reading group meets this afternoon (that is, your intention is held for this reason—what you intend is not to take-the-book-for-this-reason, but simply to take the book). If you take the book by executing this intention then you take the book for the reason that the reading group meets this afternoon. But if you take the book without executing this intention, you do not. For example, if you take the book by picking up your bag, which, unbeknownst to you, contains the book, then, although you take the book, and intend to do so for the reason that the reading group meets this afternoon, you do not take the book for the reason that the reading group meets this afternoon.

15 For versions of the first assumption, and responses to objections, see Asarnow, “Rational Internalism”; Hieronymi, “The Wrong Kind of Reason”; McHugh and Way, “Fittingness First”; Paakkunainen, “Can There Be Government House Reasons for Action?”; Setiya, *Reasons without Rationalism*; Schroeder, *Slaves of the Passions*; Silverstein, “Reducing Reasons”; Williams, “Internal and External Reasons.” For versions of the second, and responses to objections, see Kiesewetter, *The Normativity of Rationality*; Kolodny, “Why Be Rational?”; Lord, *The Importance of Being Rational*; Parfit, *On What Matters*; Raz, *From Normativity to Responsibility*; Shah, “A New Argument for Evidentialism”; Way and Whiting, “Reasons and Guidance (or, Surprise Parties and Ice Cream)”; and Williams, “Internal and External Reasons.”

16 In the rest of this section, we generally allow context to disambiguate between motivating and normative reasons.

17 Cf. Davidson, “Intending”; Setiya, *Reasons without Rationalism*; and Hieronymi, “Reasons for Action.”

Thus, acting for a reason entails acting from an intention held for that reason. But in general, if you base an attitude on a consideration, you treat that consideration as a normative reason for that attitude. So in intending to ϕ for the reason that p , you treat p as a reason to intend to ϕ . Furthermore, if it is a reason to ϕ then it is appropriate that you treat it in this way.¹⁸

Thus, if p is a reason to ϕ then p can be appropriately treated as a reason to intend to ϕ . But if it is appropriate to treat something as a reason to intend to ϕ , then it *is* a reason to intend to ϕ . Reasons are considerations that can appropriately play certain roles in our thinking and reasoning. To treat a consideration as a reason is to treat it as an appropriate occupant of such a role. If it is appropriate to treat it in this way, then it *is* an appropriate occupant of this role—that is, it is a reason. (Compare: if a precedent is appropriately treated as legally authoritative, then it *is* authoritative.)

Let us sum up the argument thus far. Reasons for action can be acted on. You act on a reason only if you intend for that reason. In doing so you treat it as a reason for intention, and appropriately so. But a consideration that can be appropriately treated as a reason for intention is one. So a reason to ϕ is also a reason to intend to ϕ .

If correct, this shows that all reasons to act are also reasons to intend. The Special Hypothesis offers a straightforward explanation: all reasons to act are reasons to intend because to be a reason to act is to be a reason to intend. Of course, by itself this only provides limited support for the Special Hypothesis, since this generalization might be explained in other ways. However, as we noted, the view that reasons to intend are instrumental, which appears to be the main alternative to the Special Hypothesis (section 1), is committed to denying, and thus cannot explain, the fact that all reasons to act are reasons to intend. As we discuss in more detail in section 3, there are cases in which intending to ϕ does not promote ϕ -ing. In these cases, a reason to ϕ is not a reason to intend to ϕ , on this view. Thus the Special Hypothesis has an important advantage over its main alternative.

Before moving on, we note one objection to this argument. We claimed that acting for a reason entails acting on an intention held for a reason. This claim is reminiscent of the “Simple View,” that intentionally ϕ -ing entails intending to ϕ , and so might be thought to face similar objections. For instance, Bratman suggests that foreseen but unintended consequences can be intentional: a mara-

18 It may not be appropriate to intend to ϕ for the reason that p , if p is an outweighed reason to ϕ . However, it will be still be appropriate to treat p as a reason to intend to ϕ . For instance, you might consider p in deliberation about whether to ϕ . Since practical deliberation aims at intention (see section 2.3), you thereby treat p as bearing on whether to intend to ϕ . More generally, if you treat p as bearing on reasoning that aims at a response R , you treat p as bearing on R .

thon runner might intentionally wear down the soles of their shoes, despite not intending to do so, if they are aware that they are doing so.¹⁹ However, whether or not this is right, the case is no counterexample to our claim, since it is clear that the runner does not run down the soles of their shoes for a reason.

Bratman's central argument against the Simple View turns on his video-game example. Here you are playing a game in which you simultaneously try to hit two targets—using each hand to guide a missile—although you know you cannot hit both. Bratman suggests that if you hit one of the targets—target 1—you do so intentionally, although you did not intend to hit target 1. Here again, we suggest, it is dubious that you hit target 1—as opposed to hitting one of the targets—for a reason. But it is also unclear that you do not intend to hit target 1. As others have observed, you are guided by a motivational state with the functional role of intention. We might thus take the case as one in which it is rational to have inconsistent intentions.²⁰ Either way then, the video-game example is not a compelling case of acting for a reason without acting on an intention held for a reason.²¹

2.2. *Reasons and Reasoning*

Our second argument aims to provide support for the Special Hypothesis through attention to the way reasons feature in reasoning. In theoretical reasoning, you aim to settle on a view of what is true—to form a belief. But you do this by considering the question of what is true—or, more specifically, of whether, say, it will rain. Thus you consider evidence bearing on this question—e.g., that clouds are gathering. If, on the basis of this evidence, you answer the question affirmatively, you thereby believe it will rain. But you do not need to think about this belief—you can just think about the prospects of rain.²²

19 Bratman, *Intentions, Plans, and Practical Reason*, 123.

20 McCann, "Settled Objectives and Rational Constraints"; and Kolodny, "The Myth of Practical Consistency."

21 Hornsby ("On What's Intentionally Done") and Heuer ("Intentions and the Reasons for Which We Act") offer putative cases of acting for reasons without acting from an intention. Hornsby focuses on "emergency cases." If a snowball is about to hit you, she suggests, you might catch it for a reason (that it is about to hit you) without intending to do so. However, forming an intention need not take long or precede the start of action. If catching the ball is not a mere reflex—if you meant to do it—then it does seem that you must have intended to do it (cf. Heuer, "Reasons to Intend"). Other putative cases involve intentions that would be in some way counterproductive. Since such cases might also be thought of as counterexamples to the Special Hypothesis, we discuss them later (section 3).

22 You can also revise an attitude as a result of higher-order reasoning about it, e.g., by considering whether your evidence is sufficient to justify belief. But in doing so you most directly revise your higher-order attitudes. It is a further step to revising your first-order attitude.

In practical reasoning, you aim to settle on a course of action—to form an intention. But you do this by considering the question of what to do—or, more specifically, of whether, say, to go to the pub. Thus you consider factors bearing on this question—e.g., that your friends are at the pub. If, on the basis of such factors, you answer the question affirmatively, you thereby intend to go to the pub. But you do not need to think about this intention—you can just think about the pros and cons of going to the pub.²³

In both theoretical and practical reasoning, then, there is a distinction between a response to which the reasoning leads and the object of that response. In the theoretical case, the response is a belief and its object is a proposition. In the practical case, the response is an intention and its object is an act type. Thus, as far as their role in reasoning goes, practical reasons seem to stand to intentions, rather than to act types, as theoretical reasons stand to beliefs; they stand to act types as theoretical reasons stand, not to beliefs, but to propositions. In the theoretical case, it is clear that what reasons support is fundamentally beliefs, not propositions. If we take their role in reasoning to be an important aspect of the nature of reasons, then, we should regard practical reasons as fundamentally supporting intentions rather than act types. In doing so we get an attractive, unified picture of how (fundamental) reasons feature in reasoning: a reason for a response is a premise of good reasoning that concludes in that response, reasoning that is about the object of the response. By contrast, if we regard practical reasons as fundamentally supporting act types, then we find deep differences between practical and theoretical reasons in respect of the way they relate, in reasoning, to the responses they support. Such differences seem hard to reconcile with the idea that their role in reasoning is an important aspect of the nature of reasons.

Our argument here might seem to ignore the Aristotelian thesis that the conclusion of practical reasoning is an action.²⁴ However, the central alternative to the Special Hypothesis is that practical reasons fundamentally support act types—i.e., types of events or processes that can but need not be intentional. But if the Aristotelian thesis is right, *intentional* actions can be conclusions of practical reasoning. Mere actions cannot be conclusions of practical reasoning

23 In characterizing practical reasoning this way, we do not deny that there is such a thing as reasoning about what you ought or have reason to do. We class this as theoretical reasoning, although it can—and should but does not always—inform practical reasoning (cf. Broome, *Rationality through Reasoning*, 250; Harman, *Change in View*, 77). Nor do we deny that in practical reasoning you might also aim to act—see below (this section).

24 For discussion, see Fernandez, “Practical Reasoning”; Dancy, *Practical Shape*, 4; Tenenbaum, “The Conclusion of Practical Reason”; and Paul, “The Conclusion of Practical Reasoning.”

by anyone's lights. Irrespective of the Aristotelian thesis, then, we can conclude that practical reasons are not fundamentally reasons for mere actions, and thus that the Special Hypothesis should be preferred to its central alternative.²⁵

3. OBJECTIONS AND REPLIES

We have made a *prima facie* case for the Special Hypothesis. We now turn to considering some objections. First, we will consider the objection that the Special Hypothesis wrongly puts the focus of our practical lives on intention rather than action (section 3.1). Second, we will address some putative counterexamples: cases where it might seem plausible to attribute a reason for action but not a reason for intention (section 3.2). Third, we will consider an objection to what we earlier granted is an implication of the Special Hypothesis: that rational criticizability for action depends on rational criticizability for intention. In some cases, it might seem, your actions make you rationally criticizable regardless of your intentions (section 3.3).

3.1. *The Wrong Focus?*

Much of our ordinary practical deliberation, our talk about practical reasons, and our practices of holding each other responsible focus on actions rather than intentions. The Special Hypothesis might seem incompatible with this fact. In taking reasons to act to hold in virtue of reasons to intend, it might seem to give practical reasons the wrong focus, or to suggest that what ultimately matters is what we intend rather than what we do.

However, this concern is misguided. First, reasons for intention are provided by features of actions—for instance, what gives me a reason to intend to go the pub is that going to the pub allows me to see my friends. It is thus not surprising that these features should be the focus of attention in practical deliberation, just as evidence for propositions, rather than beliefs as such, is the focus of ordinary theoretical deliberation. Second, practical deliberation in its basic form is framed by the question of what to *do*, rather than what to intend, just as theoretical deliberation is framed by the question of what is the case, rather than what to believe. We form intentions in deliberation by answering the question of what to do, just as we form beliefs in deliberation by answering the question of what is true. Third, the Special Hypothesis does not entail that what we intend *matters* more than how we act. It is a claim about what is metaphysically fundamental,

²⁵ We thus acknowledge that our argument here might be used to support the view that practical reasons are fundamentally for intentional actions. But this would require a defense of the Aristotelian thesis.

not about what is more important, or what we have more reason to care about. It might be that what matters most is how we act, even though reasons to act derive from reasons to intend, and even though we are criticizable for our actions only insofar as we are criticizable for our intentions.

3.2. *Reasons to Act without Reasons to Intend?*

Some philosophers hold that there are reasons to act that are not reasons to intend. We will discuss two common kinds of examples.

3.2.1. *Pointless Intentions*

Heuer suggests that, if you will do something regardless of whether you intend to do it, then you can have reasons to do it but no reasons to intend it.²⁶ Her examples include reasons not to kill and reasons not to go into space without a spacesuit. The thought seems to be that since you are already going to do the thing there is no point in forming the intention. Examples need not be limited to omissions: for instance, you might have reason to exercise today but since you will get exercise by walking to work, you need not intend to exercise.²⁷

But this is not convincing. Heuer's claim here must be that you have *no* reason for such intentions. But suppose you deliberate about whether to kill someone, whether to go into space without a spacesuit, or whether to exercise today. Clearly, there are relevant considerations that might rationally lead you to intend not to kill anyone, not to go into space without a spacesuit, and to exercise. You would be making no mistake in basing such intentions on these considerations. In doing so, you would be treating them as reasons for so intending. But, again, something that can be appropriately treated as a reason is one. So, even if you are not required to have these intentions, it is not true that you lack reasons for them.

Any impression to the contrary may be due to a tacit assumption that reasons for intentions are instrumental from reasons to act. But, as we explained in section 1, this is a substantive view, and it is one that the foregoing considerations count against. In deliberation about whether to ϕ , it can be fine to base an inten-

26 Heuer, "Reasons to Intend."

27 Heuer suggests that habitual actions might also be examples ("Reasons to Intend"). However, as she allows, it is less clear that action from habit is done without intention. We therefore focus on the clearer examples. For discussion of habitual action see Owens, "Habitual Agency." We take the points in this section to also answer further putative counterexamples that could be drawn from Heuer's discussions, including cases in which you have a reason to do something but it does not matter whether you do it intentionally ("Reasons to Intend," 877), and cases where a complex action would be impeded by intending each part of it ("Intentions and the Reasons for Which We Act," 306).

tion to ϕ on the worthwhile features of ϕ -ing, even if intending does not make it more likely that you will ϕ , because you are going to ϕ anyway.

Another potential source of confusion might be a failure to distinguish reasons for intending to ϕ from reasons for deliberating about whether to ϕ . You may have no reason to deliberate about, say, whether to kill. Even if so, it does not follow that you have no reason to intend not to kill. It is a general feature of reasons for attitudes that you can have them even though you have no reason to deliberate. Deliberation demands time and resources that you should invest carefully by directing them to important questions that cannot be left to quick and automatic cognition. You have reasons bearing on many questions that do not call for this investment.

Heuer's argument thus fails. But her examples also form the basis for a different objection to the Special Hypothesis. Even if it is accepted that there are reasons to intend not to kill and not to go into space without a spacesuit, there still seems to be a difference in the overall normative status of these intentions and the corresponding omissions. You *ought* not go into space without a spacesuit and you *ought* not kill. But it is okay not to intend these things. And since what you ought to do is, very plausibly, determined by reasons, it follows that there must be reasons to act that do not derive from reasons to intend—for otherwise, the overall status of actions and the corresponding intentions could not come apart.

However, this argument is too quick. If you ought to ϕ but need not intend to ϕ , this need not be because there are reasons to ϕ that are not reasons to intend to ϕ . Indeed, on the face of it, this is not what is going on in these cases. The factors that seem to make it okay not to intend not to kill, and not to intend not to go into space without a spacesuit, are that forming these intentions may require deliberation or conscious attention. This is a use of valuable cognitive resources, which is here a cost that is not compensated for by the need to form an intention. By contrast, the actions of not killing, and not going into space without a spacesuit, do not require deliberation or conscious attention. These factors—the costs of deliberation and attention—are not plausibly understood as reasons to ϕ that are not reasons to intend to ϕ . For instance, that forming an intention not to kill would be costly is not a reason not to kill.

How should we understand these factors and the role they play in explaining the overall status of action and intention? One possibility is that they are reasons against intending to ϕ that are not reasons against ϕ -ing. The Special Hypothesis is compatible with the existence of such reasons, since it takes reasons for and against acting to be reasons *for* intending (to act or not to act). However, an appeal to such reasons might seem in tension with the way we have argued for

the Special Hypothesis.²⁸ Fortunately, there are other possibilities. Perhaps such factors affect the overall status of intention indirectly by providing reasons to form no intention on the matter. Or perhaps it remains true in these cases that you ought to intend, but it is also true that you ought not deliberate, and will not intend unless you deliberate. If so, you may not be criticizable, even though you lack an intention you ought to have.

Each of these suggestions offers a way to explain seeming differences between the overall statuses of actions and intentions without appealing to reasons for action that are not reasons for intention. Properly assessing them would require an account of how considerations determine what you ought to do and intend, something we cannot provide here. For present purposes the point is just that it is only on a non-mandatory, and indeed implausible, treatment of the target cases that they represent counterexamples to the Special Hypothesis.

3.2.2 *Self-Undermining Intentions*

Sometimes an intention to act can be self-undermining, in the sense that acting on that intention is incompatible with doing what it is an intention to do. For example, if you intend to act modestly then you will fail, since acting modestly is incompatible with acting on such a self-regarding intention. Similarly, acting spontaneously seems incompatible with acting on an intention to be spontaneous. Heuer suggests that in these cases, too, you can have reason to act but no reasons to intend to act.²⁹

But here we need to be careful about exactly what reasons it is plausible to attribute to agents. Take modesty. You can have a reason to do what would in fact be modest in a particular situation—to share credit for some achievement, say. And in doing so you might manifest the virtue of modesty. But the reason to act that way will not be that doing so is modest. It will be, for example, that others contributed to the achievement and should get due credit. Nor will this, or any other consideration, be a reason for doing what is modest as such. It will just be a reason for whatever the act type is, e.g., sharing credit. And it will also

28 In particular, our second argument takes reasons to be potential premises of reasoning and practical reasoning to be reasoning about what to do (section 2.2). It is not clear how reasons against intention that are not reasons against acting could feature in such reasoning. For relevant discussion, see Schroeder, “The Ubiquity of State-Given Reasons”; Hieronymi, “The Use of Reasons in Thought (and the Use of Earmarks in Arguments)”; and Shah and Silverstein, “Reasoning in Stages.”

29 Heuer, “Reasons to Intend.” There are also cases where intending to act undermines the putative *reason* to act—as when you will get a reward for falling over, but only if you do so unintentionally. We would treat such cases in the same way, so we do not discuss them separately.

be a reason for intending this act type. So there are no plausible reasons here that constitute counterexamples to the Special Hypothesis.

This is not to deny that there are modesty-related reasons. For instance, you can have reasons to want to be modest, and to cultivate your modesty. But the intention to cultivate your modesty is not self-undermining.

What about spontaneity? We can say similar things. A “reason to do something spontaneous” might be a reason to choose certain sorts of options in your situation. Or it might be a reason to choose without deliberating at length. There are also reasons to desire and cultivate spontaneity. None of these responses involve self-undermining intentions. There is thus no obvious need to posit reasons that would be counterexamples to the Special Hypothesis.³⁰

We take this way of treating putative reasons for modesty and spontaneity to be plausible on its face. Note that it is also supported by the response constraint: the condition that reasons must be such that they can be responded to (section 2.1). If the intention to ϕ would be self-undermining, a reason to ϕ cannot be responded to.³¹

3.3. Criticizability

Reasons matter. When they are conclusive, you are ordinarily criticizable for failing to comply with them. As noted in section 1, the Special Hypothesis seems to imply that you are rationally criticizable for your actions only insofar as you can be rationally criticized for your intentions. We suggested that this implication is plausible.

This might be doubted. Suppose that you promise to feed your neighbor’s cat. If all else is equal, this gives you conclusive reason to feed your neighbor’s cat. And so, if all else is equal, you will be criticizable if you fail to feed your neighbor’s cat. This might seem to be so regardless of your intentions. It is not enough

30 Cf. Smith, “The Ideal of Orthonomous Agency, or the How and Why of Buck-Passing,” 62–63.

31 For further relevant discussion, see Paakkunainen, “Can There Be Government House Reasons for Action?” Pink offers a different kind of counterexample to the Special Hypothesis (*The Psychology of Freedom*). He argues that *the fact that you now intend to ϕ at a later time t* can itself generate reasons to ϕ at t , since it makes it likely that between now and t you will stake things on your ϕ -ing at t (e.g., by making plans whose success depends on your ϕ -ing at t , and inviting others to do so; cf. Kolodny, “The Myth of Practical Consistency”). Yet the mere fact that you have this intention cannot give you a reason to have it. While we lack the space to discuss this in detail, we suggest that the fact, if it is one, that things will be staked on your ϕ -ing at t is a reason both for ϕ -ing and for intending to ϕ . That seems plausible: if failing to ϕ threatens disaster, then you have a reason to intend to ϕ . However, if you are deliberating about whether to ϕ then you do not yet have the settled intention to ϕ , so no reasons of the sort Pink describes are generated.

just to intend to feed the cat: you also have to execute this intention. Thus the Special Hypothesis does not capture the significance of reasons to act.

This objection requires that there are cases in which you intend as you have conclusive reason to intend, fail to act as you have conclusive reason to act, and thereby merit criticism. In the remainder of this section, we will argue that it is unclear that there are such cases. We will do so by considering different ways in which you might fail to feed your neighbor's cat, despite intending to do so.

Suppose that you now intend to feed the cat but then later change your mind, or forget, and so do not feed the cat. You might then be criticizable for failing to feed the cat, despite having intended to do it. But the Special Hypothesis can explain this. As noted (section 1), reasons are indexed to times, which may be earlier than the time of the response that is called for. You might now have reason to feed the cat later, or you might now have reason to feed the cat now. But you merit criticism for failing to do what you have reason to do only once the time for the response has come. If you have conclusive reason to feed the cat later, then you do not merit criticism just because you are not yet feeding the cat. You merit criticism only if you have conclusive reason to feed the cat now. But in the case at issue, you lack the right intention at the time you are meant to be feeding the cat. Thus the Special Hypothesis can explain why you are criticizable.

This shows that any problematic cases for the Special Hypothesis must be cases in which you fail to act on an intention to act at that very time.³² But there might seem to be such cases. You might intend to feed the cat at midnight but fail to do so because you are not aware that it is midnight. Or you might intend to feed the cat at midnight and try to do so but fail—perhaps something prevented you, or perhaps you needed a bit of luck to succeed, which was not forthcoming.

To address such cases, we need the notion of having control over whether you ϕ . Roughly, where ϕ -ing is an act type, you have control over whether you ϕ if trying to ϕ (or not to do so) will ensure that you ϕ (or do not ϕ). You lack control over whether you ϕ if this is not so. If you lack control over whether you ϕ then it is not up to you whether you ϕ . This might be because you cannot ϕ at all, or cannot but ϕ . More interestingly, it might be because, although you can ϕ , trying to ϕ will not ensure that you do so—you might need a bit of luck.³³

32 Note that such an intention need not be a present-directed intention, i.e., an intention directed at the present, conceptualized as such.

33 For trying to ϕ to ensure that you ϕ is not for it to be necessary that, if you try to ϕ , you ϕ . A weaker relation is required—perhaps that in nearby worlds you ϕ if you try to ϕ . Note also that by “trying” we mean “trying appropriately.” This is because ϕ -ing is not necessarily out of your control if you might fail to ϕ despite a weak attempt, or because you try “too hard,” or in the wrong way. We take it that failed attempts of this sort will reflect potentially criticizable features of underlying attitudes (cf. the discussion of tracing in the main text). It will,

Consider now the case in which you intend to feed the cat and try to do so but fail. In this case, feeding the cat was not under your control. But, ordinarily, if you lack control over whether you do something, you cannot be rationally criticized for failing to do it (though you might be criticized for failing to try). The fact that it was not up to you whether to feed the cat excuses you for failing to do so, so long as you tried your best.

Of course, a lack of control does not always excuse. In particular, if you are responsible for the lack of control—perhaps because you are drunk, or failed to take earlier opportunities to ensure that feeding the cat would later be under your control—then you may still be criticizable. But this is an example of the familiar phenomenon whereby current criticizability can be traced to earlier mistakes—as when the drunken driver is criticizable for hitting the pedestrian, not because he could have done anything about it at the time, but because he decided to drive when drunk.³⁴ It is plausible that criticizability of this sort requires the agent to have had erroneous intentions at the earlier time.

The phenomenon of tracing shows the need for an important clarification of our account. While rational criticizability for failing to ϕ must derive from rational criticizability for (lacking) some intention, it need not be the intention to ϕ . It may derive from criticizability for lacking an intention to ψ , where ψ -ing would have facilitated ϕ -ing. We do not take this to be a problem for the Special Hypothesis, but rather a plausible upshot that emerges when we combine it with considerations about tracing.³⁵

What about a case in which you fail to act because of ignorance—e.g., of what the time is, or of which cat is your neighbor's? Well, if you are unaware of your ignorance or you lack control over whether you correct it, then you also lack control over whether you feed the cat. You might luckily manage to do it, but trying does not ensure that you will. Again, then, if you fail to feed the cat, you will not merit criticism, unless you are responsible for your lack of control.

If, on the other hand, you are aware of your ignorance and it is under your control whether to correct it, then, at least if you are rational, you will correct your ignorance and try to feed the cat. If you do not, that will be because there is

of course, be a difficult matter to specify precisely the required notion of trying. For relevant discussion, see, e.g., Marušič, *Evidence and Agency*, chs. 1, 6; and Portmore, "Opting for the Best," ch. 3.

34 E.g., Fischer and Tognazinni, "The Truth about Tracing."

35 We take these points to help with cases where an agent is rationally criticizable for an unintended outcome brought about through negligence or recklessness. Although in such cases the agent does not intend the bad outcome, they will be criticizable only if their intentions are criticizable in some respect.

an intention that you lack, e.g., to correct your ignorance, and your criticizability for failing to feed the cat will be traceable to it.

The qualification “if you are rational” suggests that there might be a further kind of case. Might you not intend to feed the cat at midnight, know that it is midnight, have it under your control whether to feed the cat, and yet fail to do so because you are irrational—perhaps, for instance, you are weak willed, or lazy, or lose your nerve?

This worry can be reinforced with a different example. Suppose you have conclusive reason to get out of bed now—you are already running late. Still, even if you intend to get out of bed you might not. And not because of ignorance or a lack of control—rather, you just fail to get yourself to do what you intend. This might seem irrational—and thus criticizable—but nonetheless possible, indeed common.

We think that paying attention to time helps here too. An intention “to get out of bed” could either be an intention to get out of bed at a later (though perhaps imminent) time, or an intention to be getting out of bed right now. As we have seen, cases in which you have an intention to act at a later time are not problematic. You can only be criticized for failing to act on a reason to act once the time for action has come.

But could you intend to be getting out of bed now and yet not be getting out of bed now? We think it is hard to make sense of this possibility, assuming that getting out of bed is under your control. As suggested earlier (section 1), it is plausible that there is a necessary connection between present-directed intention and trying: if you are not now trying to ϕ then you do not intend to be ϕ -ing, even if you think you do, and even if you did until a moment ago.³⁶ But if it is in your control to move now, and you try to move, you will move. In support of this, note that the conventional expression of a present-directed intention to ϕ is “I am ϕ -ing.” If you are aware you are not moving, and you are not even trying to move, you cannot sincerely say “I am getting out of bed.”

We conclude that there are no clear cases in which you intend as you have conclusive reason to intend (at the appropriate time), fail to act as you have

36 Cf., e.g., Bratman, *Intentions, Plans, and Practical Reason*, 16; Broome, *Rationality through Reasoning*, 151–52; O’Shaughnessy, “Trying (as the ‘Mental Pineal Gland’),” 380–83; and McDowell, “Some Remarks on Intention in Action,” 3. Asarnow takes the phenomenon of not getting out of bed to be a counterexample to claims of this sort (“On Not Getting Out of Bed”). While we agree with Asarnow that there is a familiar experience of, as he puts it, “irrationally dawdling in bed,” we think that it is not obvious that it involves a failure to act on an intention to be getting out of bed rather than a failure to act on a series of intentions to get out of bed imminently. Unfortunately we lack the space to engage with Asarnow’s discussion in more detail.

conclusive reason to act, and are thereby criticizable. Thus we have not seen a compelling case that the Special Hypothesis cannot explain the significance of reasons for action. Rather, it seems that in all cases in which you are rationally criticizable for failing to act, there is also a problem with your intentions. The Special Hypothesis can thus maintain that this explains why you are criticizable.

4. UPSHOTS

We have been defending the Special Hypothesis, the claim that reasons for action are fundamentally reasons for intention. After some discussion of what the claim amounts to, we presented two arguments for it and responded to several objections. This also amounts to a case for the General Hypothesis, that all reasons are fundamentally reasons for attitudes. As we began by noting, the General Hypothesis runs contrary to much theorizing about reasons. It thus has upshots for a range of issues. In closing, we indicate some of these.

Some theorists hold that reasons are always grounded in some value possessed or promoted by the responses they support.³⁷ This claim has *prima facie* plausibility when applied to reasons for action. When you have a reason to go to the pub or donate to charity, that is plausibly because of some good that is likely to come of it. The claim is less plausible when applied to reasons for attitudes: it seems you have reasons to believe truths and fear dangers even when no good will come of it. However, if the Special Hypothesis is correct, then the claim even gets reasons for action importantly wrong. According to the Special Hypothesis, reasons for action are fundamentally reasons for intention. And reasons for intention are not based in the value of intending. They might often be based in the value of the action intended, but even then the reasons for intending need not be conditional on the intention's promoting this value. As we saw with the case of intending not to enter space without a spacesuit, you can have reasons to intend even when intending does not promote doing what you intend, and so need not itself be of value. Thus, even reasons for action are not grounded in the value of the responses they fundamentally support.

This illustrates how theorizing about reasons can be shaped by the sorts of examples we start from. If we start from action as the paradigm case, the view that reasons are grounded in value will seem plausible. However, when we consider reasons for attitudes, it seems more natural to explain them in terms of their role in good reasoning or the fittingness conditions of attitudes.³⁸ If the General Hy-

37 E.g., Maguire, "The Value-Based Theory of Reasons"; Finlay, *Confusion of Tongues*; and Wedgwood, *The Value of Rationality*.

38 Cf. Hieronymi, "The Wrong Kind of Reason"; Howard, "The Fundamentality of Fit";

pothesis is correct, accounts of the latter sort have better prospects for capturing the fundamental nature of reasons.

Another upshot concerns the connections between reasons and control. When we focus on actions, we seem to find connections between reasons and *voluntary* control. You can have reasons to go to the pub but not to, say, win the lottery, because only the former is under your voluntary control. And it is plausible that you are rationally criticizable for not going to the pub only insofar as doing so was under your voluntary control. However, it is not a general feature of reason-governed responses that they are under voluntary control: belief is not under voluntary control, nor is, for instance, desire, admiration, or fear. Nor, importantly, is intention. We do not form and revise intentions by way of intentions to do so. Rather, like other reason-governed attitudes, intention is under a more fundamental, nonvoluntary form of control, in which we directly form and revise our attitudes in response to reasons.³⁹ If the General Hypothesis is correct, this suggests that it is this nonvoluntary *attitudinal* control that is most fundamentally relevant to reasons and rational criticizability. Voluntary control matters for action because it is by executing intentions, and thus exercising voluntary control, that we respond to reasons for actions. But that is an idiosyncratic case, not a paradigmatic one.

More broadly, the General Hypothesis highlights the fundamental role of attitudes in our constitution as rational, responsible, world-engaged agents. Attitudes constitute our orientation toward the world, in which we can get things right or wrong, be justified or not. Our actions change the world in more or less beneficial ways—but it is only insofar as they express or embody our orientation toward the world that they fall within the realm of reason and normativity.⁴⁰

University of Southampton
c.mchugh@soton.ac.uk
j.way@soton.ac.uk

McHugh and Way, “Fittingness First”; Setiya, *Reasons without Rationalism*; and Thomson, *Normativity*.

39 Hieronymi, “Controlling Attitudes”; and McHugh, “Attitudinal Control.”

40 For comments and discussion, we are grateful to Samuel Asarnow, Olivia Bailey, Tom Dougherty, Luke Elson, Giulia Felappi, Alex Gregory, Marie Guillot, Grace Helton, Brad Hooker, Yair Levy, Richard Rowland, David Shoemaker, Philip Stratton-Lake, Daniel Whiting, Daniel Wodak, several anonymous referees, and audiences at Cardiff University, Humboldt University Berlin, NYU Abu Dhabi, Ryerson University, Tulane University, the University of Essex, the University of Reading, the University of Southampton, and the University of Zurich. Work on this paper was supported by a fellowship at the Murphy Institute at Tulane University.

REFERENCES

- Asarnow, Samuel. "On Not Getting Out of Bed." *Philosophical Studies* 176, no. 6 (June 2019): 1639–66.
- . "Rational Internalism." *Ethics* 127, no. 1 (2016): 147–78.
- Bratman, Michael. *Intentions, Plans, and Practical Reason*. Cambridge, MA: Harvard University Press, 1987.
- Broome, John. *Rationality through Reasoning*. Oxford: Blackwell, 2013.
- Dancy, Jonathan. *Practical Shape: A Theory of Practical Reasoning*. Oxford: Oxford University Press, 2018.
- Davidson, Donald. "Intending." In *Philosophy of History and Action*, Philosophical Studies Series in Philosophy 11, edited by Yirmiahu Yovel, 41–60. Dordrecht: Springer, 1978.
- Fernandez, Patricio A. "Practical Reasoning: Where the Action Is." *Ethics* 126, no. 4 (July 2016): 869–900.
- Finlay, Stephen. *Confusion of Tongues: A Theory of Normative Language*. Oxford: Oxford University Press, 2014.
- Fischer, John Martin, and Neal A. Tognazinni. "The Truth about Tracing." *Noûs* 43 (September 2009): 531–66.
- Foley, Richard. *The Theory of Epistemic Rationality*. Cambridge, MA: Harvard University Press, 1987.
- Gibbard, Allan. *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press, 1990.
- Glüer, Kathrin, and Åsa Wikforss. "Against Content Normativity." *Mind* 118, no. 469 (January 2009): 31–70.
- Harman, Gilbert. *Change in View: Principles of Reasoning*. Cambridge, MA: MIT Press, 1986.
- Hedden, Brian. *Reasons without Persons*. Oxford: Oxford University Press, 2015.
- Heuer, Ulrike. "Intentions and the Reasons for Which We Act." *Proceedings of the Aristotelian Society* 114 (2014): 291–316.
- . "Reasons to Intend." In *The Oxford Handbook of Reasons and Normativity*, edited by Daniel Star, 865–90. Oxford: Oxford University Press, 2018.
- Hieronymi, Pamela. "Controlling Attitudes." *Pacific Philosophical Quarterly* 87, no. 1 (March 2006): 45–74.
- . "Reasons for Action." *Proceedings of the Aristotelian Society* 111 (2011): 407–27.
- . "The Use of Reasons in Thought (and the Use of Earmarks in Arguments)." *Ethics* 124, no. 1 (October 2013): 114–27.

- . “The Wrong Kind of Reason.” *Journal of Philosophy* 102, no. 9 (September 2005): 437–57.
- Hornsby, Jennifer. “On What’s Intentionally Done.” In *Action and Value in Criminal Law*, edited by Stephen Shute, John Gardner, and Jeremy Horder, 55–74. Oxford: Clarendon Press, 1993.
- Howard, Christopher. “The Fundamentality of Fit.” In *Oxford Studies in Metaethics*, vol. 14, edited by Russ Shafer-Landau, 216–36. Oxford: Oxford University Press, 2019.
- Kavka, Gregory S. “The Toxin Puzzle.” *Analysis* 43, no. 1 (January 1983): 33–36.
- Kiesewetter, Benjamin. *The Normativity of Rationality*. Oxford: Oxford University Press, 2017.
- Kolodny, Niko. “The Myth of Practical Consistency.” *European Journal of Philosophy* 16, no. 3 (December 2008): 366–402.
- . “Why Be Rational?” *Mind* 114, no. 455 (July 2005): 509–63.
- Lord, Errol. *The Importance of Being Rational*. Oxford: Oxford University Press, 2018.
- Maguire, Barry. “There Are No Reasons for Affective Attitudes.” *Mind* 127, no. 507 (July 2018): 779–805.
- . “The Value-Based Theory of Reasons.” *Ergo* 3, no. 9 (2016): 233–62.
- Marušič, Berislav. *Evidence and Agency: Norms of Belief for Promising and Resolving*. Oxford: Oxford University Press, 2015.
- McCann, Hugh J. “Settled Objectives and Rational Constraints.” *American Philosophical Quarterly* 28 no. 1 (January 1991): 56–36.
- McDowell, John. “Some Remarks on Intention in Action.” *Amherst Lecture in Philosophy* 6 (2011): 1–18. <http://www.amherstlecture.org/mcdowell2011/>.
- McHugh, Conor. “Attitudinal Control.” *Synthese* 194, no. 8 (August 2017): 2745–62.
- . “Judging as a Non-Voluntary Action.” *Philosophical Studies* 152, no. 2 (January 2011) 245–69.
- McHugh, Conor, and Jonathan Way. “Fittingness First.” *Ethics* 126, no. 3 (April 2016): 575–606.
- O’Shaughnessy, Brian. “Trying (as the ‘Mental Pineal Gland’).” *Journal of Philosophy* 70, no. 13 (July 1973): 365–86.
- Owens, David. “Habitual Agency.” *Philosophical Explorations* 20 (2017): 93–108.
- Paakkunainen, Hille. “Can There Be Government House Reasons for Action?” *Journal of Ethics and Social Philosophy* 12, no. 1 (September 2017): 56–93.
- Papineau, David. “Normativity and Judgement.” *Aristotelian Society Supplementary Volume* 73, no. 1 (July 1999): 17–43.
- Parfit, Derek. *On What Matters*. Oxford: Oxford University Press, 2011.

- Paul, Sarah K. "The Conclusion of Practical Reasoning." *Canadian Journal of Philosophy* 43, no. 3 (2013): 287–302.
- Pink, Thomas. *The Psychology of Freedom*. Cambridge: Cambridge University Press, 1996.
- Portmore, Douglas W. *Commonsense Consequentialism*. Oxford: Oxford University Press, 2011.
- . *Opting for the Best: Oughts and Options*. Oxford: Oxford University Press, 2019.
- Raz, Joseph. *From Normativity to Responsibility*. Oxford: Oxford University Press, 2011.
- Ryan, Sharon. "Doxastic Compatibilism and the Ethics of Belief." *Philosophical Studies* 114, nos. 1–2 (May 2003): 47–79.
- Scanlon, T. M. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1998.
- Schroeder, Mark. *Slaves of the Passions*. Oxford: Oxford University Press, 2007.
- . "The Ubiquity of State-Given Reasons." *Ethics* 122, no. 3 (April 2012): 457–88.
- Setiya, Kieran. "Intention." *Stanford Encyclopedia of Philosophy* (Fall 2018). <https://plato.stanford.edu/archives/fall2018/entries/intention/>.
- . *Reasons without Rationalism*. Princeton: Princeton University Press, 2007.
- Shah, Nishi. "How Action Governs Intention." *Philosophers' Imprint* 8, no. 5 (July 2008): 1–19.
- . "A New Argument for Evidentialism." *Philosophical Quarterly* 56, no. 225 (October 2006): 481–98.
- Shah, Nishi, and Matthew Silverstein. "Reasoning in Stages." *Ethics* 124, no. 1 (October 2013): 101–13.
- Silverstein, Matthew. "Reducing Reasons." *Journal of Ethics and Social Philosophy* 10, no. 1 (February 2016): 1–22.
- Smith, Michael. "The Ideal of Orthonomous Agency, or the How and Why of Buck-Passing." In *Thinking about Reasons: Themes From the Philosophy of Jonathan Dancy*, edited by David Bakhurst, Margaret Olivia Little, and Brad Hooker, 50–75. Oxford: Oxford University Press, 2013.
- Steglich-Peterson, Asbjørn. "How to Be a Teleologist about Epistemic Reasons." In *Reasons for Belief*, edited by Asbjørn Steglich-Peterson and Andrew Reisner, 13–33. Cambridge: Cambridge University Press, 2011.
- Steup, Matthias. "Doxastic Freedom." *Synthese* 161, no. 3 (April 2008): 375–92.
- Tenenbaum, Sergio. "The Conclusion of Practical Reason." In *Moral Psychology*, edited by Sergio Tenenbaum, 323–43. Amsterdam: Rodopi, 2007.

- Thompson, Michael. *Life and Action*. Cambridge, MA: Harvard University Press, 2008.
- Thomson, Judith Jarvis. *Normativity*. Chicago: Open Court, 2008.
- Way, Jonathan, and Daniel Whiting. "Reasons and Guidance (or, Surprise Parties and Ice Cream)." *Analytic Philosophy* 57, no. 3 (September 2016): 214–35.
- Weatherson, Brian. "Deontology and Descartes's Demon." *Journal of Philosophy* 105, no. 9 (September 2008): 540–69.
- Wedgwood, Ralph. *The Value of Rationality*. Oxford: Oxford University Press, 2017.
- Williams, Bernard. "Internal and External Reasons." In *Moral Luck*, 101–13. Cambridge: Cambridge University Press, 1981.
- Wilson, George, and Samuel Shpall. "Action." *Stanford Encyclopedia of Philosophy* (Winter 2016). <https://plato.stanford.edu/archives/win2016/entries/action/>.

PRACTICAL COMMITMENT IN NORMATIVE DISCOURSE

Pekka Väyrynen

EVALUATIVE AND DEONTIC JUDGMENTS play a practical role in our thought. Judging that something is good, or right, or what we ought to do, ranks things in ways that guide us toward the higher-ranked things. When I judge that I ought to eat more iron, normally I reveal (or make) myself to be motivated or at least committed to eating more iron, feeling bad if I do not, and so on. When we judge that we have most reason to slow down climate change, this normally “stops the buck” in deliberation and decision about what to do. It is controversial whether such a role in practical reasoning is essential to the evaluative or deontic concepts deployed in these judgments. It is controversial to what extent “thick” concepts like TACTFUL, KIND, and CHASTE play such a practical role in our thought. And it is controversial just what the practical role of a given normative concept is. But it is widely agreed that, at minimum, descriptively “thin” normative concepts play some such practical role for rational agents with a normal grasp of those concepts.¹ Thin normative concepts are, roughly, normative concepts that encode little descriptive information. They are characteristically expressed by terms like *good*, *right*, and *ought* in English and their equivalents in other languages. For convenience, call these “thin normative terms.”²

This paper concerns how the practical role of normative thought is reflected

- 1 I will use small capitals to denote concepts and italicized words to denote lexical items. The use of italics for emphasis instead will be clear from the context. I use “normative” to cover both the evaluative and the deontic. What I call a practical role is called “practical role” by Ridge (*Impassioned Belief*, 21), “normative role” by Eklund (*Choosing Normative Concepts*, 9), and “guiding or regulative role” by Wedgwood (*The Value of Rationality*, 15). The term “normative role” may be misleading insofar as it may refer merely to how a normative concept is associated with other normative concepts—not my topic here. Finally, to keep things manageable, I will bracket epistemic judgments, such as those concerning what one ought to believe.
- 2 These terms have other uses as well. I assume we have a reasonably determinate grasp of which uses of these terms express thin normative concepts. What I say will not hang on having a precise delineation of that class.

in *language*. Normative language manifests a parallel phenomenon. I reveal (or make) myself to be motivated or committed to eating more iron no less when I assert “I ought to eat more iron” than when I judge that I ought to eat more iron. It is natural to wonder whether the phenomenon is semantic or pragmatic, especially since it may not be cleanly delineable in natural language. If the phenomenon were specific to bare first-person assertions like “I ought to eat more iron,” there would be little initial pressure to think it semantic. Even if normative assertions in the first person are important to understanding normative thought or its relationship to action, it does not follow that they are linguistically special. And yet the standard assumption in metaethics is that at least thin normative terms are associated with their practical role somehow as a matter of meaning.³ This assumption is usually implicit and rarely articulated precisely. But it faces a serious problem in any form. As we will see, terms like *ought* and *good* often figure in assertions where their customary practical role is absent. Such cases pose a challenge: either offer some plausible explanation of why the relevant practical upshots do not show up in these cases despite featuring in our overall semantic theory for these expressions, or else do not build those practical upshots into such a theory.

I will argue that existing accounts of normative language have no adequate explanation of how the association between thin normative terms and practical role can be broadly semantic and yet tolerate the exceptions I will highlight. In closing I will consider the prospects of a pragmatic account of the phenomenon. This is an option to take seriously because generalizations that do not rise to the level of full generality are normally more apt to be explained in pragmatics.

My enterprise is hermeneutic: it concerns how the association between thin normative terms and their practical roles is in fact reflected in natural language. This gives the issues at stake and my arguments significant broader interest and importance. If the association between thin normative terms and practical role is not broadly semantic, this puts significant pressure on standard forms of metaethical expressivism and inferentialism. For expressivists typically account for the meaning of normative language in terms of practical states of mind, such as plans, desires, or sentiments. And inferentialists typically account for the meaning of normative language in terms of practical commitments to not only think but also act and feel in certain ways. Existing accounts of neither sort have an adequate solution to the challenge I will raise. If the association between thin normative terms and practical role is not broadly semantic, this will also suggest certain general lessons about how normative thought and language relate. For instance, our normative discourse can be in perfectly good working order with-

3 Exceptions include Strandberg, “A Dual Aspect Account of Moral Language”; and Finlay, *Confusion of Tongues*; see section 5.1. My argument will be more general than theirs.

out containing predicates that are associated with certain practical roles in any broadly semantic way. Even if there are distinctively practical OUGHT concepts that are conceptually or constitutively linked to certain practical upshots, we have reason to doubt that this link will be encoded in the correct overall semantic theory for natural languages. Such concepts may be privileged in normative theory, but not in our overall semantic theory for normative language. We might then also wonder how crucial such concepts really are for normative thought. Philosophers of normativity will therefore need to think more carefully about the relation between normative thought and normative talk.

1. SEMANTIC VIEWS OF THE PRACTICAL ROLE OF NORMATIVE DISCOURSE

One of the few widely shared convictions among moral philosophers is that normative judgments play a distinctive role in our practical thinking. We may disagree about just what the practical role of a given normative concept is and be unable to state it in uncontroversial terms. But we have some grasp of the notion, since we routinely compare the practical roles of different normative concepts.⁴ For instance, moral evaluation is different from aesthetic evaluation, and the practical role of CHASTE is different from that of RIGHT. Some normative concepts may even be individuated, at least in part, by their practical role. I will not deny that people can have such concepts.⁵ But do words not express concepts? So if there were distinctively practical concepts, would it not then be part of the meaning of the words that express these concepts to have this connection to practical role? Sincere moral claims by rational agents, for instance, are often thought not to tolerate the absence of corresponding practical upshots. This line of thought might also help distinguish normative terms (and not just normative concepts) from nonnormative terms. What could the difference be, if not that something about the meaning of normative terms connects them to a practical role?

This is all too quick. If words express concepts, it does not follow that those concepts directly determine the meaning of those words. In the normative case, the following is possible, for instance: (i) if a concept plays a role *R* for a thinker, then it picks out moral rightness but (ii) the word *right* is not constrained by the rules of the language to stand *only* for such a concept. In other contexts, it might be used to express concepts that lack a practical role, such as what is right according to my parish priest. So even if some normative concepts are essentially

4 Eklund, *Choosing Normative Concepts*, 45–46.

5 Whether analyzing the nature of normative concepts in terms of their practical point would yield any substantive results is a further question. For a cautionary discussion, see Sharadin and Van Someren Greve, “Is Deontic Evaluation Capable of Doing What It Is For?”

associated with certain practical roles, it does not follow that this connection is reflected in the meanings of the words used to express those concepts. In general, linguistic meaning and mental content might not line up in any neat way.⁶ It remains similarly open whether what makes normative discourse different from nonnormative discourse is how their meanings relate to some distinctive kind of practical import.

What might it be for thin normative terms to be associated with their practical roles as a matter of meaning? Our overall theory of meaningful linguistic communication is normally taken to have three main parts (whose boundaries may not be sharp): descriptive semantics, metasemantics, and pragmatics. When I ask whether the practical role of thin normative terms is a “broadly semantic” phenomenon that holds “as a matter of meaning” and is to figure in our “overall semantic theory,” the question is whether the phenomenon is a descriptive-semantic or metasemantic rather than pragmatic phenomenon. A pragmatic account would take an account of the meaning of a normative sentence and say that the relevant practical upshots are some further kind of information that may be conveyed by its utterances, given its meaning and the context: conversational implicature, pragmatic presupposition, or the like.

On descriptive-semantic accounts, the association between thin normative terms and practical role would be a feature of assignments of semantic value to thin normative terms. “Semantic value” is a neutral term for the things that, according to a semantic theory, “provide the interpretations of simple expressions and are the arguments and values of the functions defined by the compositional rules that interpret the complex expressions.”⁷ I will take descriptive semantics to include an account of the lexical properties of simple expressions and other conventional aspects of their meaning in a language, since these will constrain assignments of semantic value.⁸ (For context-sensitive expressions, context provides further constraints.) So, for instance, if the practical role of *ought* were a lexical property of the word, it would be a descriptive-semantic phenomenon. Assignments of semantic value to expressions by a theory are supposed to ac-

6 Cf. Yalcin, “Semantics and Metasemantics in the Context of Generative Grammar”; Glanzberg, “Lexical Meaning, Concepts, and the Metasemantics of Predicates.”

7 Stalnaker, “Reference and Necessity,” 535.

8 Conventional aspects of meaning are (very roughly) relatively arbitrary regularities of a population using a certain linguistic form (e.g., a word or sentence) to mean a certain thing. (Conventions are arbitrary in that no general account than a historical one is available for them.) For a discussion of various complications, including that conventionality may have multiple dimensions and come in degrees, see Simons and Zollman, “Natural Conventions and Indirect Speech Acts.”

count for features like the entailments licensed by the expressions and the truth conditions of the sentences they compose.⁹

On metasemantic accounts, the association between thin normative terms and practical role would instead be a feature of that in virtue of which thin normative terms have the semantic values they do. Metasemantics (sometimes called “foundational semantics”) is about “what the facts are that give expressions their semantic values, or more generally, about what makes it the case that the language spoken by a particular individual or community has a particular descriptive semantics.”¹⁰ It involves explaining both why expressions have the context-invariant, descriptive-semantic features they do and what it is about the situation, behavior, or mental states of the speaker or their audience that makes it the case that a context-sensitive expression has the contextually determined semantic value it does. For instance, if the semantic value of *ought* were determined by its inferential role and its inferential role were essentially practical, then the practical role of *ought* would be a metasemantic phenomenon (see section 4.4).

Canonical statements of semantic or metasemantic accounts of the practical role of normative discourse are hard to find. But there are suggestions in the ballpark. Hare thinks that the “primary meaning” of *ought* is “prescriptive”—one that, by Hare’s stipulation, entails an imperative.¹¹ Similarly, the primary meaning of *good* is an evaluative meaning that entails a recommendation. This, he thinks, is what anchors meaningful connections between the applications of these terms to very different sorts of things by different people.¹² Dreier’s “speaker relativism” analyzes “*x* is good” as meaning “*x* is approved of by *M*,” where *M* is a contextually specified moral system. He takes it to be built “into the very semantics of moral terms” that *M* is picked out in part by certain kinds of motivational and affective factors.¹³ Blackburn writes that “since moralizing and valuing are distinctive activities, the words we use to communicate our morals and our values will have their distinctive meanings” and that capturing the way that ethics is “essentially practical” requires “showing how [ethical] judgment

9 The dominant approach to semantics in this sense is truth-conditional semantics: the semantic value of a sentence just is its truth conditions, and the semantic value of an expression is the contribution it makes to the truth conditions of the sentences in which it features. But of course there are alternatives. Expressions may also have various non-compositional conventional features. Here I will have to bracket views on which moral claims play a practical role via conventional implicature (Copp, “Realist-Expressivism”) or semantic presupposition.

10 Stalnaker, “Reference and Necessity,” 535.

11 Hare, *The Language of Morals*, 159, 164.

12 Consider the example of the missionary and the cannibals in Hare, *The Language of Morals*, 148–49.

13 Dreier, “Internalism and Speaker Relativism,” 6, 23–24.

has a content or truth-condition that is itself magnetic.”¹⁴ Wedgwood proposes that the meaning of *good* and *ought* each is given by the basic rules of rationality governing its use, where these rules concern commitments to certain preferences or planning states.¹⁵ According to Gibbard, the meaning of *ought* is explained by its tie to planning states; “plan-laden concepts” have “much to do” with what ordinary normative terms like *ought* express.¹⁶ These proposals all affirm a connection between what thin normative terms mean and some kind of practical role. As we will see, they share a significant assumption. But they resist summary into a single hypothesis.

One respect in which these proposals differ concerns what kind of thing the practical role of normative language is. Some focus on practical *dispositions*: at least all else equal, judging that one ought to ϕ in circumstances *C* (etc.) tends to motivate one to ϕ in *C*, or to close deliberation about whether to ϕ , or the like. Perhaps one will also tend to feel guilty if one does not ϕ or resent or blame others if they do not ϕ in relevantly similar circumstances. Others focus on practical *commitments*: judging that one ought to ϕ in *C* (etc.) commits one to thinking, feeling, or acting in certain ways, such as perhaps forming the intention to ϕ in *C*, or making ϕ -ing in *C* part of one’s ideal plan about what to do. Perhaps it also commits one to disapproving, resenting, or blaming those who do not ϕ in *C*, and so on.¹⁷ Of course, it might be that the practical role of some normative terms involves motivations while that of some others involves commitments, and the practical role of some terms might involve both. In what follows I will mainly talk of practical commitments, since this is a logically weaker notion. Affirming a practical commitment requires less of an agent than having a matching disposition to comply.¹⁸

Another respect in which the proposals noted above differ concerns what

14 Blackburn, *Ruling Passions*, 87, 115.

15 Wedgwood, “Conceptual Role Semantics for Moral Terms” and *The Nature of Normativity*, ch. 4.

16 Gibbard, *Thinking How to Live*, 139.

17 Practical commitments differ from various species of theoretical commitment. They are not merely *intranormative*. Judging that I (morally, legally, etc.) ought to ϕ in *C* might commit me to thinking there is (moral, legal, etc.) reason for me to ϕ in *C*. Whether this commits me to acting, feeling, or thinking on the basis of those reasons is a further question. Nor can these commitments be merely *alethic* or *doxastic*. If I judge that a measure of voting intention is unreliable, I commit myself to the truth of that proposition, but not any intention or sentiment. The same goes for the commitment to believe *q* generated by judging that *p* and believing that $p \rightarrow q$. Practical commitments are more like the following: if I promise that I will be a better parent, I commit to being a better parent in the future, and the same goes if I endorse the thought that I shall be a better parent.

18 Either kind of practical role is richer if some form of internalism about it is true, such as

kind of connection is supposed to obtain between thin normative terms and their practical role. Hare and Blackburn seem to make claims in descriptive semantics. Dreier, Gibbard, and Wedgwood seem to have in mind metasemantic views. But they all seem to imply that thin normative terms' association with their practical roles is some indispensable feature of their meaning that any competent user of these expressions must master.¹⁹ This is a useful starting point; it will help articulate a problem that any broadly semantic view must address somehow or other.

2. THE PROBLEM OF DIVERSE USES

I will now argue that thin normative terms are often used sincerely and felicitously to state requirements or recommendations in ways that do not commit the speaker to thinking, feeling, or acting in accordance with those verdicts.²⁰ If the practical role of these terms were a part of their conventional profile in a language, it should not be subject to such exceptions but instead should be present in all literal uses in normal contexts.²¹ This raises what I will call the Problem of Diverse Uses: How do you reconcile the diversity of uses to which thin normative terms may be put with the claim that their association with their normative roles is broadly semantic? The problem prompts a challenge: either offer some plausible explanation of cases where the relevant practical upshots are absent that reconciles these claims, or else do not build such upshots into our overall semantic theory for thin normative terms. I will be driving us to the latter option.

motivational internalism about normative judgment. Each is distinct from Klimczyk's idea of "authored practicality" ("Compositional Semantics and Normative 'Ought'").

- 19 In due course I will ask whether all broadly semantic views carry this implication in full. Some recent metasemantics for normative discourse might not. These include Ridge, *Impassioned Belief*; Chrisman, *The Meaning of "Ought"* and "Two Nondescriptivist Views of Normative and Evaluative Statements"; Köhler, "Expressivism, Meaning, and All *That*"; Bedke, "Practical Oomph"; and Tiefensee, "Inferentialist Metaethics, Bifurcations, and Ontological Commitment" and "Metasemantics for the Relaxed." I will discuss Ridge, Chrisman, and Tiefensee in sections 4.3 and 4.4.
- 20 As the examples below make clear, I am not talking about sentences like "Martha says I ought to tell the truth," "Either I ought to tell the truth or it is not the case that I ought to tell the truth," or the embedding of thin normative terms under negation or possibility modals like *might*, or in the antecedent of a conditional.
- 21 This holds irrespective of whether practical upshots are features of propositions that speakers communicate when making claims using normative terms, or of acts of making such claims, such as their propriety conditions. An example of the latter kind of view would be the ethical neo-expressivism of Bar-On and Chrisman, "Ethical Neo-Expressivism."

One sort of evidence for the relevant kind of diversity of uses is that any suggestion of practical commitment is absent in cases like 1–2:

1. One ought to prioritize profit over fairness. But is that really the thing to do?
2. Client: What is my legal obligation, and what do you expect me to do?
Lawyer: You have to report your liability, but I do not know if you will; you may prefer to push the limits of the law and just conceal it.²²

Claims like 1 can make perfectly good sense in conversations about capitalist worldviews. The corresponding interpretation of *ought* need only be implicit in the context for the overt utterance to sound coherent. Variants of 1–2 arise by varying the modal verbs: *ought* for *have to*, and so on.²³ The point of these examples is that practical commitments do not result conventionally from using *wrong* or *ought* to state requirements or recommendations. Conventional features of utterances are not as easily defeasible as the absence of practical upshots in 1–2 would require.

One might claim that 1–2 are insufficient to show that the association between thin terms and practical role is not semantic. Perhaps not all sincere assertions of thin normative claims in normal contexts must carry practical upshots for such association to be a broadly semantic matter. For instance, the phenomenon seems clearest in bare first-person uses, but 1–2 are not such cases. However, there is no general reason why a difference in personal pronoun should make this particular kind of semantic difference. Moreover, practical upshots are absent also in some bare first-person assertions. To many people's ears, sincere assertions of 3–4 are indicative of a substantive normative mistake rather than linguistic incompetence:

3. I would be wrong to kill. But I am OK with killing and do not feel bad about it.
4. I ought to finish grading. I have absolutely no intention to do so, though.²⁴

If a sincere assertion of “I ought to finish grading” always committed you to plan-

22 Mandelkern, “Practical Moore Sentences,” 43.

23 See, e.g., Silk, *Discourse Contextualism*, 40: “Ernie ought to be home by 10. Aren’t his parents stupid? I’d stay out if I were him.”

24 Woods discusses these kinds of examples in the context of expressivism (“Expressivism and Moore’s Paradox”). Ruiz and Stojanovic suggest that such examples need a contrasting connective like *but* and this is evidence that if a rational and sincere speaker says something positive about an object, she normally expresses a certain practical stance (“On Linguistic Evidence for Expressivism,” 159–63). I can agree, but would deny (in part for reasons I will

ning to do the thing, then asserting it while also rejecting such plans, as in 4, should induce a strong feeling of incoherence. But it need not. Even if *wrong* and *ought* in many contexts express concepts whose application is practically committal in this way, instances of 3–4 need not express such concepts. The first half of 4 would often express something more like a role obligation.

So 3–4 are evidence that a practical commitment does not result conventionally even from first-person assertive utterances of lexical items like *wrong* and *ought*. Similar evidence comes from psychologically realistic characters who take delight in evil for evil's sake, are committed to sheer malignity under that description, and so on. When Iago sincerely identifies his conduct as demonic, he is specifying a feature of his conduct that he is committed to pursuing and strongly inclines him to engage in the conduct.²⁵ Further first-person cases where practical commitment is absent include:

5. I should do the shopping today (as far as I know).²⁶
6. CEO: What is our legal obligation, and what we should we do?
Head of Legal: We have to report our liability, but could decide to push the limits of the law and just conceal it.

In 5 the speaker states a normative requirement that follows (as far as she knows) from her household's shopping arrangements. Her commitments to responding accordingly are simply a further issue. From 6 we see that not all first-person uses are singular. A special convention for bare singular first-person present-tense assertions is all the less likely, and would not in any case be a conventional aspect of the meaning of thin normative terms themselves.

Examples 1–6 illustrate that thin normative terms are often used noncommittally, to state normative verdicts in ways that do not carry corresponding practical commitments.²⁷ My sample theorists acknowledge some such cases. Hare says

mention in section 5.1) that this tells us as much about the meaning of normative terms as they take it to do.

25 Cf. Kramer, "Shakespeare, Moral Judgments, and Moral Realism."

26 I adapt 5 from a second-person example in Kaufmann, *Interpreting Imperatives*, 58.

27 Some readers might be wondering about uses of normative terms to express imperatives. On the relationship between strong necessity modals like *must* and *have to* and imperatives, see, e.g., Ninan, "Two Puzzles about Deontic Necessity"; Kaufmann, *Interpreting Imperatives*, sec. 2.3.1; and Mandelkern, "Practical Moore Sentences." Some of my examples suggest that these modals do not always (even in their deontic uses) express orders or commands. Although my focus is not on orders or commands, it is worth noting that one type of semantics of imperatives analyzes "Close the door!" as meaning "You must/should close the door" and locates the pressure to act that imperatives often induce on their addressees in certain pragmatic features that govern their felicitous use, not in an explanation of their meaning. (See, e.g., Kaufmann, *Interpreting Imperatives*.) However, the debate about imperatives is ongoing.

that *ought* and *good* can have “secondary” descriptive meanings. Dreier, Blackburn, and Gibbard allow that even sincere first-person *ought* statements do not always imply that the speaker is in a corresponding conative state.²⁸ But none of them is very clear about just how thin normative terms’ association with their practical role can allow exceptions and yet be a broadly semantic phenomenon.²⁹ (I will extract some suggestions shortly.) Examples like 1–6 imply that such an association is not the sort of conventional matter that these views take it to be.³⁰ Similarly, insofar as assertive uses of thin normative terms carry directive or commissive illocutionary force, such force will not derive from their lexical meaning.

3. DIVERSE USES AND DESCRIPTIVE SEMANTICS

Any account of the diverse uses of thin normative terms that we have observed will take one of two forms. Practically committal and noncommittal uses of thin normative terms either have a uniform descriptive semantics, or they do not. These are the options whether or not their practical role is a broadly semantic phenomenon. It will be instructive to begin by considering why some simple broadly semantic views of each type fail to solve the Problem of Diverse Uses.

One proposal that posits semantic uniformity says that deontic uses of terms like *ought* affirm the practical commitments of some agent or other. The thought might be that committal uses convey the speaker’s own commitments, and noncommittal uses play the relevant practical role indirectly by committing the

28 In Wedgwood, *The Nature of Normativity*, the acknowledgment is implicit. His account is officially limited to a certain “deliberative” concept that *ought* can be used to express (97). He does not say how to generalize his analysis so that it would account for the meaning of *ought*, given the many other deontic concepts that the word can be used to express.

29 Hare, Blackburn, and Gibbard describe uses of *ought* for which they take motivational internalism to be a conceptual truth. But *ought* can be used in deontic contexts to express normative concepts for which motivational internalism does not hold. Dreier analyzes “*x* is good” as meaning “*x* is approved of by *M*,” where *M* is a contextually specified moral system. According to him, it is built into the meaning of *good* that *M* is chosen from the context by balancing the content and subject matter of a set of rules and those rules’ motivational and affective connections with the speaker (Dreier, “Internalism and Speaker Relativism,” 24–25). But the meanings of these terms do not require us to interpret 1–6 by considering what rules have such connections. (Also, *good* has other than moral uses.)

30 My arguments allow that to understand the lexical meaning of a word, we may have to look at a group of semantically related words, such as *may*, *must*, and *have to* in the case of *ought*. How concepts relate in general to lexical meaning is a more complex issue to which I cannot do justice; a rich discussion of this is Glanzberg, “Lexical Meaning, Concepts, and the Metasemantics of Predicates.”

speaker to thinking or feeling in a certain way in such and such a contingency.³¹ But this will not do. To interpret the deontic content of “We ought to report our liability, but could decide to push the limits of the law and just conceal it” in 6, we need not suppose that the Head of Legal is committed to reporting the liability in the case where the firm decides not to push the limits of the law. Or consider a value system that a science fiction author just invented. There is no pressure to interpret the *ought* claims that describe that value system’s verdicts as expressing the practical commitments of the members of this culture. In such cases, the relevant agent could only be specified as “someone who endorses the norms at issue.” In general, we understand thin normative claims so long as we can identify the relevant standards: a capitalist value system in 1, the local legal code in 2 and 6, and so on. Stating the recommendations of a standard does not require assumptions regarding who if anyone is committed to those standards. And while such recommendations can always be represented by a preference ordering, this does not require preferences to encode motivations or commitments. They need only function as abstract ranking devices.

A correspondingly simple proposal without semantic uniformity is that although noncommittal uses of thin normative terms do not directly convey practical commitments, these uses are ultimately to be explained by reference to committal uses. A proposal that is familiar from other contexts is to analyze noncommittal uses as parasitic on committal uses. They might be claimed to report or otherwise allude to other people’s committal normative judgments.³² However, this solution undergeneralizes. Normative language can be used to make noncommitted statements about what follows from some normative system that have never occurred to anyone before.³³

A different strategy that implies lack of semantic uniformity is to argue that cases where the relevant practical upshots are present and absent are distinguished by linguistic convention after all. One view is that thin normative terms have multiple distinct but related meanings. It is controversial whether modal expressions like *ought* are polysemous across different flavors of modality, such as epistemic, deontic, and dynamic.³⁴ But examples 1–6 all concern *ought* in de-

31 Cf. Gibbard, *Thinking How to Live*; Ridge, *Impassioned Belief*, 19.

32 Cf. Hare, *The Language of Morals*, 124–26, 164–65; Blackburn, *Ruling Passions*, 59–68, 110.

33 Raz, *The Concept of a Legal System*, 235.

34 The dominant tradition in the semantics of modals, developing the framework in Kratzer (“Modality”), aims to give a unified semantics of deontic, epistemic, and other flavors of modality. Viebahn and Vetter argue that each modal expression has distinct but systematically related meanings for different modal flavors, while their flexibility within each modal flavor is due to context sensitivity (“How Many Meanings for ‘May’?”). This is compatible with my arguments. Even if the lexical entry for *ought* includes a variable whose different

ontic contexts. It is much less common to think that the flexibility of modal expressions within a modal flavor is due to polysemy, rather than context sensitivity (on which more in section 4).³⁵ Features of a mere subset of uses of deontic expressions are not evidence about their meaning.

A more sophisticated version of this strategy is to claim that deontic claims that do not play a practical role are not genuinely normative claims, and so deontic terms are not used normatively in 1–6. What “normative” means in the talk of normative terms and claims is an infamously fraught issue.³⁶ The notion of the “genuinely” normative also requires explication. But interpreting “normative” to include only the genuinely normative is unduly narrow for the present purposes in any case. Suppose the sense in which a sentence like “One ought to do that” is normative only applies to uses that essentially involve practical commitments. This class of “normative” sentences would exclude many uses of *ought* that also are distinctive from typical uses of paradigmatically descriptive, nonnormative words like *tree* or *tall*.³⁷ The deontic claims in 1–6 are normative in the broader sense that they concern standards that are such that falling short of them opens one up to certain kinds of criticism—legal standards in the case of 2 and 6, and so on. Even if both “I ought to finish grading” (as asserted in the context of 4) and “Matt is tall” are descriptive in *one* sense, the implications of falling short of

values give *ought* an epistemic, deontic, or goal-oriented flavor, the practical role of deontic *ought* does not have a comparable status. See section 4.2.

- 35 See also Kaufmann, who uses polysemy avoidance to make a case for a uniform semantics for descriptive and performative modal verbs (*Interpreting Imperatives*, 60–62). There is a debate in metaethics about whether *ought* is polysemous between a “practical”/“deliberative” sense and an “evaluative” sense (Schroeder, “Ought, Agents, and Actions”; Finlay and Snedegar, “One Ought Too Many”; Chrisman, *The Meaning of “Ought,”* 124–33). But the distinctions cited in that debate differ from those on the table here. Perhaps the deliberative *ought* is always indexed to a particular agent and time (Wedgwood, *The Nature of Normativity*, 90). But *ought* can be so indexed when it is used to state what follows from some assumed standard for some other agent’s situation at a given time. Or perhaps the deliberative *ought* “matters directly for advice” and “is the right kind of thing to close deliberation” (Schroeder, “Ought, Agents, and Actions,” 17). But an *ought* claim describing what follows from a set of standards can constitute advice that is suitable for closing deliberation without the relevant practical commitments being in place; recall 2 and 6. Cf. Bronfman and Dowell, “The Language of ‘Ought,’ and Reasons,” 107–8.
- 36 Finlay, “Defining Normativity.”
- 37 The more narrowly we understand “normative,” the less clear it is that there strictly speaking are normative words or sentences in English. There may just be words and sentences that can be used normatively (in the various senses of “normative”) or nonnormatively without changing meaning. Alwood (“Should Expressivism Be a Theory at the Level of Metasemantics?” 15) makes this point about the notion of “normative” employed in Ridge, *Impassioned Belief*. I will discuss Ridge’s view in section 4.3.

the standards invoked by 4 are different from those of falling short of the contextually supplied standards of tallness. The lack of practical upshot may mean that 1–6 express requirements whose normativity is merely “formal” rather than “substantive” or “authoritative.” But these cases show that one function of these words in English is to express formally normative claims. So the sense of “normative” that is relevant to our semantic theory for expressions like *good* and *ought* had better include also such formal normativity.

We could instead seek progress by characterizing the different uses to which normative terms may be put. We distinguish (as I did earlier) “engaged” or “committal” and “detached” or “noncommittal” uses of normative language.³⁸ We might gloss this further by saying that normative terms have noncommittal uses that state requirements or recommendations that follow from some assumed body of standards, and committal uses that somehow (semantically or pragmatically) also convey corresponding practical commitments. One proposed test is that a use is at least normally noncommittal if the relevant utterance is reasonably interpreted as if it contained an implicit *according to*-type phrase.³⁹ This is a fallible test. When the doorman says to me, “According to the rules of our club, patrons may not wear sneakers,” this is not a merely descriptive use. The doorman is committed to not letting me in. But we recognize the distinction, however particular cases fall with respect to it. (In this example, the practical upshot of the doorman’s utterance is clearly a kind of pragmatic meaning.)

It is one thing to grant this distinction between committal and noncommittal uses of normative terms, quite another to claim that it is somehow marked by linguistic convention or otherwise drawn in our semantic theory for normative terms. I see little reason to suppose that the distinction between committal and noncommittal uses of terms like *ought* and *good* has such a status. The distinction does not involve a difference in linguistic form, whereas linguistic conventions pair forms with meanings. So even if there are complex regularities regarding when an assertion of “I ought to finish grading” is committal and when it is noncommittal, it is not at all clear how they could be candidates for linguistic convention.⁴⁰ Whether an assertive use of “I ought to finish grading” is committal or noncommittal is a function of features of the context of utterance that do

38 Raz, *The Concept of a Legal System*, 234–35.

39 Cf. Silk, *Discourse Contextualism*, 130–32; and Chrisman, “Two Nondescriptivist Views of Normative and Evaluative Statements,” 412. Ruiz and Stojanovic propose various other criteria for distinguishing “expressive” and “factual” uses of terms (“On Linguistic Evidence for Expressivism”). Not all of these, however, seem likely to carry over.

40 For a useful exercise, compare how different this case is from whether indirect speech acts (such as using “Can you pass the salt?” to issue a request) are conventional. See especially Simons and Zollman, “Natural Conventions and Indirect Speech Acts,” 13–22.

not normally ground linguistic convention. The distinction looks more apt for speakers to track by means of their world knowledge and pragmatic reasoning. Nor are there clear analogues in other segments of natural language. For instance, to be a competent user of “tummy” in English, I need to grasp not only that it denotes stomachs but also that it belongs to infant-directed speech; “tummy” is unsuitable by virtue of its meaning for me to use at a gastroenterologist’s.⁴¹ The difference between committal and noncommittal uses of normative language is not this kind of conventionalized difference in register. Although I will be a fair target of criticism if I assert “I ought to do that” in a seemingly committal way without being practically committed in the relevant ways, my *faux pas* (or, rather, lack of any *pas*) is just a special case of violating a general sincerity condition on speech acts, nothing specific to normative language.

It does not help the case to claim that committal uses express concepts that are crucial to a certain kind of practical thought. For instance, if noncommittal uses of normative sentences just update our picture of what is to be done relative to some standard, their acceptance need not involve more than recognizing those verdicts. Even if the acceptance of a normative sentence involves more in other cases, features specific to those uses are not evidence about the meanings of normative terms.⁴² There is also no systematic link between whether *ought* is used committally or noncommittally and what kind of normativity it is used to express. People can get very committed about what they should do by the formally normative standards of feline excitement or espresso excellence. In sum: to solve the Problem of Diverse Uses, it is not enough just to distinguish committal and noncommittal uses of normative terms.

What I take my discussion so far most strongly to suggest is that thin normative terms’ association with their practical role is unlikely to be a phenomenon at the level of descriptive semantics. We might have expected this from the start. That the speaker is practically committed to eating more iron seems not to be part of what “I ought to eat more iron” says. In that case we should expect committal and noncommittal uses of thin normative terms to be uniform in their descriptive semantics. This is dialectically significant. For if the relevant practical upshots are absent in many contexts, then semantic uniformity would be easy to explain if the presence of practical upshots in other contexts were not a broadly semantic phenomenon. The pressure is therefore on for the alternative broadly semantic view that thin normative terms’ association with their practical role is a metasemantic phenomenon instead.

41 Ferguson, “Baby Talk in Six Languages.”

42 Contrary to Gibbard, *Thinking How to Live*, 153.

4. PRACTICAL ROLE AND THE METASEMANTICS OF *OUGHT*

To assess the metasemantic view, we require some idea of what kind of descriptive semantics thin normative terms have. I will use as my sample the standard sort of descriptive semantics for deontic modal expressions in English. I will argue that nothing in this formal semantics or plausible supplementary assumptions supports treating those uses of deontic terms that are associated with a practical role as semantically or metasemantically exceptional.

4.1. *The Standard Semantics for Deontic Modals*

The dominant view in linguistics and philosophy of language is that expressions such as *ought/must/should/have to* are context-sensitive modal operators. For instance, in *Ought(p)*, *ought* is a nonpropositional device for shifting the circumstances relative to which *p* is to be semantically evaluated at a world of evaluation *w*. Relative to *w*, “I eat more iron” might be false but “I ought to eat more iron” true. On the account due to Kratzer that is the starting point for many further developments, *ought* determines the set of worlds relevant to modeling those circumstances along two contextual parameters.⁴³ Intuitively, the first is which worlds matter, the second is how we rank them. In the terminology of the theory, the first is a *modal base*: a function *f* from *w* to a set of worlds compatible with *w*. This is a restriction on *w* by whatever background conditions may be selected as relevant in context *c*.⁴⁴ The second is an *ordering source*: a function *g* from *w* to the set of best worlds in the modal base. This yields an ordering over worlds in terms of whichever standards are selected in *c*. Then “*fg*-compatible worlds” will be the set of best worlds determined by *c* and *w*.

The semantic value of *Ought(p)* relative to a context-world pair can now be represented by the following modal rule:

[[*Ought(p)*]]^{*fg*} is true iff all of the *fg*-compatible worlds are *p*-worlds.⁴⁵

43 Kratzer, “The Notional Category of Modality” and “Modality.”

44 The relevant notion of context here includes both linguistic elements (such as previous utterances and topics of conversation) and extralinguistic elements (such as the intentions of the speaker, objects discernible in the environment, the social setting, and common presuppositions).

45 This formulation is from Chrisman, “Two Nondescriptivist Views of Normative and Evaluative Statements,” 410. Double brackets stand for a function that maps a linguistic expression to its semantic value. My main points will go through even if *ought* need not order possible worlds (Ridge, *Impassioned Belief*, 33–35). I suspect that they will also go through even if the Kratzer semantics is replaced by one that connects deontic modals to verdicts of a theory of practical rationality, e.g., in ways proposed by Cariani, “Deontic Modals and Probabilities”; Charlow, “Decision Theory”; and Lassiter, “Linguistic and Philosophical Considerations on

For instance, suppose that “You ought to abstain from premarital sex” is uttered when the question is what it takes to live a Roman Catholic life. The modal rule for *ought* then tells us that the sentence is true just in case you abstain from premarital sex in all worlds that are compatible with something like the following two restrictions. First, a restriction to worlds in which you can engage in premarital sex (*f*). Second, a restriction of those worlds to ones that rank highly by Roman Catholic values (*g*). Interpreted relative to this context, the sentence is, plausibly, true. But now take a different context where the question is what it takes to live a life of high sensory pleasure. Interpreted relative to this context, the sentence is, plausibly, false. The modal rule allows *ought* to be interpreted relative to any ordering with certain formal properties: how well something conforms with the 1911 Boy Scouts Handbook; how excited my cat gets; what is better for a given agent; what is likely to happen anyway (normative fatalism is a thing for some); the standards endorsed by the speaker; some specific moral ideal; the most fundamental moral standards (whatever they may be); or whatever standard (objective or otherwise) is expressed in “I know I morally ought to ϕ , but ought I really to ϕ ?”⁴⁶ This captures our recognition that nearly anything can be felicitously called by terms like *ought*.

4.2. Two Metasemantic Options

Suppose *ought* has roughly the above kind of descriptive semantics. What kind of metasemantic function might the practical role of *ought* play with respect to it? Its practical role will not explain any particular context-invariant features of its meaning, such as why its meaning includes parameters for which worlds matter and how we rank them. Two options seem to remain. One has to do with the metasemantics of the context-sensitivity of *ought*. Perhaps its practical role contributes to explaining its semantic value specifically in its committal uses. The other option is quite different. Perhaps the practical role of *ought* instead plays some significant role in explaining why a given formal semantics is an appropriate model for *ought* in the first place.

Bayesian Semantics.” For these accounts appeal just to formal decision-theoretic constraints on the ranking of options. Issues about deontic logic raised by Broome may be more tricky (“A Linguistic Turn in the Philosophy of Normativity?” 11–13). For responses to various challenges to Kratzer-style semantics for deontic modals, see Bronfman and Dowell, “The Language of ‘Ought,’ and Reasons.”

46 As Worsnip puts it, the contextually selected standards need not be “parochial” but can be “aspirational” (“‘Ought’-Contextualism Beyond the Parochial”). As these examples suggest, standards need not be reified, they may lack handy natural language labels, and their contents may not be immediately obvious or transparent. Cf. Evers, “Moral Contextualism and the Problem of Triviality,” 295–96.

The first view is initially attractive. It promises to solve the Problem of Diverse Uses. Whether a use is committal or not is a difference in context. We might then think that when *ought* is used in a committal way, this can make a difference to the values of its contextual parameters. In this way, the practical role of *ought* might contribute to explaining its semantic value in some cases but not others. To assess this, it is important to see that the modal rule makes no semantic difference between committal and noncommittal uses of *ought*. Nothing in the standard semantics *requires* the contextually supplied ordering sources or modal bases to be such that the results of feeding them into the modal rule for *ought* align with anyone's practical commitments.⁴⁷ In noncommittal uses, the values of contextual parameters are fixed by factors other than practical commitments. Even in committal uses, appeal to practical role should not single out any specific standard to be fed into our formal semantics. Often the selection of an ordering source is a substantive normative matter, at least beyond specifications such as the "correct standards" or the "most fundamental standards." By the same token, however, nothing in the standard semantics *rules out* that contextually relevant motivations or practical commitments may contribute to fixing the semantic value of *ought* in committal contexts. So do they?

If you think that the practical role of normative language sets it apart from nonnormative language in some broadly semantic way, you will want its practical role to be a distinctive and comparatively significant feature of its meaning. Locating the difference in the metasemantics of context sensitivity fails to fulfill this promise. The problem arises from perfectly general considerations. What is

47 One might claim that such a requirement can be motivated as an external constraint on the semantics. Things can be ranked in terms of badness rather than goodness. Formally speaking such a ranking is eligible to help determine the semantic value of *ought* in certain contexts. But could there really be a context in which a competent speaker who says that we ought to do something is thereby asserting that it is what we do in all the maximally *bad* worlds compatible with the modal base? If not, our overall semantic theory should posit a restriction on what rankings can supply the ordering source. One might then suggest that such a restriction has something to do with the relation between these rankings and the practical commitments of some contextually determined agent. In response, let us consider again Iago, who is committed to sheer malignity for its own sake. Sincerely uttering, "You ought to kill, and it is better if you kill than if you study" is morally misguided and perhaps for that reason uncommon. But it need not be linguistically incoherent if the ordering source is the same. (The sentence is true if killing is ranked highly by the contextually selected ordering source, and false if not.) By contrast, "You ought to kill, but it is better if you study than if you kill" is incoherent if the ordering source is the same. Various semantic connections between *good* and *ought* constrain their interpretations independently of the practical roles of these terms (Lassiter, *Graded Modality*, ch. 8). So explaining why *ought* claims rarely assert what is done in maximally bad worlds does not require invoking the practical role of *ought* even in committal uses. It cannot require that in noncommittal uses anyway.

required for interpreting utterances featuring context-sensitive terms? A variety of contextual inputs must combine to set the values of contextual parameters and thereby transform nonpropositional semantic values into compositional semantic values that are truth apt. This may be fairly straightforward for simple demonstratives like *this*, or even conventionalized for indexicals like *I* and *now*. In general, however, just which contextual factors can contribute to determining the semantic value of an expression and how they may combine to do so defies simple generalization. Factors that are in general eligible are many. They include (but may not be limited to) common presuppositions, social influences, the structure of the preceding discourse, and salient objects.⁴⁸ The metasemantics of the ordering source parameter for *ought* is indirect in the same way: multiple factors influence how its value is set in context.⁴⁹ That must be so since in noncommittal uses the semantic value of *ought* is determined by factors other than its practical role anyway. Suppose we have been discussing capitalism and I say, “We ought to prioritize profit over workers’ interests.” It may well be clear from the preceding discourse, or the pins on my jacket, that I am a committed critic of capitalism. I should then be construed as talking about what follows from capitalist values.

If factors other than the practical role of *ought* suffice to determine semantic value in noncommittal contexts, we can expect them sometimes to override contextually relevant practical commitments even in committal contexts. In general, different factors may pull in different directions and, thus, require balancing. For instance, social influences and speaker intentions may conflict. No particular factor enters such a calculation from a position of privilege, unless the conventional meaning of the expression in question so dictates. But we have seen that the practical role of *ought* is not part of its conventional meaning, nor is its presence marked otherwise by linguistic convention. So when practical commitments help us to limit the set of contextually relevant ordering sources in some way, any other way of limiting them in that way would have done just as well. Thus, even when the practical role of *ought* plays an indirect metasemantic role, it plays no distinctive or comparatively significant role in determining the semantic value of *ought*.⁵⁰ So the claim that the practical role of thin normative terms is a particularly significant feature of their meaning cannot be vindicat-

48 Glanzberg, “Context, Content, and Relativism.”

49 On indirect metasemantics in general, see Glanzberg, “Context, Content, and Relativism,” and “Indirectness and Intentions in Metasemantics.” King defends a metasemantics on which the values of contextual parameters are fixed by speakers’ intentions (“The Metasemantics of Contextual Sensitivity”). But even on that view, it is a substantial issue whether or when the ordering source parameter is fixed by practical commitments.

50 Practical upshots can also play an indirect metasemantic role even if they are heterogeneous across contexts in ways that would not support explanatory generalizations.

ed in the metasemantics of their context sensitivity.⁵¹ My reasons for this claim may not be conclusive. The issues here are subtle and the metasemantics of context sensitivity is an underexplored topic in general. But the reasons look strong enough to recommend looking elsewhere.

The second metasemantic view is that the practical role of *ought* is part of what explains why the dominant sort of formal models for modal language provide a good descriptive semantics for terms like *ought* in the first place. For instance, an inferentialist metasemantics might hold (roughly) that a given formal semantics is a good model for *ought* because it explicates a certain kind of inferential structure in which *ought* is nested.⁵² Or, metasemantic expressivism might hold (again, roughly) that a given formal semantics is a good model for *ought* because it appropriately mirrors the structure of mental states that *ought* expresses.⁵³

If the association between thin normative terms and practical role is to be metasemantic in this way, a further claim is required. The explanation of why a certain formal semantics is a good model for *ought* must invoke specifically practical inferences or motivational states. But, on the face of it, an explanation that unifies the committal and noncommittal uses of *ought* does not need this further claim.⁵⁴ The standard semantics implies that committal and noncommittal uses of any deontic sentence of the form *Ought*(*p*) have the same semantic value when the modal base and the ordering source are the same. Explaining noncommittal uses only requires invoking theoretical commitments and cognitive states. By parity, that should suffice also for explaining committal uses. The standard semantics does not care about this distinction between uses. So, on the face of it, explaining why it is a good model for *ought* should not require invoking practical role. (It really is dialectically significant if committal and noncommittal uses of *ought* are uniform in their descriptive semantics!) If that is right, it would complete my case that nothing in the standard semantics for *ought* supports treating

51 In section 5.1, I will suggest that pragmatic accounts of the practical role of *ought* can accommodate the above kind of weak and indirect metasemantic role.

52 Chrisman, *The Meaning of "Ought"*; Tiefensee, "Inferentialist Metaethics, Bifurcations, and Ontological Commitment" and "Metasemantics for the Relaxed."

53 Ridge, *Impassioned Belief*; Köhler, "Expressivism, Meaning, and All That."

54 Silk's general framework for theorizing about normative language looks compatible with this alternative (*Discourse Contextualism*). Silk proposes that "normative uses" of *ought* present the speaker as endorsing the standards that justify the *ought* claim in question (*Discourse Contextualism*, 130–32, 137). Justification is an intranormative notion, and endorsement (which Silk usually applies to endorsing a body of *information*) does not entail affirming specifically practical commitments. Nor is that entailed by his account of normative uses as *presupposing* a set of standards endorsed in the context, analyzed as a contextual parameter whose value speakers can manage by exploiting their mutual grammatical and world knowledge and general pragmatic reasoning skills (*Discourse Contextualism*, 131, 145).

those uses that are associated with a practical role as semantically or metasemantically exceptional. I will now argue that existing forms of metasemantic inferentialism and expressivism do not support assigning the practical role of *ought* a significant metasemantic function of this type. The Problem of Diverse Uses still pushes us not to build the practical role of thin normative terms into our overall semantic theory for these expressions.

4.3. Ideational Expressivism

First consider “ideational expressivism” due to Ridge. Ideationalism is a general metasemantic theory according to which “facts about the semantic contents of meaningful items in natural languages are constituted by facts about how those items are conventionally used to express states of mind.”⁵⁵ What Ridge’s expressivism adds to this general framework is the claim that normative sentences have their meaning in virtue of expressing certain kinds of nonrepresentational, motivational states while descriptive sentences have their meaning in virtue of expressing representational states.⁵⁶ My discussion will not touch ideational expressivism in its most general form, but only those specific forms that assign significant metasemantic work to the practical role of normative terms.

To illustrate, consider Ridge’s broadly Kratzer-style descriptive semantics, on which any use of *ought/must* says roughly something of the following form: “Any standard of contextually specified kind *S* would, relative to a contextually specified set of background information or facts *B*, recommend/require *X*.”⁵⁷ Formally normative standards, such as legal standards, have no essential link to a practical role. But some standards do. If a sentence like “You ought to give to charity” is used to make an “all things considered practically normative” claim, it can be paraphrased as “Any acceptable standard of practical reasoning would, given contextually specified circumstances, recommend that you give to charity.”⁵⁸ Normative contexts are to be understood narrowly as those in which the relevant sort of standard is “any acceptable standard of practical reasoning.”⁵⁹ In such contexts, Ridge claims, *ought* and *must* express certain kinds of motivational states.⁶⁰ In other contexts, *ought* and *must* are used nonnormatively.

55 Ridge, *Impassioned Belief*, 107.

56 Ridge, *Impassioned Belief*, 110–11.

57 Ridge, *Impassioned Belief*, 28.

58 Ridge, *Impassioned Belief*, 40.

59 A standard of practical reasoning is a standard that is treated by the agent as a standard of practical reasoning and that can function in a certain action-guiding way (Ridge, *Impassioned Belief*, 40).

60 Ridge, *Impassioned Belief*, 19–21.

So Ridge acknowledges that terms like *ought* can be put to diverse uses. However, using ideational expressivism to explain why the standard semantics is a good model for *ought* requires no reference to the practical role of *ought*. No such reference is required for noncommittal uses of *ought* to state recommendations. A further good question is whether one can use *ought* noncommittally to state what would be recommended by any acceptable standard of practical reasoning.⁶¹ (Why not?) To assess *ought* sentences that count as normative under Ridge's narrow definition, consider his preferred account of expression:

Accountability Expression: A declarative sentence “*p*” in sense *S* in a natural language *N* used with assertive force in a context of utterance *C* expresses a state of mind *M* if and only if conventions which partially constitute *N* dictate that someone who says “*p*” in sense *S* in *C* with assertive force is thereby *liable* for being in state *M*.⁶²

Whether an assertive use of *ought* states what would be recommended by any acceptable standard of practical reasoning is a function of features of the context of utterance that do not normally ground linguistic convention.⁶³ So the relevant linguistic conventions would have to enter someplace else. Ridge does allude to “a theory of how the linguistic conventions provide a function from contexts of utterance to states of mind for any given sentence of the language.”⁶⁴ But what linguistic conventions of English dictate that assertive utterances of sentences of the form *Ought*(*p*) to state what would be recommended by any acceptable standard of practical reasoning make the speaker liable to be, specifically, in a motivational state? We saw in section 4.2 that they will not be conventions per-

61 According to Ridge, to decide that a course of action is acceptable in a given set of circumstances is “in some sense to decide that the course of action is not ruled out for purposes of your deliberation—that it is still ‘on the table’” (*Impassioned Belief*, 41). Note that some noncommittal uses of *ought* may satisfy this description. In 5, for instance, the speaker is not (yet) engaged in deliberation but treats going shopping as being still on the table.

62 Ridge, *Impassioned Belief*, 109.

63 Ridge notes that we can use modifiers like *really* and *genuine* to flag that we are using *ought* or *must* in his narrowly normative sense. For instance, we can acknowledge that etiquette requires something but query whether that is what we *really* ought to do or have any *genuine* reason to do (*Impassioned Belief*, 20). But this effect of *really/genuine* may be better explained in the pragmatics. It tells us nothing much about the meanings of assertive utterances of *ought* sentences without such modifiers. Nor is a narrowly normative meaning of “really ought to” a function of semantic composition. What *genuine* and *really* contribute to composition in these contexts is the same as when we ask whether refraining from noting a smudge on someone's face is *genuinely* a polite thing to do, or whether going on a bar crawl against my parents' wishes is *really* a fun thing to do.

64 Ridge, *Impassioned Belief*, 130.

taining to how semantic value relative to context is computed from context as an operation on the modal rule for *ought*.⁶⁵ A more general point can be made even in the absence of the kind of general theory to which Ridge alludes. Whether ideational expressivism explains why the standard semantics is a good model for *ought* does not depend on there being conventions that map the relevant uses of *ought* specifically onto motivational states. It is hard to see what conventions of English should rule out accounting for assertions concerning what would be recommended by acceptable standards of practical reasoning in terms of nonmotivational states of mind, such as beliefs about what such standards recommend. In this vicinity lie also good but underexplored questions about how philosophically committal we should want our metasemantic theories to be.

Given Accountability Expression, a sentence can express *M* without its being the case that the speaker actually is in *M*. One might then reply that *ought* sentences have the semantic values they have in virtue of expressing the relevant kind of motivational attitudes in all contexts. What happens in noncommittal uses is that further contextual information implies that the speaker's state of mind does not exemplify that structure. But this reply is problematic. First, Accountability Expression does not support it. In section 3, we saw some reason to think that the liabilities incurred by noncommittal uses of *ought* are not plausibly

65 Ridge's own view may be more like the view that the practical role of *ought* does work in the metasemantics of its context sensitivity. In the main text I discuss the prospects of recruiting his view to a different service, and so am not attributing all of the moves I discuss to Ridge himself. But ideational expressivism faces trouble also with showing that the practical role of *ought* plays a special role in explaining semantic value relative to context. Ridge individuates orderings finely: "A moral standard provides a different ordering from a legal standard, even in the case in which the law requires all and only the morally required actions" (*Impassioned Belief*, 36). Consider two speakers who both assert the same *ought* sentence, such as, "One ought to report one's liability." The contextually specified standards may be isomorphic in the orderings they induce but such that one speaker treats the standard as an acceptable standard of practical reasoning whereas the other does not. (Two speakers can differ in this way with respect to moral standards, legal standards, and more.) Ridge's view implies that the contexts involve different standards, so the two utterances have distinct semantic values. That may be fine if the contexts are separate, though I myself find it odd if solely a difference in the motivational states that the speakers are liable to be in implies that they say different things. But suppose the speakers are involved in a conversation. It would be implausible to posit constant context shifts depending on whether the speaker at the time treats a specific standard as an acceptable standard of practical reasoning. Modeling such a conversation requires just a shared conception of what actions are recommended and a way of tracking what the speakers are planning to do about that. Moreover, it should be possible to report the conversation by saying "Timmy and Tammy both think that one ought to report one's liability. Tammy is planning to report hers; Timmy is not." If their utterances said different things, such a report should be either false or odd. But it is an accurate report that would be felicitous in the described context.

parasitic on those incurred by its committal uses. We still have no evidence that a parasitic relation holds in such a way that noncommittal speakers are liable to be in a motivational state by linguistic convention. So far as the conventions that constitute English go, deontic assertions can uniformly express states representing what verdicts follow from contextually specified standards. Second, the standard semantics treats deontic talk as expressing what requirements follow from such standards. Nothing in it suggests that this semantic function of deontic talk is somehow parasitic on some more immediately practical function.⁶⁶ Explaining why it is a good model for *ought* thus does not require invoking the practical role of *ought*. I conclude that ideational expressivism does not help broadly semantic views to solve the Problem of Diverse Uses.

4.4. Inferentialist Metasemantics

Now turn to inferentialist metasemantics for normative language. Chrisman takes the truth conditions predicted by the Kratzer-style modal rule to articulate how speakers “have to think and reason in order to satisfy the implicit conceptual commitments affirmed by using “ought” to make an assertion in ordinary discursive practice.”⁶⁷ My discussion will not touch metasemantic inferentialism in this general form, but only those specific forms that assign significant metasemantic work to the practical role of normative terms.

Existing inferentialist metasemantics for normative language take this form. Chrisman acknowledges that some uses of *ought* merely commit one to concluding the verdicts of the contextually selected ranking. But he proposes that the commitments affirmed by “genuinely normative” uses could “be conceived as commitments to reason practically in certain ways.”⁶⁸ The key idea is that the general metasemantic function of *ought* is still the same across these local differences. Tiefensee similarly proposes to understand evaluative terms such as *good* in terms of a general metaconceptual function of structuring and explaining the legitimacy of certain language exit transitions to intentions, actions, and so on.⁶⁹ Transposed to the deontic key of this paper, this is to analyze *ought* as a

66 Ridge suggests that normative discourse might be like imperatival discourse, in that although imperatives can serve subsidiary communicative functions, these are “somehow parasitic on the more immediately practical function of the imperative form” (*Impassioned Belief*, 21). Here I question the parallel. See also note 27 for the possibility that the practical function of the imperative form is a kind of pragmatic phenomenon.

67 Chrisman, “Two Nondescriptivist Views of Normative and Evaluative Statements,” 415; cf. Chrisman, *The Meaning of “Ought,”* ch. 5.

68 Chrisman, “Two Nondescriptivist Views of Normative and Evaluative Statements,” 416.

69 Tiefensee, “Metasemantics for the Relaxed.” For reasons of space, I simplify a great deal. I bracket Tiefensee’s view that we need to begin with words like *good* because *ought* presup-

linguistic instrument for explicating certain commitment structures. Asserting “I ought to help my sister,” for instance, explains why a commitment to a premise like “Babysitting the boys will help my sister” provides a reason to enter into the commitment that I shall babysit the boys, and likewise for premises concerning what one has promised, and so on.⁷⁰

These views acknowledge that terms like *ought* can be put to diverse uses. But they, too, struggle with the Problem of Diverse Uses. Recall that, on the standard semantics, an *ought* sentence will have the same truth conditions across committal and noncommittal uses when contexts supply the same modal base and ordering source. “We ought to report our liability,” for instance, says that reporting our liability ranks highly on the contextually determined standard. On the face of it, explaining why the standard semantics is a good model for a (deontic) *ought* should then be insensitive to differences in practical upshot. Metasemantic inferentialism would be more general and unified if it modeled *ought* as explicating structures whose language exit transitions consist in theoretical commitments. Examples would be commitments to form certain beliefs about what is required or recommended and to recognize reasons for action that correspond to the same ordering source.⁷¹ This would account for one central function of deontic vocabulary of English: stating what requirements or recommendations follow from various sorts of standards to which the interlocutors need not be practically committed. Its further function to convey practical commitments could well figure in some other part of our overall theory of meaningful communication, such as pragmatics.⁷²

For metasemantic inferentialism to solve the Problem of Diverse Uses, one of two different views must instead be true. One is the view that Chrisman gestures at: committal and noncommittal uses of *ought* explicate different commitment structures; the structures involved in committal uses feature not only the

poses evaluative orderings (“Metasemantics for the Relaxed”). As far as I can see myself, an ordering need not itself have evaluative content, and only a notion of ranking highly on an ordering (rather than, e.g., being among the *best* in any evaluatively laden sense) is required for identifying the set of recommended items.

70 Cf. Tiefensee, “Metasemantics for the Relaxed,” 122, 125.

71 Contrary to what Tiefensee seems to suggest, it is not sufficient for practical commitment in the relevant sense that a claim of the form “I ought to ϕ ” gives a *pro tanto* justification or reason to ϕ (“Inferentialist Metaethics, Bifurcations, and Ontological Commitment,” 2444). Recall note 17.

72 Note a related issue about the metasemantic adequacy of versions of inferentialism that appeal to indispensably practical commitment structures. If modal expressions are polysemous across different “flavors” of modality, the polysemy does not seem to arise from differences in commitment structures. The arguments for modal polysemy in Viebahn and Vetter are not like that, for instance (“How Many Meanings for ‘May’?”).

oretical but also practical commitments. The other is the view that they explicate the same commitment structures and these are indispensably practical. Anyone making a deontic assertion using *ought* would normally be understood as undertaking the commitments that the sentence makes explicit, but further contextual information may suggest that the speaker's assertion only reports that structure. However, neither view explains why the standard semantics is a good model for *ought*. Nor does either view support the distinct view that the practical role of *ought* constrains our overall semantic theory in some other way.

If practical commitments were indispensable to the commitment structures that *ought* explicates or affirms, the opposite thesis should fall to a *reductio*. But it does not. Suppose that the language exit moves in the commitment structures that deontic *ought* makes explicit consist wholly in theoretical commitments. Deontic *ought* claims would then make the same kind of commitment structure explicit in both committal and noncommittal uses, and language exits would still be to states that are about actions. This variant of metasemantic inferentialism could equally well say that anyone making a deontic assertion using *ought* would normally be understood as undertaking the commitments that the sentence makes explicit, but further contextual information may suggest that the speaker's assertion only reports that structure. For it is perfectly compatible with all of the above that speakers who assert *ought* sentences would normally be understood, on the basis of pragmatic inferences, as undertaking *also* some practical commitments. So the appeal to practical commitments does not seem indispensable in explaining why the standard semantics is a good model for *ought*.

This leaves the view that committal and noncommittal uses of *ought* explicate different commitment structures. For this to solve the Problem of Diverse Uses, the distinction between committal and noncommittal uses of normative language must be semantically significant. Chrisman suggests that "genuinely normative" uses of *ought* differ conceptually from the rest. In some sense that must be right, since there is a distinction between committal and noncommittal uses to be drawn. But this minimal sense is not enough here. We have seen that the distinction is not lexically encoded or marked by linguistic convention. We have also seen that prospects are poor for showing that noncommittal uses of normative language are semantically parasitic on some more immediate practical function of normative language. Nor have we been shown any other evidence that the distinction plays a role in explaining why the standard semantics is a good model for *ought*. As we have seen, if a word can be used in certain contexts to affirm or explicate certain implicit conceptual commitments, this is not automatically something to be reflected in our overall semantic theory. So even if inferentialism explains why the standard semantics provides a good model for

terms like *ought*, we have yet to see why that explanation must feature not only theoretical but also practical commitments. Differences between committal and noncommittal uses might instead be captured in some other part of our overall theory of meaningful communication.

Time to conclude. In this section I introduced the dominant sort of formal semantics for deontic modal expressions like *ought*. I then argued that explaining why deontic *ought* has this kind of descriptive semantics does not require appeal to the practical role it plays in many contexts. My argument went through all the places I can think of where a practical role might show up: the metasemantics of the context-invariant semantic features of *ought*; the metasemantics of its context-sensitivity; and explaining why a given kind of formal semantics is in general a good model for *ought*. Existing expressivist and inferentialist metasemantics for normative language do not support the claim that the practical role of such language is a distinctive and particularly significant feature of its meaning. And yet metasemantics seems to be the most promising place for vindicating broadly semantic accounts of the practical role of thin normative terms. I cannot claim to have conclusively ruled out that the practical role of normative language is a metasemantic phenomenon. Caution is due: the issues here are subtle and underexplored and my argument has many moving parts. But at minimum my argument sets a demanding bar to be met.

5. BROADER IMPLICATIONS

I will close by discussing the prospects of explaining the practical role of normative language in pragmatic rather than broadly semantic terms and highlighting some broader implications of my negative argument for metaethics and the philosophy of normativity.

5.1. *Looking Ahead to Pragmatics*

The challenges that I have raised against broadly semantic accounts of the practical role of thin normative terms are significant because generalizations that do not rise to full generality are normally more apt to be explained in the pragmatics. Explanations of linguistic phenomena that appeal to general principles of communication instead of special theoretical posits are more unified and parsimonious. So if a phenomenon can be explained in the pragmatics, then (all else equal) it should not be explained semantically. The obvious follow-up question is whether the practical role of thin normative terms can be adequately explained as a pragmatic phenomenon. If it cannot, that would support introducing suitable posits into our overall semantic theory.

The general character of pragmatic reasoning concerns what can be inferred from the following sort of premise: the speaker believed that uttering a given particular sentence in the given context was best for promoting her conversational ends. Determining whether the practical role of normative speech can be derived through broadly pragmatic reasoning thus requires answering the following sorts of questions: In pursuit of what conversational ends may we engage in normative speech? By means of what kind of normative speech do we take those ends to be best achieved? Answering these questions is well beyond the scope of this paper. But there is reason to be optimistic about a pragmatic account. Everyone will need an account of what makes particular *uses* of words normative and what makes them practically oriented. A very wide range of non-normative terms can be contingently associated with practical upshots, given a suitable context. For instance, saying that a knife is sharp can convey a disposition or commitment to use it when it is common ground that a hard loaf of bread needs slicing. Any association between uttering “This knife is sharp” and such practical upshots is pragmatic. Plenty of theoretical space thus exists for a pragmatic account of the practical role of normative language.

One possibility is to adapt the pragmatic account from Finlay.⁷³ Finlay develops a certain kind of “end-relational” semantics of normative terms. He then argues that we can predict and explain the practicality of normative talk as a feature of “how we use normative language in context, pursuing our desired ends” by using “only maximally simple and conservative principles of pragmatics.”⁷⁴ The broad kind of pragmatic story might well be similar regardless of whether the relevant inputs from context are ends, rankings, or whatnot. Here is another possibility. Suppose we are particularly concerned to explain why we normally do not need to work it out from scratch whether the speaker of a bare first-person normative utterance is committed to acting or reacting accordingly. Here we might be able to adapt accounts that treat certain features of normative discourse as “generalized” pragmatic features. When a pragmatic feature is generalized, it arises from saying a certain thing in the absence of conversational moves or other special circumstances that would revoke the implication.⁷⁵ Strandberg argues that the motivational implications of moral assertions are generalized conversational implicatures.⁷⁶ Väyrynen argues that the evaluative implications of “thick” terms, such as *chaste* and *generous*, are default but defeasible “not-at-is-

73 Finlay, *Confusion of Tongues*, ch. 5.

74 Finlay, *Confusion of Tongues*, 116.

75 Here I generalize from Grice’s notion of generalized conversational implicature (*Studies in the Way of Words*, 37–38).

76 Strandberg, “A Dual Aspect Account of Moral Language.”

sue contents” of their assertive uses in normal contexts.⁷⁷ The practical upshots with which thin normative terms are associated might be susceptible to one of these types of explanation.

A more full, pragmatic account of the practical role of thin normative terms is beyond the scope of this paper. But another optimistic note worth sounding is that the practical role of thin normative terms can both be a pragmatic phenomenon and play the kind of indirect metasemantic role described in section 4.2. Suppose someone makes a claim about what we ought to do but you do not immediately get just what they are saying. One way you can try to understand them is by searching for practical commitments that would make sense of the utterance. For instance, information about the speaker’s or another salient agent’s practical commitments might help to identify standards that would recommend the action in question.⁷⁸ (Other ways include a closer look at the preceding discourse, and so on. Recall the discussion of indirect metasemantics of context sensitivity in section 4.2.) Speakers often can expect their audience to interpret their utterances in this way by exploiting their grammatical knowledge (that *ought* requires a ranking, and so on), world knowledge, and general pragmatic reasoning skills. The general point is that a context-sensitive expression with an indirect metasemantics may well be such that a factor can help fix the value of a contextual parameter in a context of its utterance and yet get communicated by the utterance as a matter of pragmatics. My arguments suggest that contextually relevant practical commitments may be one such thing.

5.2. *Implications for Metaethics and the Philosophy of Normativity*

If my arguments against broadly semantic accounts of the practical role of normative language are on the right track, they put significant pressure on a range of prominent accounts of normative language. As my discussions of Hare, Blackburn, Gibbard, and Ridge suggest, they put pressure on expressivist accounts of the meaning of normative language in terms of plans or other conative states.⁷⁹ As my discussions of Wedgwood, Chrisman, and Tiefensee suggest, they put similar pressure on versions of conceptual or inferential role theories that aim to explain the meaning of normative language in terms of its practical role. These

77 Väyrynen, *The Lewd, the Rude and the Nasty*, ch. 5.

78 For one more detailed discussion of these kinds of aspects of normative discourse, see Silk, *Discourse Contextualism*, 126–32.

79 Some recent work interprets expressivism more broadly as a view according to which states of mind are fundamental in a theory of meaning; cf. Ridge, *Impassioned Belief*; Charlow, “Prospects for an Expressivist Theory of Meaning”; Silk, *Discourse Contextualism*. As emphasized in section 4.3, my arguments touch such views only insofar as they appeal to essentially practically oriented states of mind.

arguments also challenge those moral error theories according to which it is a nonnegotiable commitment of our moral discourse that such discourse plays a certain kind of practical role.

My arguments also show that we need to be much more careful in moving between claims about normative thought and claims about normative talk. To be clear, the arguments are compatible with there being normative concepts to which a practical role is essential. There are downstream questions, such as how we might manage to express such concepts if expressing them is not a semantic or metasemantic feature of normative language. But the immediate lesson is just that even if there are OUGHT concepts that are conceptually linked to motivation or practical commitment, this link is reflected neither in our descriptive semantics nor our metasemantics for the corresponding segment of English. Normative discourse can then be in perfectly good working order without predicates that are associated with certain practical roles in any broadly semantic way. This does not seem descriptively inadequate. For instance, my arguments raise no deep problems for explaining why *ought* is characteristically used to offer advice, *must* to issue more insistent sort of requirements and orders, and so on. If their characteristic semantic function is to describe recommendations or requirements that follow from some assumed ranking, they will be naturally suited for directive conversational effects.

How my arguments bear on the philosophy of normativity is more complex. Some deny that discussions focused on how normative terms work in natural language are of much import to philosophers of normativity. For example, Broome argues that philosophers of normativity “do not have to be much bothered by the fluid and contextual nature of ‘ought’ in common English.”⁸⁰ There is a central practical meaning of *ought*, specified by an “enkratic” rational requirement on intentions. The philosophy of normativity is principally concerned with the metaphysical nature of this ought, what determines when it obtains, and so on.

I will not here discuss whether Broome is right that the enkratic OUGHT is at the center of normativity. The main lesson of my paper is that even if a given OUGHT is privileged with respect to the philosophy of normativity, it will not be linguistically privileged. For example, even when *ought* is used to express an authoritative OUGHT that resolves conflicts between explicitly relativized notions like “moral ought” and “prudential ought” and specifies what we ought to do without qualification, that is just a special case of the standard semantics. For it to be a well-behaved use of *ought*, it must induce a ranking of the things that matter in the context.⁸¹ We might not have a dedicated name for such a ranking. But

80 Broome, “A Linguistic Turn in the Philosophy of Normativity?” 10.

81 Broome allows that there may be such a “final ordering”; “A Linguistic Turn in the Philoso-

if there was not one, this OUGHT would not be able to resolve conflicts between various explicitly relativized oughts. Regarding the enkratic OUGHT, it would not be clear why rationality should require that you intend to do what you judge you ought to do (in this sense) unless it was something that mattered highly in the context.

We probably will not be able to say much at all about the content of the ranking induced by any such privileged OUGHT without doing substantive normative theory. But that is as it should be. We should not expect semantic theory to tell us much about normative theory. Nor, however, should normative theory do all that much to shape or drive semantic theory. So I counsel caution and reconciliation. Even if philosophers of normativity have identified a specific concept that lies at the center of normativity and can be expressed by *ought*, this implies little regarding how its practical role is reflected in language. Even if a given OUGHT is individuated by a practical role that involves distinctive conceptual commitments, it does not follow that those commitments are to be reflected in our overall semantic theory. Just as semanticists should avoid exaggerated claims about normativity, philosophers of normativity should be sensitive to how the core features of their central concepts are reflected in language.⁸²

University of Leeds
p.vayrynen@leeds.ac.uk

phy of Normativity?" 10.

- 82 This paper was four years in the making, which feels long for something that began as a fun side project and remained a side project. My bumbling early runs at its core argument probably count as a distinct paper even by a fairly coarse-grained individuation of philosophy papers. That predecessor ("Evaluative Terms and Normative Role") received valuable feedback from audiences at University of Leeds, University of Geneva, Uppsala University, Humboldt University of Berlin, and the New York Philosophy of Language Workshop. Something more like the present paper was helped along by discussions at University of Cambridge and workshops at NYU Abu Dhabi and Frankfurt School of Finance and Management. I'm grateful to Alex King as my commentator in Abu Dhabi and to Sarah Buss, Matti Eklund, Will Gamester, Camil Golub, Matthew Kramer, Nick Laskowski, Tristram McPherson, Eliot Michaelson, Christine Tiefensee, Robbie Williams, and Jack Woods for conversations and/or comments on various versions of the material. Comments from anonymous referees for *JESP* and many other journals along the way also led to significant improvements. Many thanks to you all.

REFERENCES

- Alwood, Andrew. "Should Expressivism Be a Theory at the Level of Metasemantics?" *Thought* 5, no. 1 (March 2016): 13–22.
- Bar-On, Dorit, and Matthew Chrisman. "Ethical Neo-Expressivism." *Oxford Studies in Metaethics*, vol. 4. Edited by Russ Shafer-Landau, 133–65. Oxford: Oxford University Press, 2009.
- Bedke, Matthew. "Practical Oomph: A Case for Subjectivism." *Philosophical Quarterly* 69, no. 277 (October 2019): 657–77.
- Blackburn, Simon. *Ruling Passions: A Theory of Practical Reasoning*. Oxford: Clarendon Press, 1998.
- Bronfman, Aaron, and J.L. Dowell. "The Language of 'Ought,' and Reasons." In *The Oxford Handbook of Reasons and Normativity*, edited by Daniel Star, 85–112. Oxford: Oxford University Press, 2018.
- Broome, John. "A Linguistic Turn in the Philosophy of Normativity?" *Analytic Philosophy* 57, no. 1 (March 2016): 1–14.
- Cariani, Fabrizio. "Deontic Modals and Probabilities: One Theory to Rule Them All?" In Charlow and Chrisman, *Deontic Modality*, 11–46.
- Charlow, Nate. "Decision Theory: Yes! Truth Conditions: No!" In Charlow and Chrisman, *Deontic Modality*, 47–81.
- . "Prospects for an Expressivist Theory of Meaning." *Philosophers' Imprint* 15, no. 23 (August 2015): 1–43.
- Charlow, Nate, and Matthew Chrisman, eds. *Deontic Modality*. Oxford: Oxford University Press, 2016.
- Chrisman, Matthew. *The Meaning of "Ought": Beyond Descriptivism and Expressivism in Metaethics*. Oxford: Oxford University Press, 2015.
- . "Two Nondescriptivist Views of Normative and Evaluative Statements." *Canadian Journal of Philosophy* 48, nos. 3–4 (2018): 405–24.
- Copp, David. "Realist-Expressivism: A Neglected Option for Moral Realism." *Social Philosophy and Policy* 18, no. 2 (Summer 2001): 1–43.
- Dreier, James. "Internalism and Speaker Relativism." *Ethics* 101, no. 1 (October 1990): 6–26.
- Eklund, Matti. *Choosing Normative Concepts*. Oxford: Oxford University Press, 2017.
- Evers, Daan. "Moral Contextualism and the Problem of Triviality." *Ethical Theory and Moral Practice* 17, no. 2 (April 2014): 285–97.
- Ferguson, Charles A. "Baby Talk in Six Languages." *American Anthropologist* 66, no. 6 (December 1964): 103–14.

- Finlay, Stephen. *Confusion of Tongues: A Theory of Normative Language*. Oxford: Oxford University Press, 2014.
- . “Defining Normativity.” In *Dimensions of Normativity: New Essays on Metaethics and Jurisprudence*, edited by David Plunkett, Scott Shapiro, and Kevin Toh, 187–219. Oxford: Oxford University Press, 2019.
- Finlay, Stephen, and Justin Snedegar. “One Ought Too Many.” *Philosophy and Phenomenological Research* 89, no. 1 (July 2014): 102–24.
- Gibbard, Allan. *Thinking How to Live*. Cambridge, MA: Harvard University Press, 2003.
- Glanzberg, Michael. “Context, Content, and Relativism.” *Philosophical Studies* 136, no. 1 (October 2007): 1–29.
- . “Indirectness and Intentions in Metasemantics.” In *The Architecture of Context and Context-Sensitivity: Perspectives from Philosophy, Linguistics and Logic*, edited by Tadeusz Ciecierski and Pawel Grabarczyk, 29–53. Basingstoke, UK: Springer, 2020.
- . “Lexical Meaning, Concepts, and the Metasemantics of Predicates.” In *The Science of Meaning: Essays on the Metatheory of Natural Language Semantics*, edited by Derek Ball and Brian Rabern, 197–225. Oxford: Oxford University Press, 2018.
- Grice, Paul. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press, 1989.
- Hare, R. M. *The Language of Morals*. Oxford: Clarendon Press, 1952.
- Kaufmann, Magdalena. *Interpreting Imperatives*. Basingstoke, UK: Springer, 2012.
- King, Jeffrey. “The Metasemantics of Contextual Sensitivity.” In *Metasemantics: New Essays on the Foundations of Meaning*, edited by Alexis Burgess and Brett Sherman, 97–118. Oxford: Oxford University Press, 2014.
- Klimczyk, Joanna. “Compositional Semantics and Normative ‘Ought.’” *Axiomathes* 31, no. 3 (June 2021): 381–99.
- Köhler, Sebastian. “Expressivism, Meaning, and All That.” *Canadian Journal of Philosophy* 48, nos. 3–4 (2018): 337–56.
- Kramer, Matthew H. “Shakespeare, Moral Judgments, and Moral Realism.” In *The Routledge Companion to Shakespeare and Philosophy*, edited by Craig Bourne and Emily Caddick Bourne, 235–45. London: Routledge, 2019.
- Kratzer, Angelika. “Modality.” In *Semantics*, edited by Arnim von Stechow and Dieter Wunderlich, 639–50. Berlin: De Gruyter Mouton, 1991.
- . “The Notional Category of Modality.” In *Words, Worlds, and Contexts: New Approaches in Word Semantics*, edited by Hans-Jürgen Eikmeyer and Hannes Rieser, 38–74. Berlin: Walter de Gruyter, 1981.

- Lassiter, Daniel. *Graded Modality: Qualitative and Quantitative Perspectives*. Oxford: Oxford University Press, 2017.
- . “Linguistic and Philosophical Considerations on Bayesian Semantics.” In Charlow and Chrisman, *Deontic Modality*, 82–116.
- Mandelkern, Matthew. “Practical Moore Sentences.” *Noûs* 55, no. 1 (March 2021): 39–61.
- Ninan, Dilip. “Two Puzzles about Deontic Necessity.” In *New Work on Modality: MIT Working Papers in Linguistics*, edited by Jon Gajewski, Valentine Hacquard, Bernard Nickel, and Seth Yalcin, 149–78. Cambridge, MA: MIT Press, 2005.
- Raz, Joseph. *The Concept of a Legal System*. 2nd ed. Oxford: Clarendon Press, 1980.
- Ridge, Michael. *Impassioned Belief*. Oxford: Oxford University Press, 2014.
- Ruiz, Andrés Soria, and Isidora Stojanovic. “On Linguistic Evidence for Expressivism.” *Royal Institute of Philosophy Supplement* 86 (October 2019): 155–80.
- Schroeder, Mark. “Ought, Agents, and Actions.” *Philosophical Review* 120, no. 1 (January 2011): 1–41.
- Sharadin, Nate, and Rob van Someren Greve. “Is Deontic Evaluation Capable of Doing What It Is For?” *Journal of Ethics and Social Philosophy* 19, no. 3 (March 2021): 203–29.
- Silk, Alex. *Discourse Contextualism*. Oxford: Oxford University Press, 2016.
- Simons, Mandy, and Kevin Zollman. “Natural Conventions and Indirect Speech Acts.” *Philosophers’ Imprint* 19, no. 9 (February 2019): 1–26.
- Stalnaker, Robert. “Reference and Necessity.” In *A Companion to the Philosophy of Language*, edited by Bob Hale and Crispin Wright, 534–54. Oxford: Blackwell, 1997.
- Strandberg, Caj. “A Dual Aspect Account of Moral Language.” *Philosophy and Phenomenological Research* 84, no. 1 (January 2012): 87–122.
- Tiefensee, Christine. “Inferentialist Metaethics, Bifurcations, and Ontological Commitment.” *Philosophical Studies* 173, no. 9 (September 2016): 2437–59.
- . “Metasemantics for the Relaxed.” *Oxford Studies in Metaethics*, vol. 16, edited by Russ Shafer-Landau, 108–33. Oxford: Oxford University Press, 2021.
- Väyrynen, Pekka. *The Lewd, the Rude and the Nasty: A Study of Thick Concepts in Ethics*. Oxford: Oxford University Press, 2013.
- Viebahn, Emanuel, and Barbara Vetter. “How Many Meanings for ‘May’? The Case for Modal Polysemy.” *Philosophers’ Imprint* 16, no. 10 (June 2016): 1–26.
- Wedgwood, Ralph. “Conceptual Role Semantics for Moral Terms.” *Philosophical Review* 110, no. 1 (January 2001): 1–30.

———. *The Nature of Normativity*. Oxford: Oxford University Press, 2007.

———. *The Value of Rationality*. Oxford: Oxford University Press, 2017.

Woods, Jack. “Expressivism and Moore’s Paradox.” *Philosophers’ Imprint* 14, no. 5 (March 2014): 1–12.

Worsnip, Alex. “‘Ought’-Contextualism Beyond the Parochial.” *Philosophical Studies* 176, no. 11 (November 2019): 3099–119.

Yalcin, Seth. “Semantics and Metasemantics in the Context of Generative Grammar.” In *Metasemantics: New Essays on the Foundations of Meaning*, edited by Alexis Burgess and Brett Sherman, 17–54. Oxford: Oxford University Press, 2014.

FREEDOM TO ROAM

Matthias Brinkmann

IN THIS PAPER, I look at the right to roam, which is the right of the public to access private lands. This might sound oxymoronic. Is it not in the nature of private property that owners can exclude whomever they want, under a wide range of circumstances? Can you even be considered to fully own anything if you do not have the right to say no to others accessing it? Assumptions in favor of property owners' rights to exclude are deeply ingrained in American law and many other legal systems, are accepted widely by left-liberal philosophers, and are part of a prephilosophical view of private property.¹

By contrast, the Swedish *allemansrätten* ("everyman's right") gives everyone the right to walk freely across any part of the Swedish countryside; a right to pick wild berries, mushrooms, and flowers; and a right to swim or boat in lakes and rivers. It also extends beyond walking to biking, skiing, and horseback riding. There are limits to the right: roamers must respect the privacy of landowners, must stay away from houses and their curtilage, and must not damage the land—in particular, they must not cross fields or other agriculturally used lands. But crucially, all land, whether privately or publicly owned, is subject to the *allemansrätten*: landowners have no rights to exclude roamers or to put up fences or other obstacles.

The right to roam is an established legal right in other countries too, primarily in Scandinavia, but weaker versions of it can also be found in some northern European countries, such as Switzerland, England, and Scotland. However, there has been no serious philosophical reflection on this right—not a single treatment in a philosophical journal. The right to roam has been discussed in legal journals, where it is a minor theater in the larger war about the right to exclude

1 There are many exceptions we can ignore. For example, the right of owners to exclude is more circumscribed where their property is used in "quasi-public" ways—e.g., shopping malls, restaurants, and hotels (Gray and Gray, "Civil Rights, Civil Wrongs and Quasi-Public Space").

in property law.² While there is much to like in this literature, a focus on law can quickly make us lose sight of the important philosophical issues.

In this paper, I offer a general defense of the right to roam. First, I give a brief overview of how the right to roam is understood and implemented in existing legal systems (section 1). Then, I show that defending the right to roam on broadly liberal grounds is quite difficult, as an interest-based defense of the right fails (section 2) and a left-liberal view of property rights favors rights to exclude (section 3). Notwithstanding these difficulties, I suggest that the right to roam can be defended on nonideal grounds (section 4): it has symbolic value, reminding the haves of their fundamentally equal co-ownership of the world with the have-nots (sections 5 and 6).

1. THE RIGHT TO ROAM

At a minimum, the right to roam is the right to traverse certain parts of the countryside on foot, irrespective of the ownership of these lands, as long as one respects the privacy of the owners and traverses responsibly and without damaging the land. A corresponding duty of landowners is that they not impede roamers—that they not put up fences and signs or other obstacles keeping roamers out. The right to roam generally excludes motorized transport, but it may extend to biking, horseback riding, swimming, hiking, and canoeing.

All countries with a right to roam exclude houses and their curtilage (the identifiable area belonging to a house, such as gardens, sheds, and patios) from its scope, and it is generally understood that roamers must respect the privacy of owners—for example, by keeping a certain distance from houses.³ A second common restriction concerns land used for agricultural purposes. The Norwegian Outdoor Recreation Act, for example, excludes “farmyards, plots around houses and cabins, tilled fields, hay meadows, cultivated pasture, young plantations and similar areas,” and permits traversing those areas only under restrictive circumstances.⁴

- 2 The most philosophically helpful paper, which I will engage with below, is Anderson, “Britain’s Right to Roam.” See also Lovett, “Progressive Property in Action.” Other references to the legal literature can be found throughout. For an account of the wider issues in property law, see Klick and Parchomovsky, “The Value of the Right to Exclude”; and Wyman, “The New Essentialism in Property.” Useful defenses of roaming aimed toward a popular audience can be found in Shoard, *Right to Roam*; and Ilgunas, *This Land Is Our Land*.
- 3 English law, e.g., imposes a minimum distance of twenty meters to be kept from houses (Countryside and Rights of Way Act, 2000, c. 37, sch. 1, pt. 1, para. 3).
- 4 Lov om friluftslivet [Outdoor Recreation Act], July 1, 1957, art. 1a (Norway), <https://www.regjeringen.no/en/dokumenter/outdoor-recreation-act/id172932/>.

Beyond this shared core, we can broadly distinguish two models of the right to roam: the Scandinavian model, which recognizes strong, far-reaching, and relatively open-ended rights to roam the countryside, and the northern European model, which recognizes only weaker and limited rights to roam. The former is the more interesting, and the one I will aim to defend.⁵

The right to roam in the northern European model tends to be a narrow right to walk across certain types of lands. A case in point is English law, which recognizes a right to walk across so-called access lands in the Countryside and Rights of Way Act 2000. Only about 6.5 percent of English land is classified as such, however, and only certain types of land can become access lands—specifically, mountains, moor, heath, down, and coastal lands.⁶ Moreover, the act includes a wide number of exceptions. For example, lakes and rivers are excluded, as are golf courses and “land habitually used for the training of racehorses.” Activity on access lands is also restricted to walking; most other recreational activities, such as camping, hiking, flower picking, and horseback riding, are excluded.

The English case is representative of other legal jurisdictions, like Germany’s, Switzerland’s, or Denmark’s. In these legal systems, the right to roam exists only in heavily circumscribed forms and does not cover much more than a right to walk across specified types of land. Germany, for example, recognizes no general right to roam but grants a right to freely access forests, as does Austria.⁷ Switzerland gives everyone a right to access forests and meadows, as well as uncultivable land such as the high alpine mountains.⁸

5 The labels for the two models are not meant to suggest precise geographical boundaries: Denmark’s situation is more akin to the northern European model, while Finland and, to a degree, Scotland adopt the Scandinavian model.

6 Klick and Parchomovsky, “The Value of the Right to Exclude,” 949. Gray gives the number as 12 percent (“Pedestrian Democracy and the Geography of Hope,” 48).

7 Gesetz zur Erhaltung des Waldes und zur Förderung der Forstwirtschaft (Bundeswaldgesetz, BWaldG), May 2, 1975, Bundesgesetzblatt [BGBl] I at 1037, as amended, art. 14 (Germany), <https://www.gesetze-im-internet.de/bwaldg/BjNR010370975.html>; Bundesgesetz vom 3. Juli 1975, mit dem das Forstwesen geregelt wird (Forstgesetz, ForstG), July 3, 1975, Bundesgesetzblatt [BGBl] No. 440/1975, as amended, art. 33 (Austria), <https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10010371>.

8 Schweizerisches Zivilgesetzbuch (ZGB), December 10, 1907, SR 210, as amended, art. 699 (Switzerland), https://www.fedlex.admin.ch/eli/cc/24/233_245_233/de#art_699; Bundesgesetz über den Wald (Waldgesetz, WaG), October 4, 1991, SR 921.0, as amended, art. 14 (Switzerland), https://www.fedlex.admin.ch/eli/cc/1992/2521_2521_2521/de#art_14.

For an overview of different European jurisdictions, see Robertson, “Public Access to Private Land for Walking.” Even within the United States, feeble cousins of a right to roam exist, such as laws regulating access to beaches in certain states (Rose, “The Comedy of the Commons”). See also Sawers, “The Right to Exclude from Unimproved Land.”

In the Scandinavian model, by contrast, the right to roam is much broader and far ranging. The Swedish *Allemansrätten*, for example, allows one to camp on private land and to build campfires; to pick berries, flowers, and mushrooms; and to use alternate modes of transportation like skiing and horseback riding. Moreover, the *Allemansrätten* is often seen as a common right, predating the legal system. Thus, most of the rules governing the *Allemansrätten* are not actually found in positive law, but rather at the level of social convention. Consequently, guidance on the *Allemansrätten* tends to be framed vaguely, in terms of open-ended guidelines, and often appeals to common sense and respect for nature and others.⁹

The right to roam must be distinguished from at least one related legal phenomenon. *Easements* grant someone a right to traverse someone else's land for a specified purpose.¹⁰ In the paradigm case, A's only means of accessing a public road from their property is via B's land—in *extremis* you can imagine that B's land surrounds A's like a donut.¹¹ In such a case, almost all jurisdictions will grant A the right to traverse B's land to access the public road. But while both the right to roam and easements grant access to others' private land, easements are highly localized and vested in particular people for a particular purpose. For example, if A's land became directly connected to a public road, their right to traverse B's land would cease; nor could A stray freely from the path that is strictly necessary for them to reach public roads.

2. THE INSUFFICIENCY OF INTERESTS IN ROAMING

In the jurisdictions where it exists, the right to roam is designed to strike a balance between the interests of landowners and the interests of roamers. The public is given an opportunity for recreation and leisure, while landowners' economic and privacy interests remain largely protected. If the right to roam wears any justification on its sleeve, it is this balancing of competing interests. For this reason, it is useful to start with considering the interests roaming serves and whether they are strong enough to justify legally recognizing a right to roam.

Jerry Anderson identifies six relevant interests of roamers: (i) transportation, (ii) enjoyment of nature, (iii) physical health, (iv) mental health, (v) connection to history and culture, and (vi) sense of community.¹² On the basis of any of

9 For a brief overview of *Allemansrätten*, see Sténs and Sandström, "Allemansrätten in Sweden."

10 For an overview of the category in English law, see Gray and Gray, *Land Law*.

11 Gray and Gray, *Land Law*, 248. This is an "easement of necessity." For discussion, see Dominiak, "Must Right-Libertarians Embrace Easements by Necessity?"

12 Anderson, "Britain's Right to Roam," 412–17. I have reordered the list for presentational ease.

these interests, or some combination of them, so the hope goes, we can justify a right to roam. Our aim is specifically to establish the freedom to roam as a *right*, not merely as a desirable policy objective. Within a liberal framework, listing a set of interests that roamers have is insufficient, by itself, to establish roaming as a right.

However, we have a ready-made and widely accepted theory of rights at hand that allows us to make the connection: the interest theory of rights.¹³ According to this theory, if an interest satisfies certain criteria—for example, if it is important enough, broadly shared, and feasible to protect through legal means—then it grounds a corresponding right.¹⁴ Note that this is the most sympathetic ground on which we can consider the issue. On alternative theories of rights, such as a status theory of rights, the right to roam is likely to be harder to defend.

The crux of the matter, then, rests on whether the interests in roaming proposed by Anderson and others fulfill the relevant criteria. I will not develop a full account of what these criteria are. However, there are at least three minimal criteria that any plausible liberal version of the interest theory will recognize:

1. Strength criterion: The relevant interests must be morally important enough to justify imposing duties on others.
2. Neutrality criterion: The relevant interests must be compatible with the state remaining neutral between different conceptions of the good life: all reasonable people must be able to agree that these interests are worth protecting and serving through the state.
3. Necessity criterion: If the relevant interests are meant to justify a right to *X*, then *X* must be practically necessary for serving the interests—there can be no alternative institutional scheme that similarly serves the interest but has significantly lesser moral costs.

These three commitments will become clearer as we go along. My central claim is that none of Anderson's six interests, and no combination of them, fulfills these three criteria.

13 The classic account is Raz, "On the Nature of Rights." A good discussion of Raz's theory, specifically with a view toward property rights, is Cruft, "Against Individualistic Justifications of Property Rights." For an interest theory of human rights, see Tasioulas, "On the Foundations of Human Rights."

14 This means that I use the interest theory as a theory of the justification of rights (as, e.g., Tasioulas does), not as a theory about the nature (e.g., Kramer, "Some Doubts about Alternatives to the Interest Theory of Rights") or directionality (e.g., Sreenivasan, "Duties and Their Direction") of rights.

2.1. *Means of Transportation*

We can start with Anderson's first proposed interest in a right to roam, roaming as a means of transportation. There is an immediate worry stemming from this interest not being strong enough in the average member of the public. The right to roam, after all, primarily applies within a rural context, and it generally does not cover motorized transport. But it is hard to think that there are many people who regularly need to roam across private, rural lands to move between, say, home and work, at least if we think of rich, urbanized societies.

Thus, from the start, this interest is not strong enough to justify a right to roam (although it might justify, say, a right to a public system of roads). The interest theory certainly cannot allow for every interest, no matter how weak, to translate into a right. After all, rights correspond to duties (or other types of Hohfeldian incidents), which demand fulfillment above any other practical concerns. Only important and widely shared interests can serve as a ground for a right, if we want to avoid extreme proliferation of rights.¹⁵

One might reply that some people and some groups have a strong interest to roam as a means of transportation, even if that interest is not widely shared.¹⁶ We can imagine a rural town in which the only plausible way to get to the market or to the factory or to church is by walking through the holdings of a large landowner. Would this not be a sufficiently strong interest to justify a right to roam? The problem with this response is that we now run afoul of the necessity requirement. We can acknowledge that specific groups of people need specific kinds of access to specific lands to fulfill important needs. But surely, granting a general right to roam—that is, a right for everyone to roam any kind of private land—would be a huge overreaction to such localized needs. This argument at best justifies specific easements, not a right to roam.

The idea behind the necessity requirement is that rights should be tailored to the interests they serve. They should cover all that is necessary for protecting and promoting the interests, but not more—especially if there are feasible alternatives that are more tailored for the relevant purpose. Once again, we want to be conservative, both philosophically and practically, in multiplying rights. Rights take primacy in our moral reasoning, imposing side constraints on how we can act and (in the case of positive rights) duties of provision. Given these considerations, it is morally preferable to deal with transportation-based interests in roaming by allowing easements, rather than by granting a general right to roam.

15 On the worry of rights proliferation, see Wellman, *The Proliferation of Rights*.

16 For an example of a group with strong interests in being able to roam, think of the Irish Travellers. We can also think of the rural populations in developing countries.

2.2. *Enjoyment of Nature*

We can see the necessity criterion in action, in a different way, when it comes to roaming as a means of enjoying nature. While not everyone feels the urge to explore nature, walking, biking, and camping in nature are an important element of welfare for many. Let us accept, perhaps generously, that this interest satisfies the strength criterion.

However, consider two alternatives to the right to roam. The first institutional arrangement is a system of public parks and other publicly provided spaces. Under this arrangement, the state provides for an abundance of freely accessible spaces, both in nature and in cities. The state might fund natural parks—for example, Shenandoah National Park in Virginia. Alternatively, some land might be set aside as “commons,” such as Port Meadow in Oxford. On well-administered versions of such schemes, there would be many diverse and geographically dispersed sites, all easily accessible to the public.

We can also imagine a different arrangement that rests on public-private partnerships. Perhaps the state offers a less extensive array of public parks, but it encourages landowners to open up their land to the public. The state might offer small financial rewards to landowners for registering their land as being accessible to roamers, or it might make other legal rules that encourage landowners to open up their lands—for example, through favorable tax treatment. Information on which private lands are publicly accessible would be made available in some central place, such as a government website.¹⁷

Contrast these two schemes with the right to roam, through the specific lens of how they serve an interest in enjoying nature. The following strikes me as true: (i) these two alternative schemes, if implemented reasonably well, will sufficiently serve our interest in enjoying nature, and (ii) these schemes have lesser moral costs than an institutionalized right to roam, because they rest on public or consensual provision of access to land while respecting the autonomy of property owners. These two claims, taken together, mean that enjoyment of nature fails the necessity criterion for establishing a right to roam.

One might respond that a right to roam better serves the relevant interest than these alternative schemes. You might reply, for example, that public parks are harder and more difficult to access, while roaming can often start right from your doorstep. But the relevant question is not whether these alternative schemes *best* serve the interest, but whether they serve it *reasonably well*. A theory that gives us a right to those means that best serve our interests rather than a right to plausible means would be much too demanding, as it would fail to take

17 For a similar proposal, see Kochan, “The Market to Roam.”

into account the burdens of such an institutional scheme. But while the alternative schemes I suggested might have disadvantages compared to roaming, it also seems clear that they serve the relevant interests reasonably well.

2.3. *Physical Health*

The problems with roaming as a means for physical health are closely similar, and we can keep discussion brief. There are well-recognized health benefits of walking and of the other activities a right to roam allows, such as skiing, hiking, and biking, so we can grant that this interest fulfills the strength criterion.¹⁸ However, it is not clear why roamers need access to private property specifically to fulfill this interest. Alternative schemes would be equally sufficient to fulfill it and would have significantly lower moral costs—for example, public subsidies for parks, trails, gyms, sports clubs, or swimming pools. Of course, these alternative schemes might not serve our interest in physical health to the same degree, or in precisely the same way, as a right to roam does, but that is no principled objection to these alternative schemes.

2.4. *Mental Health*

Anderson's next claim is that roaming nature is good for our mental health.¹⁹ "Mental health" here can be interpreted in a narrow and a broad way. In the narrow, medicalized sense, the argument would be that roaming is a good way to prevent, alleviate, or treat diagnosed mental illnesses. It is questionable, however, whether there is much evidence that roaming is good for mental health in this narrow sense (do long walks substitute for a good therapist?). Thus, for the argument to succeed, mental health must be interpreted in a broader sense, to include a wider set of psychological benefits that might come from roaming, such as peace of mind, mental tranquility, a sense of personal freedom, feeling oneness with nature, or a patriotic connection to one's *Heimat* ("homeland").²⁰

There are troubles here, once again, with the strength and necessity criteria. Instead of pursuing these lines of objection, I want to highlight a different problem. Consider that some people have a strong, unquenchable desire to know

18 E.g., Bowler et al., "A Systematic Review of Evidence for the Added Benefits to Health of Exposure to Natural Environments." See Ilgunas, *This Land Is Our Land*, for more references to the health benefits of walking.

19 Ilgunas also stresses the physical and mental health benefits of roaming (*This Land Is Our Land*, chap. 6).

20 Gray, for example, marshaling a variety of poets in support, speaks loftily of nature's ability to support "self-discovery, the quest for meaning and significance, about locating oneself in time and place amidst what is often otherwise a rootless existence" (Gray, "Pedestrian Democracy and the Geography of Hope," 56).

what is going on in other people's lives. In this descriptive sense, nosiness can be a quite psychologically strong desire. But we do not think that such an interest is morally weighty: interests in being nosy have no claim to be enshrined as moral or legal rights. The general point is that pure psychological strength is no indicator for whether some interest should be recognized as the ground for a right, as some of these interests fail important moral constraints.

One such constraint is the neutrality criterion. If the state were to recognize a right to roam on the basis of roaming's supposed tranquilizing effect, for example, then our political institutions would ascribe to a specific account of the nature of the good life. Not everyone will agree that tranquility is a desirable, or particularly valuable, state of mind; or even if everyone agrees with that claim, many people will reasonably deny that tranquility is closely connected to roaming nature or that its importance is so great as to trump the autonomy of landowners. Thus, if the right to roam were justified on any such grounds, it would be justified not on public grounds that are accessible to everyone but on a specific vision of what a good life looks like.

I think that all attempts to cash out what we mean by "mental health" in the broad sense will fall prey to this problem. There will be at least some reasonable visions of the good life on which there is no importantly strong, mental health-based interest in roaming the countryside. Insofar as we aim to develop a recognizably liberal justification of the right to roam, then, this path is barred.

2.5. Connection to History and Culture

Let us move on to the second-to-last interest mentioned by Anderson, roaming as a way to connect with one's history and culture. There certainly is value in having access to sites that play an important part in one's culture or history: there is value in the English seeing Stonehenge, the Greeks visiting the Acropolis, or Indians accessing the Ganges. Similarly, there are some sites that stand out for their stunning and unique natural beauty, like the Grand Canyon. For places like these, we have a strong interest in being able to access them. Furthermore, the interest also pertains to these lands specifically, as nothing else can serve as a substitute.

However, this interest once again fails the necessity criterion. What this observation establishes is a strong presumption, perhaps a requirement, that sites of cultural and historical significance or outstanding natural beauty should be accessible to everyone. Many legal systems are based on an implicit understanding that such lands should serve the public, or that if they are privately owned, there should be heavy restrictions on what can be done to the land. The right to roam, however, applies to even the most mundane of land—the kind of place

that has never appeared on a postcard and where people put up sarcastic signs, such as “Nothing happened here on November 7, 1839.”

Another way to gather support for roaming is through the vague concept of *Heimat*, which expresses the idea of the deep social, emotional, and cultural connection to the identity-forming surroundings in which one grew up.²¹ On this view, not being able to wander, say, the village of one’s youth or the forests and meadows of one’s community would be a crucial injury to one’s ability to make a place of one’s own in the world. The advantage of this argument is that it plausibly succeeds on the necessity criterion. You need access to these specific lands—say, the English countryside—and access to only some of it will not do, either. However, the problem is that it fails the neutrality criterion. Aside from the vagueness of the idea, it rests on deeply contestable ideas about national and cultural ties to land, imbuing them with a mythical aura that we cannot appeal to in liberal argument.

2.6. *Sense of Community*

Lastly, sometimes the idea is advanced that roaming supports communal bonds. Ken Ilgunas, on the basis of his extensive experiences hitchhiking across the United States, argues that through roaming,

you learn that the American people are not nearly as scary and dangerous as our news stories would have us believe. You come to realize that Americans generally are hospitable, friendly, and kind. Once this realization sets in, you begin to move through the world in a lighter, more carefree way. You feel more trusting, more relaxed, more free.²²

Roaming, Ilgunas argues, increases hospitality and social trust. It is a way to meet strangers and neighbors, to form social bonds, and to overcome the atomism that characterizes modern life. I leave it as an exercise to the reader to consider how this argument fails each of the three criteria so far mentioned; it seems clear to me that it does.

3. LEFT-LIBERAL ACCOUNTS OF PROPERTY

I have offhandedly mentioned the moral costs of the right to roam. It is now time to describe what these costs are. I will do so by outlining a left-liberal account of private property on which the right to roam looks *prima facie* suspicious. To avoid being dragged down by detail, I will give a generic account, not the view

21 For the notion, see Blickle, *Heimat*.

22 Ilgunas, *This Land Is Our Land*, chap. 6.

advanced by any particular left-liberal thinker. However, I am confident that my account fairly represents mainstream left-liberal thinking about private property in political philosophy.

At its core, the left-liberal view is committed to a combination of two normative principles:

Principle of Autonomy: Our fundamental concern is to distribute opportunities to exercise one's autonomy—the ability to freely create, pursue, and realize one's own life plans. What counts as a good life is to be decided by each individual, and political institutions should not take a stance on these issues.

Principle of Distribution: Opportunities to realize one's life plans should be distributed in an equitable way (e.g., equally or in a way that benefits the worst-off or in a way that guarantees everyone a decent minimum), and political institutions should ensure that they are distributed in such a way.²³

It is now common to think of rights to private property as a bundle of rights, such that we can analytically distinguish, for example, a right to transfer property from a right to benefit from it.²⁴ In particular, we can distinguish between rights to transfer, rights to reap benefits from, and rights to control property. The right to roam would correspond to a diminished set of landowners' rights to control their property, while it would leave other rights pertaining to their land intact. The crucial question, then, is what bundle of control-rights owners would have under the left-liberal view.

The left-liberal cannot think of the standard bundle of private property rights as including *all* logically possible incidents. Specifically, left-liberals will need to say that private property rights do not include rights to reap the (full) income from what one owns or the right to transfer one's property as one pleases, as achieving the aims set by the principle of distribution would otherwise become impossible. The left-liberal state needs to be able to raise taxes and, more generally, to take from some to give to others. John Christman has argued this point in

23 In this characterization, I am indebted to a similar description in Fabre, "Justice, Fairness, and World Ownership."

24 Some commentators insist that the bundle-of-rights view is uninformative or wrong (e.g., Penner, "The Bundle of Rights Picture of Property"). These "new essentialists" insist that a right to exclude forms part of the core of the concept of private property. (For a good overview, see Wyman, "The New Essentialism in Property.") I will set this debate aside, however.

The Myth of Property, in which he rejects the “absolutist” view of private ownership on which it is a kind of absolute sovereignty over what one owns.²⁵

One might think that a rejection of the absolutist view of ownership puts the defender of the right to roam in a good position. If transfer and benefit-reaping rights can be limited in the service of distributional aims, it seems that we can be optimistic that control rights are similarly limited. However, most left-liberals have suggested that control rights should be treated differently.²⁶ As an example, we can turn to Dworkin’s luck egalitarianism. Dworkin famously develops his theory of distributive justice through a hypothetical auction, the aim of which is to establish equality of resources. Dworkin notes, however, that such an auction is indeterminate unless bidders know what they will be able to do with their resources—that is, what rights attach to the resources they win at the auction.²⁷

Dworkin explicitly rejects a picture on which various rights on resources are auctioned off alongside resources (such that, for example, you might bid on a tree *and* the exclusive right to pick its fruit).²⁸ Instead, he suggests that the auctioneer must divide up resources and accompany them with a fixed bundle of rights, in a way that allows for “the greatest possible flexibility in fine-tuning bids to plans and preferences,” before the auction starts.²⁹ Bidders must be allowed to bid for a profile of resources that best allows them to achieve their life plans (we can see the principle of autonomy in action here). This, Dworkin claims, is best achieved if the resources people bid for are accompanied by very strong control rights: “The principle of abstraction insists that people should in principle be left free ... to use the resources they acquire ... in whatever way they wish.”³⁰ Thus, once you acquire a resource in the Dworkinian auction, you have broad control rights over what ought to be done with it, and you do not need to specially bid for such rights.

In short, it is highly likely that rights to exclude will be a constitutive feature of property rights in a Dworkinian system. Thus, Dworkin is in agreement with right-liberal philosophers who think that the right to exclude is part of the core of ownership and that property rights are important because they allow owners

25 Christman calls it the “liberal” view, but I think this label is misleading. See also Murphy and Nagel, *The Myth of Ownership*.

26 Including Christman, *The Myth of Property*.

27 Dworkin, *Sovereign Virtue*, 143.

28 Dworkin, *Sovereign Virtue*, 139–45.

29 Dworkin, *Sovereign Virtue*, 151.

30 Dworkin, *Sovereign Virtue*, 152. For a good discussion of Dworkin’s principle of abstraction, see Otsuka, “Liberty, Equality, Envy, and Abstraction.”

to “set an agenda” for the things owned.³¹ Dworkin’s disagreement with right-liberal and libertarian thinkers is primarily over how property rights should be distributed, not over the thing that should be distributed.

Even left-liberals who do not share the specifics of Dworkin’s position will be strongly drawn to the outcome he reaches due to the combination of the two principles that characterize left-liberalism. Left-liberals are concerned with giving people equitable chances of leading autonomous lives. This suggests that we should assign to people that set of property rights that best allows them to realize their autonomy, subject to the constraint that everyone have equal (or otherwise justly distributed) chances of doing so. Exclusive rights over a certain part of the world seem to be an important ingredient in doing so. The right to roam takes away from landowners the ability to exclusively determine the agenda for the things they own. For this purpose, it does not matter that roaming leaves one’s property and privacy largely untouched. It is the reduction in one’s ability to control what one owns that is the moral insult and that is grave enough to justify exclusion. On the left-liberal picture of property, it is thus likely that we will oppose any such diminishment of control rights.

4. A NONIDEAL ARGUMENT

Putting the arguments from the previous two sections together, we can see that our chances of defending the right to roam on left-liberal grounds look bleak: there are significant moral costs to recognizing a right to roam and few, if any, convincing benefits that pass the test of the interest theory. I will start in this section by emphasizing the role that a right to roam can play in a nonideal world—that is, a world that is far removed from what is demanded by the principle of distribution.

The distribution of resources and, more broadly, opportunities to realize one’s autonomy is enormously uneven. This is particularly true of land. In the face of these violations of the principle of distribution, we should look for feasible schemes that bring the distribution of resources and opportunities closer to the distributive ideal. The obvious way of doing so would be a large-scale redistribution of resources, or some radical change in our economic and social institutions. However, there are deep and robust political and social obstacles to such ambitious attempts to equalize resources. In the medium term, we must look for more modest, feasible policy options, even if those might not be part of the ideal package of policies that we should adopt.

It is on this ground that we can offer a first attempt to defend the right to

31 For these ideas, see Schmidtz, “Property and Justice,” 80; and Katz, “Exclusion and Exclusion in Property Law.”

roam. The right to roam provides roamers—generally speaking, the have-nots—with increased opportunities to exercise their autonomy, while it takes a small smidgeon of that ability away from owners who already have ample opportunities, and much more opportunity than justice demands, to realize their autonomy. Instituting the right to roam, then, is a move toward a more equitable distribution of opportunities for autonomy under nonideal circumstances.

Encouraging evidence for the feasibility of realizing a right to roam comes from the English and Scottish cases, which implemented such schemes through statutory law relatively recently. Both nations had center-left governments who found the political capital to implement such reforms, with the Scottish law being particularly ambitious. By contrast, the alternative schemes I have discussed previously—for example, a system of public parks—would require much more time, effort, and money to realize, as they would necessitate prolonged negotiations and expensive buybacks from private owners.³²

This first attempt at a nonideal argument gets the basics of a convincing justification of the right to roam correct: it must matter that we live in a world characterized by extreme distributive injustice. But the argument, as offered, is too simple. A first, general problem is that instituting a right to roam is a minor way of remedying inequality. The poor have a claim, at the bar of justice, to have their life chances improved; the freedom to traverse the lands of others seems a minor, perhaps even insultingly miniscule, moral response to that claim. One might reply that the right to roam would be part of a package of policies tackling inequality, but even then, it remains mysterious why we should put so much emphasis on it and why our political energies should not be directed elsewhere.

A second observation is that the current argument deals in two different kinds of autonomy that are hardly of the same nature and value. The autonomy that owners possess under robust right-to-exclude schemes makes them akin to small-scale tyrants over a part of the world. Under a roaming scheme, this form of autonomy is lost. But the autonomy that roamers gain in return is at best a pale cousin of this robust autonomy. A right to roam, even in its Scandinavian form, allows only a limited range of activities: it does not allow residency, commercial activity, altering the land, or determining how the land ought to be used. Roaming is certainly one form of freedom, but it seems insignificant in comparison to the autonomy of full ownership. Thus, we might worry that the current argument has a leveling-down feel to it: we take away from some (the owners) a highly valued form of autonomy, but in return no one gets to enjoy it.³³

32 Feasibility considerations of this nature are stressed in Ilgunas, *This Land Is Our Land*, chap. 6.

33 For a famous statement of the leveling-down objection in a different context, see Parfit, “Equality and Priority.”

5. THE SYMBOLIC ARGUMENT

Despite the weaknesses in this first attempt, we are not far from a successful argument. The core idea missing is that the right to roam serves as a *symbolic* inscription of a fundamental form of social equality, namely the fundamental co-ownership of all land. I will start with the first idea, of something being a symbolic inscription, before I turn to the second, more complicated notion.³⁴

5.1. *Symbolic Inscriptions*

Let us start with an analogy. North Korea is one of the worst regimes you can imagine, and no liberal country should support such a regime politically. At the same time, North Korea has faced, and is likely to face again, humanitarian crises, especially with respect to its ability to feed its own population. What might be called for, in such a case, is a form of reluctant humanitarian aid—aid that addresses the humanitarian problem and helps the North Korean people while also expressing the serious reservations about the regime to, and through, which such aid is given.

In particular, imagine that the European Union provides food, seeds, tractors, and other machinery to North Korea as a humanitarian gift, but that it inscribes them with the slogans “Provided by the European Union” and “Freedom—Democracy—Human Rights” (in Korean). The inscriptions are symbolic: first, they express, opposing North Korean propaganda, that North Korea is not self-sufficient and that it cannot claim this help for itself; and second, they express that the aid is a reluctant gift and not an endorsement of the oppressive policies of the North Korean state. It need not detain us how realistic this example is. What is important are two features of the symbolic inscriptions. First, they do not diminish the essential usability of the aid given, in its material and economic terms: the tractors still work, the seeds still plant, the food still nourishes. What the inscriptions do, however, is to provide the aid with a certain expressive meaning: they establish certain social meanings and aim to prevent others.

Similarly, I would suggest, we can see the freedom to roam as a symbolic inscription in the rights given to property owners. The inability of landowners to keep out roamers serves as a symbolic reminder that all land is, under a left-liberal scheme, fundamentally co-owned and that landowners are in some way socially equal to nonowners (more on that idea in a moment). However, this does not diminish the core usability of private property rights in land. Those who privately own land can still decide whether it shall be field or golf course, shed or skyscraper, garden or wasteland; they can reap profits from the land, transfer it,

34 On symbolic value, see Adams, “Symbolic Value”; and Sneddon, “Symbolic Value.”

abandon it, and so forth. They can, in other words, determine the agenda for the land they own in almost any respect other than whom to keep out.

The right to roam is different, of course, from a physical inscription or object that we classically think of as symbols. As a legal rule, it is a complex social convention that expresses itself, in particular, in the behavior of public officials.³⁵ This does not diminish its ability to bear symbolic meaning, however. What is needed for something to be legible as a symbol is some community of interpreters who roughly converge in their standards of interpretation. Legal rules are legible, in principle, in such a way, and we have no problem assigning symbolic meaning to legal rules.³⁶

5.2. *Fundamental Co-ownership*

Second, we need to work out what the right to roam symbolizes. The obvious place to look is the left-liberal principles of justice, and in particular, the principle of distribution. Every distribution of benefits and burdens, according to that principle, needs to be justified against some distributive (patterned) ideal. This entails that no external piece of property is *fundamentally* owned: the possession of any piece of land is only justified insofar as it is compatible with equal opportunities for all.

Imagine that *A* possesses some tract of land *L*. Let us imagine a simple non-luck-egalitarian account of distributive justice on which everyone is entitled to an equally valued part of the available resources in the external world.³⁷ If so, *A* is justified in holding *L* just in case, and because, *L* is part of a package of resources that has equal value to what everyone else holds. The point is that all land, like all other resources, is fundamentally co-owned, with no land being exempt from distributive claims. There is no antecedent, fundamentally private ownership of land—no procedure of acquisition, as Locke or Nozick recognize it, to make unowned land one's own. Any claim that *A* has to *L* is the upshot of an empirical, contingent distribution of legal rights to realize moral claims on a more fundamental level, on which no land is owned.

Fundamental co-ownership in this sense is a *regulative* idea, not one that expresses a historical or legal reality. In particular, the claim is not that under any current legal structure, everyone can make a claim to own all parts of the world

35 The background assumption here is a form of legal positivism à la Hart, *The Concept of Law*.

36 I return to some of these issues in section 6.

37 More sophisticated egalitarian theories are likely to make the morally optimal shares of land choice sensitive, such that *A* can possess land of higher value than *B* if that difference goes back to voluntary choices that *A* and *B* have made. This would not change the basic point made in the text, however.

in combination with all others. Whether any individual possesses legal rights in common with others depends on the particulars of an individual society's legal conventions. In its regulative function, however, fundamental co-ownership expresses an important truth about the justificatory bar against which any legally instituted system of ownership must be measured. Even if *A* might legally possess *L*, *A* must offer reasons other than initial acquisition or just transfer why he should be recognized as legally owning it. Instead, *A* must show that the system by which we distribute the totality of external resources serves certain distributive aims.

Fundamental co-ownership in this sense expresses a form of fundamental social equality. We are not just formally equal, in that moral principles cover us equally and that there are no natural hierarchies between us. We are also fundamentally and substantively equal, in that each of our lives bears equal weight, and all differences between us must be justified to others at the bar of the liberal principles of justice. We can keep the idea of social equality vague, as a multitude of interpretations of that idea can be offered; the same goes, to a degree, for the idea of fundamental co-ownership.³⁸ Different left-liberal conceptions will fill out these ideas in different ways, corresponding to different ways to understand the ideals of autonomy and distributive justice. But any such view, I presume, will be committed to some account of fundamental social equality with respect to external resources.

5.3. *The Symbolic Value of Roaming*

Let us connect the dots. We live in societies that do not realize, and are not remotely close to realizing, a just distribution of resources and opportunities. In such a society, limiting private property owners' control right to exclude by giving the public a right to roam is a symbolic reminder that any rights granted under our legal system are ultimately only justified against a background of social equality. Roaming does not take away owners' land or the essential usability of that land, but it reminds owners and nonowners alike of the fundamental moral claims that others have against them. It is the egalitarian thorn in a legal system that otherwise encourages hierarchical relations between haves and have-nots.

Why is there value in reminding us of fundamental social equality in our property rights? There are two parts to this question: first, why there is value of reminding us of fundamental social equality at all in our rights, and second, why there is value in doing so with respect to control rights concerning land in particular. The answer to the first part goes back to the fact that we live in a nonideal

38 For a recent account of co-ownership of the world, see, e.g., Risse, "Common Ownership of the Earth as a Non-Parochial Standpoint."

world. A system of strong private property ownership coupled with enormous inequalities in ownership encourages an absolutist, quasi-libertarian conception of ownership—a picture on which you absolutely own what is yours and owe only slender responsibilities to the moral community.

However, there is value in reminding ourselves that any deviations from the ideal of distribution are accepted only grudgingly and temporarily until we can achieve a better distribution. The error of the Dworkinian left-liberal is precisely to think that a change in the principle of distribution leaves the things to be distributed unchanged. Symbolic reminders leave the essential usability of property intact, thus granting owners a still significant amount of autonomy, while also highlighting the fundamentally egalitarian basis on which any claim to property is justified.

The second part, then, is why we should symbolically remind ourselves of social equality through control rights concerning land in particular. After all, there are many different ways in which we could institute symbolic reminders. There is no deep, *a priori* answer to be offered in response, only a series of observations based on the common workings of human society. First, insofar as we all need space to live and realize the many valuable activities of human life, land appears as a finite and scarce precondition for achieving valuable lives. Second, most societies tie economic wealth, security, and social recognition closely to land ownership. Inequality in land ownership—and our inability to access private land in particular—is one of the most visible ways in which wider social and economic inequalities express themselves in many societies. All this makes land a focal resource worth regulating for the left-liberal.

6. OBSERVATIONS AND OBJECTIONS

I will now consider some advantages of this argument, as well as respond to some objections.

6.1. *Interests, Again*

I argued that the interest-based argument failed because the proposed interests grounding a right to roam were not important enough, were inaccessible within public reason, or failed the necessity requirement. The revised nonideal argument has no similar weaknesses. First, a right to roam is *not* directly justified on the basis of anyone's interests in roaming. Instead, it is justified with reference to a fundamental liberal principle of justice. These principles do not commit us to any particular conception of the good life, and thus pass the neutrality requirement. In the same way, we can answer the worry that the right to

roam does not rest on moral considerations that are weighty enough. On the argument considered, the right to roam is a symbolic reminder of an important liberal principle—namely, the fundamental co-ownership of the world. Lastly, the argument answers why a universal right to roam, rather than, say, specific easements or public-private partnerships, is necessary. None of these alternative schemes symbolically denotes *all* land as being fundamentally co-owned; indeed, public-private partnerships reinforce the notion that the sovereignty of owners is in some way supreme.

Still, interests play some role in the current argument. There are different ways that the fundamental co-ownership of the world could be symbolically represented in our legal norms and conventions, some of which serve our interests more than others. Another way we could symbolically represent social equality, for example, is to give everyone ownership of an equal plot of land on the moon.³⁹ We can assume for argument's sake that moon co-ownership works just as well symbolically as a right to freely roam the earth. Of course, no one's tangible interests are served by having a plot of land on the moon. That the right to roam not only symbolically represents social equality but also serves important interests thus gives us reason to prefer it over moon co-ownership. That roaming serves interests thus enhances the argument for it, without being its primary foundation.

6.2. *Autonomy, Again*

We can also review the symbolic argument in light of the failures of the simple nonideal argument (section 4). The first objection was that roaming seems like a minor and insignificant remedy for the pressing inequalities we face. The advantage of the current argument is that it does not justify roaming as a serious attempt to remedy distributional equality through redistribution. Rather, roaming is offered as important in its symbolic capacity. Thus, even if the right to roam did not shift burdens and benefits in a meaningful way toward equality, we could still justify it on the basis of its symbolic capacity.

What about the suspicion that roaming leads to a repugnant form of leveling-down, by making it impossible for anyone to realize the superior form of autonomy that comes with full control rights over land? Here, I think, the left-liberal defender of the symbolic argument should, to a degree, bite the bullet. Under a roaming scheme, you cannot unilaterally say no to others entering your property. (Note, however, the many exceptions: you can still say no to others invading your privacy or diminishing the economic viability of your land, etc.) Thus, we all lack a certain kind of strong autonomy over our own property.

39 I owe this example to an anonymous reviewer.

Lacking complete autonomy in this sense is precisely the symbolic reminder of our falling short of an equitable distribution of resources.

6.3. *Further Doubts about Symbolism*

We need to be careful in distinguishing the symbolic meaning of some legal rule (i.e., how a competent interpreter in a given community would understand it) from its expected impact on social conventions (i.e., how that legal rule being in force is likely to impact the behavior and beliefs of individuals subject to that rule).⁴⁰ I have suggested that a right to roam would be an important symbol. However, one might now raise doubts of two kinds: (i) whether the symbol would be recognized in its intended role, and (ii) whether, even if it were recognized in its intended role, it would have any significant effects on changing social conventions and behavior.

With respect to the first doubt, we should admit that the mere existence of a right to roam, as a legal rule, is probably insufficient for that right to be interpreted as a symbol for social equality. There could be societies that allow roaming but do not symbolically interpret it to express any kind of social equality. To be successful, the right to roam requires a larger social, cultural, and political context that makes the rule symbolically legible in the right way.

We can even imagine scenarios in which the right to roam is perceived opposite from the way it is intended, in a fashion detrimental to left-liberalism. Perhaps the right is interpreted, for example, as a symbol of disrespect for autonomy, or in favor of state paternalism, or as some other form of overbearing state interference. If this perception is widespread and does not rest on willful ignorance or misinterpretation of the policy, then it speaks against adopting the right to roam.⁴¹

Inversely, we can imagine that anti-egalitarian arrangements are interpreted in an egalitarian fashion. Imagine that a private company starts buying up large tracts of land, finally becoming a monopoly in a certain area, and calls itself “The People’s Land.” The company advertises itself (misleadingly) as not being interested in profit and uses the slogan, “The People’s Land Belongs to Everyone.” We can imagine that such advertising contributes to a perception that its land is indeed owned by its leaseholders. Some observers might mistake “The People’s Land” for a public service and falsely believe that it has egalitarian aims it

40 The point is stressed in Adler’s critique of expressive theories of law (“Expressive Theories of Law”).

41 For this paragraph and the next two, I am indebted to very insightful objections from an anonymous reviewer. The idea of “The People’s Land,” in particular, is adapted from their comments.

does not in fact have. We can assume that residents on “The People’s Land” will perceive each other to be highly socially equal to each other, even though there is an extreme difference in wealth between them and the owners of the holding company.

If symbolism possesses any independent weight, then it seems that left-liberals should at least sometimes endorse a social institution like “The People’s Land,” even though it is detrimental to substantive egalitarian aims. This is a counterintuitive result. I think our answer must be that the value of symbols is subject to some internal coherence requirement, such that symbols must be compatible, in some way, with the things they mean to symbolize. On this view, some symbol *X* representing a thing *Y* possesses independent value only if *Y* is intrinsically valuable and *X* itself is compatible with *Y*. On this view, there would be no independent value in “The People’s Land,” because the operation of this institution contradicts the value it is meant to symbolize. The right to roam, on the other hand, does not obviously violate the coherence requirement. Thus, there is independent value in legally recognizing such a right.⁴²

Putting such scenarios aside, I am confident that we can find the cultural resources, at least within Western legal cultures (which I know best), that allow the right to roam to be interpreted in an egalitarian fashion. The Land Reform (Scotland) Act 2003, for example, was certainly perceived as a model to equalize access to land, and it built on (among other things) an important historical tradition in which the Scottish Highlands were perceived to be owned by the Scottish people in common.⁴³ Moreover, even if such resources for interpretation do not exist in a given society, they can be created—first and foremost, politically. Instituting a right to roam would then be part of a larger campaign of reforming the perception of property rights.

We should also admit that a legal rule like the right to roam has neither a single nor an unambiguous symbolic meaning. Instead of the technical notion of fundamental co-ownership, a right to roam is likely to be understood also in vaguer terms of equality, democracy, shared heritage, or community. What matters, however, are not the precise labels that would be used in interpretation but whether a legal right favoring roamers overall sends a message that endorses, or is at least compatible with, the principle of fundamental co-ownership.

Another worry is the second—namely, even if roaming is interpreted in its intended fashion, it remains a powerless symbol. What would roaming matter,

42 More would need to be said about the relevant notion of “compatibility,” and there is a host of further questions that we would need to tackle. But I will set these issues aside.

43 For a highly sympathetic treatment of the Scottish case along these lines, see Alexander, “The Sporting Life.” Cf. Lovett, “Progressive Property in Action.”

this objection goes, if it did not incline anyone to act or think in more egalitarian ways? The reply to this objection, I think, should be twofold. First, changing behavior is not the primary object of a symbolic reminder. Think back to the humanitarian aid given to North Korea by a reluctant European Union. It would certainly be a problem if the symbolic inscription were not recognized at all. At the same time, there is no expectation of success in significantly changing the North Korean government or its policies. Similarly, if the right to roam can be recognized as a slight to quasi-libertarian perceptions of property rights, and sometimes it does get recognized in such a way, this might be sufficient to fulfill its purpose.

However, if the right to roam really turned out to be a fundamentally powerless symbol, which engaged the members of a given society so marginally as to not significantly change anyone's behavior or perception, then we should indeed rethink its aptness. Where the social and cultural context is such that reminders of the fundamental co-ownership of the world are best expressed through alternative means, those should be chosen. But while it is a limitation, I think this is not a weakness: it simply shows that the argument is sensitive to concrete social and cultural contexts, as it should be.

7. CONCLUDING REMARKS

In lieu of a summary, let me finish by noting some open questions.

First, I have not defended the left-liberal view of property rights. The argument of this paper is thus one internal to left-liberalism (which, I should add, I am not myself committed to in many respects). Further philosophical work would look at how this argument plays out in the context of different left-liberal accounts of justice. I have mostly operated, for example, with an egalitarian principle of distribution; we might wonder what would change in the argument if we move to a more sufficientarian account.

Second, following much of the literature on the right to roam, I have implicitly conceived of it as applying in a rural context, as the right to roam nature. But we might ask whether there are not also implications in an urban context. For example, we might think that a similar argument can be hoisted against gated communities or the privatization of public amenities.

Lastly, we can also consider whether the argument I have given can be extended to other practical contexts. The most general lesson of the paper is that the left-liberal principle of distribution should also affect, via the inclusion of symbolic reminders, the things to be distributed. If I am correct, then all rights should bear symbolic inscriptions of their egalitarian purpose. This is likely to

have implications for the ownership of other external resources—for example, the means of production—as well as how certain services and opportunities should be offered, such as education. But tracing these implications is the story for another day.⁴⁴

Ludwig-Maximilians-Universität Munich
 mail@matthiasbrinkmann.de

REFERENCES

- Adams, Robert Merrihew. "Symbolic Value." *Midwest Studies in Philosophy* 21, no. 1 (September 1997): 1–15.
- Adler, Matthew. "Expressive Theories of Law: A Skeptical Overview." *University of Pennsylvania Law Review* 148, no. 5 (May 2000): 1363–501.
- Alexander, Gregory S. "The Sporting Life: Democratic Culture and the Historical Origins of the Scottish Right to Roam." *University of Illinois Law Review* 2016, no. 2 (2016): 321–70.
- Anderson, Jerry L. "Britain's Right to Roam: Redefining the Landowner's Bundle of Sticks." *Georgetown International Environmental Law Review* 19, no. 3 (2007): 375–436.
- Blickle, Peter. *Heimat: A Critical Theory of the German Idea of Homeland*. Rochester, NY: Camden House, 2004.
- Bowler, Diana, Lisette Buyung-Ali, Teri Knight, and Andrew Pullin. "A Systematic Review of Evidence for the Added Benefits to Health of Exposure to Natural Environments." *BMC Public Health* 10, no. 456 (2010).
- Christman, John Philip. *The Myth of Property: Toward an Egalitarian Theory of Ownership*. New York: Oxford University Press, 1994.
- Cruft, Rowan. "Against Individualistic Justifications of Property Rights." *Utilitas* 18, no. 2 (June 2006): 154–72.
- Dominiak, Łukasz. "Must Right-Libertarians Embrace Easements by Necessity?" *Diametros* 16, no. 60 (June 2019): 34–51.
- Dworkin, Ronald. *Sovereign Virtue: The Theory and Practice of Equality*. Cambridge, MA: Harvard University Press, 2002.
- Fabre, Cécile. "Justice, Fairness, and World Ownership." *Law and Philosophy* 21, no. 3 (May 2002): 249–73.

44 I am thankful to Amin Ebrahimi and an anonymous reviewer for this journal for helpful comments.

- Gray, Kevin. "Pedestrian Democracy and the Geography of Hope." *Journal of Human Rights and the Environment* 1, no. 1 (March 2010): 45–65.
- Gray, Kevin, and Susan Francis Gray. "Civil Rights, Civil Wrongs and Quasi-Public Space." *European Human Rights Law Review* 4 (1999): 46–102.
- . *Land Law*. 7th ed. Core Text Series. Oxford: Oxford University Press, 2011.
- Hart, H.L.A. *The Concept of Law*. Oxford: Clarendon Press, 1961.
- Ilgunas, Ken. *This Land Is Our Land: How We Lost the Right to Roam and How to Take It Back*. New York: Penguin, 2018.
- Katz, Larissa. "Exclusion and Exclusivity in Property Law." *University of Toronto Law Journal* 58, no. 3 (Summer 2008): 275–315.
- Klick, Jonathan, and Gideon Parchomovsky. "The Value of the Right to Exclude: An Empirical Assessment." *University of Pennsylvania Law Review* 165 (2016): 917–66.
- Kochan, Donald. "The Market to Roam: Using Sharing Economy Platforms for Expanding Roaming Access to Land Resources." *Natural Resources Journal* 59, no. 1 (Winter 2019): 89–134.
- Kramer, Matthew H. "Some Doubts about Alternatives to the Interest Theory of Rights." *Ethics* 123, no. 2 (January 2013): 245–63.
- Lovett, John. "Progressive Property in Action: The Land Reform (Scotland) Act 2003." *Nebraska Law Review* 89, no. 4 (2010): 739–818.
- Murphy, Liam, and Thomas Nagel. *The Myth of Ownership: Taxes and Justice*. Oxford: Oxford University Press, 2002.
- Otsuka, Michael. "Liberty, Equality, Envy, and Abstraction." In *Dworkin and His Critics: With Replies by Dworkin*, edited by Justine Burley, 70–78. *Philosophers and Their Critics* 11. Malden, MA: Blackwell Pub, 2004.
- Parfit, Derek. "Equality and Priority." *Ratio* 10, no. 3 (December 1997): 202–21.
- Penner, J.E. "The Bundle of Rights Picture of Property." *UCLA Law Review* 43, no. 3 (1996): 711–820.
- Raz, Joseph. "On the Nature of Rights." *Mind* 93, no. 370 (April 1984): 194–214.
- Risse, Mathias. "Common Ownership of the Earth as a Non-Parochial Standpoint: A Contingent Derivation of Human Rights." *European Journal of Philosophy* 17, no. 2 (June 2009): 277–304.
- Robertson, Heidi Gorovitz. "Public Access to Private Land for Walking: Environmental and Individual Responsibility as Rationale for Limiting the Right to Exclude." *Georgetown International Environmental Law Review* 23, no. 2 (2011): 211–62.
- Rose, Carol. "The Comedy of the Commons: Custom, Commerce, and Inher-

- ently Public Property." *University of Chicago Law Review* 53, no. 3 (Summer 1986): 711–81.
- Sawers, Brian. "The Right to Exclude from Unimproved Land." *Temple Law Review* 83, no. 3 (Spring 2011): 665–96.
- Schmidtz, David. "Property and Justice." *Social Philosophy and Policy* 27, no. 1 (January 2010): 79–100.
- Shoard, Marion. *A Right to Roam*. Oxford: Oxford University Press, 1999.
- Sneddon, Andrew. "Symbolic Value." *Journal of Value Inquiry* 50, no. 2 (June 2016): 395–413.
- Sreenivasan, Gopal. "Duties and Their Direction." *Ethics* 120, no. 3 (April 2010): 465–94.
- Sténs, Anna, and Camilla Sandström. "Allemansrätten in Sweden: A Resistant Custom." *Landscapes* 15, no. 2 (November 2014): 106–18.
- Tasioulas, John. "On the Foundations of Human Rights." In *Philosophical Foundations of Human Rights*, edited by Rowan Cruft, S. Matthew Liao, and Massimo Renzo, 45–70. Oxford: Oxford University Press, 2015.
- Wellman, Carl. *The Proliferation of Rights: Moral Progress or Empty Rhetoric?* Boulder, CO: Westview Press, 1999.
- Wyman, Katrina M. "The New Essentialism in Property." *Journal of Legal Analysis* 9, no. 2 (Winter 2017): 183–246.

THE INDEFENSIBLE SELF-DEFENSE ARGUMENT

Howard Hewitt

THE SELF-DEFENSE ARGUMENT is an argument for a woman's right to abortion on the grounds that it is an exercise of her right to defend her bodily autonomy. Even if we assume that the fetus is an innocent person, it is still the case that he occupies his mother's body in a particularly intimate way and, in an unwanted pregnancy, without her ongoing consent. These facts require us to concede that abortion is an exercise of a woman's right to defend herself from those who would use her body without her consent, regardless of the moral status of the fetus or any of the other facets of the relationship between a woman and her unwanted fetus. If sound, this argument would show that abortion is morally permissible in any case of unwanted pregnancy prior to the viability of the fetus. The self-defense argument is a robust, uncompromising defense of abortion rights, and while I will argue it fails to establish its conclusion, its failure has something to teach us about the abortion controversy.

I will argue for three claims in what follows: First, that the self-defense argument is unsound. Second, that a crucial assumption motivating the self-defense argument—namely, that we are to understand the mother-fetus relationship as a relationship governed by the same norms that govern intimate bodily relationships between normal adult human beings—not only fails to support a woman's right to abortion but also leads to the opposite conclusion. A commitment to understanding the mother-fetus relationship in this way leads to the conclusion that abortion in typical cases of unwanted pregnancy is morally impermissible, at least on the assumption that the fetus is a person. This leaves us with a choice: accept the moral impermissibility of abortion or reject the commitment. I argue that we all, abortion defender and critic alike, have good reason to reject the commitment, the third claim I aim to establish.

The paper goes like so: In section 1, I explain the self-defense argument and how its premises must be interpreted in order for it to be plausible. In section 2, I provide a counterexample to one of the self-defense argument's premises. The premise is that in any case where another person is using one's body in an intimate way without one's consent, one may use lethal force to stop that use, given

that no other means are available. I show that this is false and that the premise cannot be successfully revised, thereby establishing that the self-defense argument is unsound.

In section 3, I mount an analogy argument from my counterexample to the conclusion that abortion is morally impermissible in typical cases of unwanted pregnancy. I show that what drives us to that conclusion is the self-defense argument's particular commitment to applying the norms governing intimate bodily contact between normal adults to the mother-fetus relationship. This commitment leads to the conclusion that abortion in typical cases of unwanted pregnancy is morally impermissible. I also show that this commitment implies that cases of voluntary conception constitute impermissible violations of a fetus's right to bodily autonomy, and so even though the commitment has implications amenable to the abortion critic, it is not something he can accept.

Section 4 concludes, in light of the foregoing, that we should reject that commitment and the self-defense argument along with it.

1. THE SELF-DEFENSE ARGUMENT

We are committed to a limited right of self-defense against persons using our bodies without our consent, including lethal self-defense in cases where the use is especially serious and there are no nonlethal means available to avoid it. The self-defense argument holds that abortion is an exercise of this right by a mother against her unwanted fetus. Abortion, then, is morally permissible, self-defensive killing, or can be justified as such for any unwanted pregnancy prior to the viability of the fetus. And this will be the case even if we assume that the fetus is a person from conception on.

The argument goes like so.

1. Any unwanted pregnancy is a serious, nonconsensual use of one person's body by another.
2. Any serious, nonconsensual use of one person's body by another justifies lethal self-defense.
3. So, any unwanted pregnancy justifies lethal self-defense.

I state the argument in this way to note a symmetry that will become important later on. Needless to say, for those who use this argument as a defense of abortion, the "user" is the fetus, and it is the mother who is thereby licensed to use lethal force against him. (Keep in mind that lethal self-defense will only be justified if there are no nonlethal means available, which, in this context, means

that the conclusion will only justify abortion that kills the fetus up to the point of viability, after which the use relation can be ended nonlethally.)

All the defenders of this argument I am aware of motivate it by making an analogy between unwanted pregnancy and rape.¹ A woman may, if no other means are available, kill a man if he uses or threatens to use her body sexually without her consent. A rapist is liable to killing, even if the woman begins a sexual encounter with a man and then changes her mind; if he does not stop, she may use lethal force to stop him, given that no lesser means could. The purveyors of the self-defense argument argue that the relationship between a pregnant woman and her unwanted fetus is, while not in every particular the same, morally comparable to the relationship between a victim and her rapist, and so the mother may permissibly use lethal force to end her unwanted pregnancy. (Note: I warn the reader that in what's coming there are descriptions of sexual assault, and, while I have tried to make the discussion as palatable as I can, the subject is impossible to avoid, given the rape analogy and the purpose it serves in underwriting the self-defense argument.)

Before moving on to assess the self-defense argument, I want to linger a bit on the analogy between sexual assault and unwanted pregnancy. The norms that govern sexual relations between normal adult human beings are not ambiguous. And they imply that no person has a *right* to use another person's body sexually. For another person to permissibly use your body in such a way requires your *ongoing* consent. And if a permissible use of your body by another requires your ongoing consent, that person does not, indeed cannot, have a *right* to so use your body. We can see this in the fact that a woman cannot be *responsible* for a sexual assault in any sense that would give her assailant a right to carry out the assault. There can be no such right. If partisans of the self-defense argument are right about the nature of the relationship between mother and fetus, then the fetus *cannot* have a right to use his mother's body, no matter what her role in bringing about the pregnancy, or what other relation she might have to the fetus. Much of the appeal of the self-defense argument is that it would dispense with the need to worry about considerations an abortion critic might adduce for the claim that the fetus has acquired a right to use his mother's body. If the rape analogy is apt, there can be no such right.

A great deal, then, rides on the analogy, and while there seem to be a number

1 See Little, "Abortion, Intimacy, and the Duty to Gestate"; Long, "Abortion, Abandonment, and Positive Rights"; McDonagh, *Breaking the Abortion Deadlock*; and Regan, "Rewriting *Roe v. Wade*." These writers make the same argument using the same analogy, but there is a difference in the rhetoric with which the analogy is pursued, with Little and Regan measured, and Long and McDonagh red in tooth and claw.

of differences between sexual assault and unwanted pregnancy, partisans of the self-defense argument believe that the differences do not, when properly considered, make a difference. We will take a look at them now.

One might think the threat the rapist poses to an innocent person and the threat the unwanted fetus poses to his mother differ in their seriousness because they involve different amounts (or kinds) of harm, and so one might deny that the particular threat the fetus poses would justify lethal self-defense, whereas the threat a rapist poses does. That is, we could view a forced pregnancy as less harmful to a woman than a rape and so question whether the use relationship involved in premise 1 is of the same seriousness involved in premise 2. But even if we grant that there are different amounts of harm involved, this will not affect the validity of the self-defense argument. If someone threatened to stick something, or someone, in your body against your will in the way a fetus is in his mother's body, then it would be morally permissible to use force to stop him—lethal force if necessary—even if the threatened violation, somehow, was not as traumatic as being raped. If there is a problem with the self-defense argument, it is not because the seriousness of the harm involved in a forced carriage does not rise to a level that might justify lethal self-defense. So, we should concede that the relation between mother and fetus is serious enough to do the work the argument needs it to do.

A more important issue is what we are to make of the claim that the fetus is *using* his mother's body. Unlike a rapist, the fetus is innocent of any wrongdoing, and the argument seems to require us to think of the fetus as an *invader* of his mother's body, a *violinist* of her bodily integrity, which she may, if she wishes, use lethal means to expel. That the fetus's interests are served by remaining in his mother's body does not make him a user of her body in the sense that he intends to treat her as an instrument to secure his ends, and if this is the sort of use that must be involved for lethal self-defense to be justified, the argument equivocates. We cannot interpret "nonconsensual use" as involving agency on the part of the user without making the first premise false. And it would appear that this is just the sort of use that, as the rape analogy suggests, paradigmatically justifies lethal self-defense.

This being so, we need a sense of "nonconsensual use" that does not involve an exercise of agency on the part of the "user," but will justify lethal self-defense in any case where the relation obtains. And this will be tricky. If "nonconsensual use" of the mother's body by the fetus means that the fetus is in her body and she wants him out, then the first premise will be true, at least for any woman that seeks to end her pregnancy by abortion, but the second premise will require

an argument.² Abortion critics will not deny that a woman seeking an abortion has a fetus in her body and wants him out, but will deny that she may, at the cost of his life, put him out. When “nonconsensual use” is read this way, the second premise seems to amount to the claim that she may do so. I will come back to this in the next section and show that on this reading premise 2 is false. But before we get to that, I want to briefly address a related worry that is in the literature on “innocent threats.”

For the self-defense argument to work, it must be permissible to kill an innocent person to prevent that person from using one’s body in a serious way. And some recent work has argued that killing innocent threats of any kind is not permissible, although perhaps is excusable in some circumstances.³ If this is indeed the case, then self-defensive killing in the abortion context will not be justified. The reasons for thinking innocent threats—innocent people whose continued existence threatens another innocent person’s life and limb—may not be killed are complex and deserve to be discussed on their own terms, which I will not attempt to do here. Suffice it to say that if those arguments are sound, then the self-defense argument is obviously unsound. If you cannot kill an innocent person whose continued existence threatens your life, then presumably you cannot kill an innocent person who poses a lesser threat. The abortion debate would then just be a question about the fetus’s moral status. That said, everything that follows here should be understood as conditional on the permissibility of killing at least some innocent persons in cases where they pose a threat to life and limb.

The philosophers who argue against the permissibility of killing innocent threats do so in the face of intuitively compelling cases where it seems like innocent threats may permissibly be killed. Thomson’s expanding baby case is used in just this way, to sway the reader to think that it would be permissible to kill an innocent person whose continued existence is a threat to one’s life.⁴ Roderick Long, a partisan of the self-defense argument, constructs a case to convince the reader that a threat to one’s bodily autonomy can also justify the killing of an innocent person. Speaking of the innocence of the fetus as a reason to doubt that unwanted pregnancy justifies a lethal response, he says:

- 2 More accurately, we should say that “the fetus is in intimate bodily contact with her and she wants it to stop.” As we will see, to be in this relation with another person does not require them (or part of them) to literally be in your body.
- 3 The most well-known defense of this sort of view is offered by Jeff McMahan. See his “Self-Defense and the Problem of the Innocent Attacker,” *The Ethics of Killing*, “Self-Defense and Culpability,” and *Killing in War*. Also see Doggett, “Recent Work on the Ethics of Self-Defense,” for a review of the recent literature on self-defense.
- 4 Thomson, “A Defense of Abortion,” 52.

Considerations of the threat's innocence or guilt may legitimately affect judgments of the moral proportionality of the response. But when the threat is as personal and intrusive as an unwanted pregnancy, it is difficult to see how the innocence of the fetus could make enough of a difference to justify forcing the mother to quietly endure nine months of what is tantamount to rape. Analogously, even if someone has been involuntarily hypnotized into becoming a literal rapist, his victim still has the right to kill him in self-defense.⁵

According to Long, it is the serious nature of the threat—the intimate bodily contact the continued existence of the fetus imposes on the mother—that generates the permission to kill. And Long's hypnotized rapist seems as compelling a case as any for the claim that a serious threat to one's bodily autonomy can justify a lethal response against an innocent person.

2. THE COUNTEREXAMPLE

I want to adapt the case that Long has given us in the above quote. So, suppose that we have a person, *A*, who has control of a device—the hypnotizer—by means of which he can control the minds and bodies of other people, thereby robbing them of their agency and making them do what he wishes. *A* then uses his device to make some innocent person, *B*, sexually assault another innocent person, *C*. To avoid any prurient details, call the act that *A* makes *B* do to *C* ϕ and stipulate that ϕ takes place over some period of time, that it involves a part of *B*'s body being in *C*'s body, and that this contact is both serious and sexual. So, for short, we have

Hypnotized Rapist 1 (HR1): *A* makes *B* ϕ *C*.

Long claims that *C* would be within his rights to kill *B*, given that that was the only way for *C* to avoid the ϕ -ing. Notice that, in this case, if there is a rapist here, it is *A*. *B* is not raping anyone; he is not doing anything. *A* is using him as an instrument to torment *C*. But, by stipulation, *C* cannot stop the event from happening without killing *B*, and regrettable as that might be, Long thinks it justified, given the nature of what *C* will have to endure if *B* is not killed. And if this is so, then we have a case where one is justified in killing an innocent person to avoid a violation of one's bodily autonomy, which is precisely what the self-defense argument maintains the mother may do to her unwanted fetus.

But even if this is so, the immediate rejoinder will be that while it is permissi-

5 Long, "Abortion, Abandonment, and Positive Rights," 188–89.

ble for a woman to kill an innocent man who has been hypnotized to “rape” her, as in Long’s case, it would not follow that she is entitled to kill a man if she did the hypnotizing. That is, HR1 is unlike the typical unwanted pregnancy in that, in the latter, the woman bears responsibility for the use relation that obtains between her and her fetus. And this suggests a different case.

Keep ϕ the same, and now imagine we have only two people involved: *A* and *B*. *B* is an innocent person, and *A* again uses the hypnotizer to control *B* such that *B* is now ϕ -ing *A*. Call this

Hypnotized Rapist 2 (HR2): *A* makes *B* ϕ *A*.

A, with the aid of the hypnotizer, causes *B* to ϕ *A* in such a way that during that period of time, the only way for the ϕ -ing to stop is for *A*, or someone else, to kill *B*. And during that time, *A* decides that he does not want the ϕ -ing to continue. So we have a case where *B* is “using” *A*’s body in a serious way and *A* wants it to stop, but *A* is not entitled to kill *B* to stop it. Thus, premise 2 of the self-defense argument is false. It is not the case that *A* is justified in using lethal force against *B*, even though the use relation obtains between them: *B* is in *A*’s body in a serious way, and *A* wants him out.

Long defends the contention that even in such a case a person would be justified in lethal self-defense.

A woman never has an obligation—or at any rate, never has an enforceable obligation—to let herself be raped. That’s moral bedrock if anything is. The notion of an enforceable obligation to let one’s body be used by a rapist is a moral obscenity; and the same holds for the notion of an enforceable obligation to let one’s body be used as an incubator by a fetus, even if the mother is responsible for the fetus’ presence there in the first place.⁶

We should agree that there can be no obligation of any kind to let oneself be raped. But in HR2, *A* is the rapist. To think that when *A* changes his mind he has thereby withdrawn his consent to what *B* is doing to him is absurd. *A* cannot withdraw his consent to what he is doing, or has done, to himself. *B* is not involved in the right way for there to be a consent issue in HR2.⁷ And so, premise 2

6 Long, “Abortion, Abandonment, and Positive Rights,” 189. An “enforceable obligation” on Long’s view is a moral obligation that any just society would enforce by law.

7 To avoid any misunderstanding, I am not saying that it does not make sense to talk about a person who has temporarily lost his agency and what he would or would not consent to. *B* does not consent to what *A* is doing to him, for example. But when one does something to oneself, the effects of which are not mediated by the agency of anyone else, then I think the idea of withdrawing one’s consent from those effects makes no sense. *A* may regret what he

of the self-defense argument is false. It is not the case that any serious, nonconsensual use of one's body by another justifies lethal self-defense.

Why is *A* not entitled to use self-defensive force against *B*? The most obvious reason is that *A* is responsible for bringing all this about; he has orchestrated this use of his body by *B*, and he is not licensed to put people into situations where they are a threat to him and then kill them in self-defense. To think otherwise would be to give a license to kill to anyone with the power to put other people in situations where they pose a serious threat.

Another possible explanation is that given *A*'s role in bringing this state of affairs about, *A* is not being violated by *B*'s use of his body. While it is true in HR2 that *B* is in his body in an intimate way and *A* wants him out, *A* is not suffering the sort of violation that *C* suffered, or was going to suffer, in HR1. *A* is not being raped. As I said before, for *A* to claim that when he decides he no longer wants the ϕ -ing to continue he thereby withdraws his consent to *B*'s ϕ -ing him is an absurd misrepresentation of what is going on between him and *B*. He cannot withdraw his consent to what *B* is doing to him because *B* is not doing anything to him. While there is sexual activity going on and *B* is involved in it, *B* is not involved in the right way for this to constitute a violation of *A*'s bodily autonomy by *B*. And one might think that only such a violation would justify lethal self-defense against an innocent person. (Notice that this is consistent with thinking that it is morally permissible for *C* to kill *B* in HR1. In that case, *B* is an innocent person who is being used as an instrument by *A* to violate *C*. *A* is doing the violating, not *B*, but we could say that one may defend oneself from any such violation, even if that defense involves killing an innocent person.)⁸

Whatever the explanation for why it is not permissible for *A* to kill *B* in HR2, the question remains whether premise 2 can be revised to avoid the counterexample and still be general enough to yield the argument's conclusion, or something close to it. I am skeptical.

One could try to draw a distinction between the bodily relation between *A* and *B* and the bodily relation between the mother and fetus, and claim that the latter is even more invasive than the former. One could say that while *B* is certainly in *A*'s body in a very intimate way, *B* is not *enmeshed* in *A*'s body the way

has done to himself, wish things were different, curse the gods, or whatever, but he cannot *withdraw his consent* from anyone in HR2.

8 I do not mean to deny that there might be cases where it would be permissible for a person to kill an innocent person who is not violating, or going to violate, another person's bodily autonomy in the sense involved here. I just want to note that this might be an explanation put forth to explain why *A* may not kill *B* in HR2. There is still work to do figuring out just which situations justify lethal self-defense against innocent persons and which not, and how violations of bodily autonomy, or lack thereof, figure in.

the fetus is with his mother, and it is *this* use relation that always justifies self-defensive killing, while other use relations need not.⁹ But trying to tighten up the relation in this way is going to beg the question. The self-defense argument is trying to prove that the bodily relationship between mother and fetus always justifies lethal self-defense, and trying to specify a bodily relation that is intertwined in just the way the fetus is with his mother is going to pull the conclusion into the premises.

We could restrict the self-defense justification in premise 2 to only those serious, nonconsensual uses of one's body that one is not responsible for having brought about. That is, we might say that if you are a mere bystander with regard to a serious, nonconsensual use of your body, then you are morally permitted to use lethal force to stop or prevent that use, but if you are responsible for the use, you are not so permitted. This would certainly avoid the counterexample. And something like this revised premise may well be true, but this will significantly restrict the scope of the argument's conclusion. The self-defense argument is an argument for (preivable) abortion on demand, and to keep the generality of its conclusion, such a revision of premise 2 will require that no woman bear responsibility for her pregnancy. Since this is not true in most cases of unwanted pregnancy, the self-defense argument would then apply only to those cases in which the woman is not responsible, and so will at best show that abortion is permissible in cases where we (or most of us, anyway) already agree abortion is morally permissible.

One might deny that the woman who has consensual sex that results in pregnancy is thereby responsible for the fetus that results (or might claim that she bears some diminished responsibility different in kind from the responsibility *A* has for putting *B* in the situation involved in HR2). And I will concede that if it is demonstrated that a woman in a typical pregnancy is a mere bystander with regard to the fetus's use of her body, then the counterexample will have been avoided. There is a great deal of literature dealing with responsibility for pregnancy and its implications (if any) for abortion.¹⁰ If the defenders of the

9 In *The Ethics of Killing*, McMahan, although not a defender of the self-defense argument, expresses this thought in the following way:

In some respects, pregnancy is even more invasive than rape. Another individual's entire body becomes lodged within the pregnant woman's body and draws sustenance from it. And, unlike rape, the invasive process continues for nine months unless it is forcibly arrested. If, therefore, the threat to bodily integrity from rape is sufficiently serious to warrant a lethal defensive response, it seems that the more deeply invasive and enduring threat from pregnancy should be as well. (399)

10 For those interested in the responsibility debate concerning abortion, see chapter 4 of David Boonin's *A Defense of Abortion* and the citations therein. One thing I will note about

self-defense argument want to take this tack, then they will have to engage with that literature and make their case. But to make the issue one of responsibility is, I think, to abandon the self-defense argument rather than defend it. The promise of the self-defense argument was that it would elide such complications by appeal to the stringent nature of a woman's right to bodily autonomy. To admit that the issue of responsibility is crucial to the defense of abortion, then, is an admission that the right to bodily autonomy is not up to the task.¹¹

That said, one could revise premise 2 by appealing to the obvious difference in the intentions of the two relevant parties in the cases. In HR2, as described, it is clear that *A* intentionally puts *B* in a serious bodily relationship with himself, whereas in the typical unwanted pregnancy, the mother does not intentionally put a fetus in an intimate relationship with herself but merely foresees that such a relationship might occur. So, we could restrict premise 2 to refer only to those serious bodily relationships one did not intentionally bring about but merely foresaw might come about as a result of one's actions. Since we generally think

responsibility and abortion is that the literature takes it that what is at issue is what your obligations are to those needy persons you create through your voluntary actions. That is, whether responsibility for creating them, as is the case in the typical pregnancy, places an obligation on the creator to provide aid to the created. Abortion critics say yes; abortion defenders, like Boonin, say no. What will make this more difficult for the partisans of the self-defense argument is that given their commitments, not only has the pregnant woman created a needy person but she has also put that needy person in a serious, nonconsensual bodily relationship akin to serious sexual activity. This, I argue in what follows, makes it much harder to avoid the conclusion that the mother's responsibility, however parsed, gives the fetus a strong claim against her killing him.

- 11 Some think that Thomson's people seed case is relevant here in that they believe she has more or less successfully argued by analogy that even if the mother is responsible for the fetus being in her body, it is still morally permissible for her to get an abortion. See Thomson, "A Defense of Abortion," 58–59. This is the wrong way to understand the people seed case. Thomson is considering an argument that abortion is morally impermissible because it violates a "special right" the fetus has to use his mother's body, and the fetus is supposed to have this right because the mother engaged in consensual sex knowing that it might result in pregnancy. Thomson asks why the fetus has this special right and anticipates the response that this is so because the mother could have done, or refrained from doing, something such that the fetus would not now be using her body. Thomson's examples in this part of her paper, including the people seeds, are an attack on the claim that this fact is sufficient to establish that the fetus has this special right to use his mother's body. That is, the people seed case is a counterexample, not an analogy. Thomson does not claim (for good reason) that it would be permissible to kill the people described in these cases; she only claims that they do not have a special right to be using the things they are using simply because one could have done something that would have kept them from using those things. The cases in this section of Thomson's paper, then, do not show that abortion is permissible *in any case*. For an illuminating discussion on how Thomson's various examples are used—and, just as importantly, not used—in her paper, see Watkins, "Re-reading Thomson."

it is worse for a person to intend to put someone in a bad way than it is for a person to do something they (merely) foresee might put someone in a bad way, this seems a plausible attempt to revise the premise to deal with the counterexample and still get to the general conclusion the self-defense argument attempts to establish.

While I grant the general moral distinction between intending and foreseeing being invoked here, that distinction is consistent with there being some things one can be held liable for, even if one did not intend for them to come about. And creating an intimate bodily relationship of the sort that motivates the self-defense argument is just such a case. No one is going to pardon *A* for killing *B* if *A* knowingly did something that would bring *B* into the bodily relationship we have in HR2, even if *A* did not intend to bring about that relationship but merely foresaw that there was a reasonable chance it might occur. I realize that “might” and “reasonable chance” are controversial claims in the abortion context, but in HR2 they are not controversial at all. If doing something that has, say, a one in twenty chance of bringing about an ϕ -ing relationship with an innocent person is something you voluntarily do, then we would hold you liable for it, even if you did not strictly intend for the ϕ -ing to come about. And so, the liability the mother has for her fetus will be the same liability that *A* has for *B* in HR2, given that she had consensual sex that she knew might result in the relationship between her and her fetus, whether she intended to bring it about or not.

There is obviously much more to be said about possible individual cases of unwanted pregnancy, responsibility, and intention: there are “cases and cases,” as Thomson reminds us. I only appeal here to HR2 and what would be permissible for *A* to do to *B*, even if it were the case that *A* did not intend to bring about the intimate bodily relationship but merely foresaw there was a reasonable chance it might occur. That one’s voluntary actions might impose *that* relationship on another person raises the stakes considerably. And since the defender of the self-defense argument views any pregnancy as morally akin to serious sexual activity, it is fair to make a comparison between a typical unwanted pregnancy, which I grant is not intended, and the relationship that *A* brings about in HR2. As we will see, understanding the pregnancy relation in this way, in addition to thwarting the proposed revision of premise 2 by reference to the distinction between intending and foreseeing, will lead us to some more radical conclusions concerning abortion.

3. AN ANALOGY ARGUMENT FROM HR2

Now, HR2 is just a counterexample. If what I have argued so far is correct, it mere-

ly shows us that a premise in the self-defense argument is false. An abortion critic could attempt to make an analogy argument from HR2 to the moral impermissibility of the typical abortion, but this looks like it will be tough going; in fact, the differences between the cases look so stark that it seems bizarre to seriously entertain the comparison. First, in HR2, *A* harms *B* by putting *B* in that intimate bodily relationship with himself, but it is far from clear that a woman harms a fetus by conceiving him. And the harm that *A* visits upon *B* is surely relevant to our judgments about what it is permissible for *A* to do to *B* in HR2. Second, one might reasonably doubt that there really is a symmetry in the use relation between the cases. After all, the woman is not using the fetus for anything; it is the fetus that is in the woman's body, and if anyone is using anyone else in the pregnancy relationship, it is the fetus using his mother. This, at least on the face of it, seems different from what is going on in HR2. Third, *B* in HR2 is a normal human being who will be aware of the harm done to him and will, if he lives, suffer the consequences long after the event itself, whereas a fetus will not be aware of anything. One could argue that this is a morally relevant difference between the cases as well. So, it looks like an analogy between HR2 and the typical pregnancy will not hold together.¹²

But there is still a worry here, and it is generated by the use relation involved in the self-defense argument. If the use relation between mother and fetus is morally akin to intimate, sexual activity, then it is going to matter who puts whom in that relation. In HR2, as I noted above, with *B* being placed in the situation without his consent, it is *B* that is suffering a violation of his bodily autonomy. He is the victim. So, any person who puts another person in that use relation without that other person's consent will be violating that person's bodily autonomy. *B* could, for example, invoke self-defense as a justification for killing *A* to stop the ϕ -ing, if only he could and killing *A* was the only way to stop it. And a third party would be justified in doing so on *B*'s behalf. Given this, why is it not the case that a mother violates the fetus's bodily autonomy by putting him in an intimate relationship with her; why is she not guilty of a violation of a kind that *A* is guilty of? And if she is, why do the same considerations not weigh against the mother's killing her fetus by abortion? After all, she put him in that intimate relationship without his consent and now plans to kill him to end it.

I am going to defend the analogy argument from HR2 to the moral impermissibility of abortion, *not* because I think it shows that abortion is morally im-

12 There is the difference in intentions between the parties in the cases that one might point to as a morally relevant difference. But if what I said above concerning the distinction between intending and foreseeing and its (ir)relevance to our judgments concerning HR2 is correct, then this will be a morally irrelevant difference between the cases.

permissible, but because I think it shows us that the conclusion the self-defense argument is supposed to establish—that previable abortion is morally permissible—is undermined by its premises. Thinking of the relationship between mother and fetus as partisans of the self-defense argument do, coupled with the facts concerning the typical unwanted pregnancy, leads to the opposite conclusion. I will work my way through the differences I mentioned above, and I will argue that the differences between the cases are not morally relevant as long as we are committed to understanding the pregnancy relationship as being comparable to serious sexual contact between two normal human beings. This commitment makes the seemingly bizarre comparison between HR2 and a typical pregnancy apt. This commitment is my target.

The first objection to the analogy argument concerns the harm that *A* visits upon *B* in HR2 that does not appear to be present in the typical unwanted pregnancy. In HR2, it is clear that *A* harms *B* by putting him in the bodily relation. *A* has kidnapped and raped an innocent person. And *A* is now, on top of all that, going to kill him. No one will shed tears over *A*'s predicament in HR2. But a mother, even if she deliberately conceives a fetus, has not thereby harmed the fetus. And the prior harm that *A* has done *B* is surely relevant to our judgment concerning the impermissibility of his killing *B*.

This objection depends on the fact that in HR2, *B* exists prior to the act that puts him in the relationship with *A*. When *A* does what he does, he violates *B*'s liberty rights and *B*'s right to bodily autonomy. And this is very different from what the mother does when she becomes pregnant by consensual sex. The fetus does not exist prior to the act that causes him to be in the relationship and cannot exist unless he remains in it. So, the objection goes, the fetus is not harmed by being put in that relationship, nor does it harm him to remain in the relationship after it is brought about.

Now, while it is true that the fetus's existence depends on his being put into that relationship, if we are to take seriously that that relationship is morally akin to serious sexual activity, we cannot maintain that being put in that relationship does not harm the fetus or, if this is different, that it does not violate his bodily autonomy. It is a commonplace that conceiving a fetus does not harm the fetus since one cannot be made worse off by being brought into existence (unless one's life is not worth living), but this will only be true as long as we do not think the relation thereby established is a violation of his bodily autonomy. In the pregnancy case, the act that results in conception is the act that creates the bodily relationship. If being put in that relationship without one's consent is a serious violation of one's bodily autonomy, the fetus will suffer from the violation, even if the same act brought him into existence. If instead of *A* kidnapping *B*, *A*

simply created him in his garage and then brought him into the bodily contact involved in HR2—or, better yet, it was the case that *A* could *only* create *B* by *also* making it the case that they were in the same sort of bodily contact described in HR2—it would not be permissible, given that *B* is an innocent person, for *A* to perform the act that created *B*, nor would it be permissible, given that the act was already performed, for *A* to kill *B* in order to end the bodily contact that *A* brought about.¹³ The very act that brings on conception would be a morally impermissible act, given that the use relation that results is akin to serious sexual activity. Since one cannot obtain a fetus's consent before the fact, conceiving a fetus will constitute a violation of his bodily autonomy.

Furthermore, it will not matter that the fetus cannot continue to exist without being in that relationship. It is true that the mother can end the relationship without the loss of her life and the fetus cannot end the relationship without losing his. But in HR2, *B* cannot survive without remaining in a condition of intimate bodily contact with *A* once the contact is initiated. *B* is just as dependent for his survival on a continued bodily relationship with *A* as the fetus is with his mother, and still suffers a violation of his bodily autonomy, regardless of his need to maintain that relationship to survive. If this is the price of existence, we might be willing to suffer it, but it is not something we could visit upon an innocent person absent his consent.

So, the fact that being brought into existence, when considered by itself, does not harm the person brought into existence is not a reason to think there is anything wrong with the analogy argument. And crucially, it is the nature of the relation between mother and fetus as the partisan of the self-defense argument understands it that does the work in holding the analogy together.

The second objection I mentioned above concerns the location and activity of the fetus in a typical pregnancy. He is *in* the woman's body, using her resources, changing her hormones, and so on. She is not using him for anything. We cannot therefore believe that there is a symmetry in the use relation between mother and fetus and the use relation between *A* and *B* in HR2, where *A* is subjecting *B* to his nefarious purposes.

However, if the use relation is what partisans of the self-defense argument say it is, it does not matter who is in whose body. For example, a woman could rape a man without putting anything *in* his body, and the man would be entitled to self-defensive action if that were the only way to stop the woman from carrying out the rape. And this is exactly the sort of relationship we have between *A* and *B* in HR2, where (part of) *B*'s body is in *A*'s body. Even though *A*'s body is in inti-

13 The garage case is one Michael Davis came up with in his "Foetuses, Famous Violinists, and the Right to Continued Aid," 277.

mate contact with *B*'s body, there is no part of *A*'s body *in* *B*'s body. But it is still the case that *A* may not kill *B* to end the intimate contact that *A* brought about. Given that there is intimate contact between mother and fetus and this intimate contact is as morally fraught as serious sexual activity, it does not matter who is in whose body; all that matters is who put whom in the relationship. And as for the fact that the fetus is "using" his mother's body to survive, we have already seen that *B*'s survival in HR2 requires his continued bodily contact with *A*. It is hard to see that a difference in the details concerning why this bodily contact is needed for the survival of the fetus and why it is necessary for the continued survival of *B* is a difference that would compromise the analogy.

The last worry about making the analogy argument is that as HR2 is described, *B* is a normal, innocent person that *A* kidnaps and rapes. Surely *B* will suffer the violation visited upon him in a way the fetus will not in an unwanted pregnancy. The fetus is not aware of what is going on, and will never be aware of it if aborted.

But whether the harm, or violation, is perceived does not matter. We could easily imagine that *B* is rendered unconscious by the process of coming to ϕ *A*, that he remains so during the ϕ -ing, and further that he is never made aware of what happened. This makes no difference to what *A* is entitled to do to him, and so it makes no difference that the fetus is not conscious of his mother's "use" of his body. If one party, either intentionally or negligently, puts the other party in that use relation without the other party's consent, they will have thereby violated the bodily autonomy of the other party. I do not see how this is to be avoided. It does not matter who is in whose body or whether the victim is aware of the violation as long as the violation occurs.

Now, analogy arguments being what they are, I cannot prove there are no other differences that make a moral difference. I can only say that I do not see any that would. But the point of the analogy argument is to put pressure on the idea that the mother-fetus relationship is subject to the same norms that govern intimate bodily relationships among normal human beings. Imposing the sort of relationship under consideration here on another without their consent is a crime. Imposing it on the fetus, given that he is a person at conception, will not differ on this score.

4. CONCLUSION

But again, this all depends on the relation between mother and fetus being what the partisan of the self-defense argument thinks it is. That it leads to a position critical of abortion is a reason for the abortion defender to let it go with both hands. That it leads to the conclusion that the typical pregnancy is a violation

of a fetus's right to bodily autonomy—that each discrete, voluntary act of the propagation of our species is an act comparable to sexual assault—gives us all a good reason to think that this rape-analogy business was a bad idea and that the self-defense argument that relies on it is indefensible.

I will not pretend to know how the relationship between mother and fetus is to be understood, given that the fetus is a person. All I hope to have done here is to show that one way of understanding it cannot be correct. That relationship cannot be understood as a relationship subject to the norms that govern intimate bodily relationships between normal adult human beings. Thinking of the pregnancy relation in this way has implications that neither party to the abortion dispute can accept.

Auburn University
hhh0001@auburn.edu

REFERENCES

- Boonin, David. *A Defense of Abortion*. Cambridge: Cambridge University Press, 2003.
- Davis, Michael. "Foetuses, Famous Violinists, and the Right to Continued Aid." *Philosophical Quarterly* 33, no. 132 (July 1983): 259–78.
- Doggett, Tyler. "Recent Work on the Ethics of Self-Defense." *Philosophy Compass* 6, no. 4 (April 2011): 220–33.
- Little, Margaret Olivia. "Abortion, Intimacy, and the Duty to Gestate." *Ethical Theory and Moral Practice* 2, no. 3 (September 1999): 295–312.
- Long, Roderick. "Abortion, Abandonment, and Positive Rights: The Limits of Compulsory Altruism." *Social Philosophy and Policy* 10, no. 1 (Winter 1993): 166–91.
- McDonagh, Eileen L. *Breaking the Abortion Deadlock: From Choice to Consent*. New York: Oxford University Press, 1996.
- McMahan, Jeff. *The Ethics of Killing: Problems at the Margins of Life*. New York: Oxford University Press, 2002.
- . *Killing in War*. New York: Oxford University Press, 2009.
- . "Self-Defense and Culpability." *Law and Philosophy* 24, no. 6 (November 2005): 751–74.
- . "Self-Defense and the Problem of the Innocent Attacker." *Ethics* 104, no. 2 (January 1994): 252–90.

- Regan, Donald H. "Rewriting *Roe v. Wade*." *Michigan Law Review* 77 (1979): 1569–646.
- Thomson, Judith Jarvis. "A Defense of Abortion." *Philosophy and Public Affairs* 1, no. 1 (Autumn 1971): 47–66.
- Watkins, Michael. "Re-reading Thomson: Thomson's Unanswered Challenge." *Journal of Libertarian Studies* 20, no. 4 (Fall 2006): 41–59.

IS EPISTOCRACY IRRATIONAL?

Adam F. Gibbons

IT IS COMMONLY THOUGHT that widespread voter ignorance is a problem for democracies.¹ Ignorant voters can inadvertently vote against their interests. More generally, political leaders can supply harmful laws and policies by catering to the ill-informed preferences of such voters. If the electorate wants bad policies, accommodating politicians will often satisfy their demands. Epistocrats think that we should mitigate the harm caused by voter ignorance by allocating comparatively greater amounts of political power to citizens who possess more politically relevant knowledge.²

One important challenge to epistocracy attacks the underlying assumption that better-informed citizens possess superior political judgment.³ According to this challenge, while better-informed citizens may possess more knowledge of politically relevant facts, much research in political psychology indicates that better-informed citizens are less rational than their ill-informed counterparts, being more susceptible to various biases, more partisan, less open minded, and more prone to engaging in motivated reasoning. Correlatively, their ill-informed counterparts are less biased, less partisan, more open minded, and engage in motivated reasoning with less frequency.⁴ Epistocratic institutions, then, might

- 1 Caplan, *The Myth of the Rational Voter*; Somin, *Democracy and Political Ignorance*; and Brennan, *Against Democracy*.
- 2 Brennan, *Against Democracy*; and Jones, *10% Less Democracy*.
- 3 Friedman, "Democratic Competence in Normative and Positive Theory"; Gunn, "Against Epistocracy"; Hannon, "Are Knowledgeable Voters Better Voters?"
- 4 Box-Steffensmeier and De Boef, "Macropartisanship and Macroideology in the Sophisticated Electorate"; Zaller, "Floating Voters in U.S. Presidential Elections, 1948–2000"; Achen and Bartels, "Blind Retrospection"; Taber and Lodge, "Motivated Skepticism in the Evaluation of Political Beliefs"; Abramowitz and Saunders, "Is Polarization a Myth?"; Taber, Cann, and Kucsova, "The Motivated Processing of Political Arguments"; Hartman and Newmark, "Motivated Reasoning, Political Sophistication, and Associations between President Obama and Islam"; Joslyn and Haider-Markel, "Who Knows Best?"; Lodge and Taber, *The Rationalizing Voter*; Kahan, Peters, Dawson, and Slovic, "Motivated Numeracy and Enlightened Self-Government"; and Guay and Johnston, "Ideological Asymmetries and the Determinants of Politically Motivated Reasoning."

just end up placing more political power in the hands of less rational citizens. There is no guarantee that such institutions would constitute an improvement over existing democratic institutions, and they may even lead to worse governance and worse outcomes.⁵ Call this the *problem of epistocratic irrationality*.

In this paper I argue that the problem of epistocratic irrationality can be overcome. A closer look at the psychological data adverted to by critics of epistocracy shows that it is unclear whether it has the implications such critics claim it has. But even if the critics are right about these implications, it still would not follow that democratic institutions will outperform epistocratic institutions. Much depends on the precise form that the envisaged epistocratic institutions take. A more considered approach to epistocratic institutional design might allow us to reap the benefits of placing more power in the hands of the more knowledgeable, while avoiding some of the costs associated with increasing the power of purportedly more dogmatic citizens.

I begin in section 1 by outlining the problem that voter ignorance presents to democracy, before then discussing the problem of epistocratic irrationality at greater length. In section 2, I argue that critics of epistocracy have downplayed and overlooked several problems with their arguments for epistocratic irrationality. First, the citizens counted as knowledgeable by the standards of the experiments that critics advert to are not always such that they would be knowledgeable according to epistocratic standards. Second, the relationship between levels of political knowledge and political irrationality revealed by such experiments is often more complex than critics of epistocracy suppose. Third, the value of traits such as open mindedness is easy to exaggerate. Even if less politically knowledgeable people were more open minded (and more knowledgeable people less open minded), it is unclear whether this would have the upshots attributed to it by critics of epistocracy. In section 3, I argue that even if the relevant data has the implications that critics claim it has, suitably amended forms of epistocracy could overcome the problem of epistocratic irrationality. I consider two potential strategies: (1) using refined selection mechanisms that avoid placing too much power in the hands of irrational citizens, and (2) increasing epistocratic influence only in those areas where such influence has a reliable track record of improving outcomes. I turn to consider some objections to my claims in section 4. Section 5 concludes by summarizing the main claims of the paper.

5 What it is to govern well is naturally a controversial question. In this paper, I focus solely on the relative *epistemic* merits of democratic institutions as against epistocratic institutions—that is, I focus on the degree to which such institutions (attempt to) create legislation grounded in objective facts, the degree to which such institutions use reliable decision-making methods, and the like.

1. DEMOCRATIC IGNORANCE AND EPISTOCRATIC IRRATIONALITY

The claim that a well-informed population is central to the health of a democracy is widely endorsed.⁶ An uninformed population may not know which policies are in their best interests. An uninformed population cannot hold their representatives meaningfully accountable.⁷ Even epistemic democrats, sanguine as they are about individually ignorant voters, accept that a minimum level of voter competence is a necessary precondition of an epistemically well-functioning democracy.⁸

However, many decades of empirical studies indicate that most voters are ignorant of even basic political facts.⁹ Not only are voters unfamiliar with simple, relatively uncontroversial findings in economics, political science, and other social sciences, they often do not know basic facts such as which politicians were responsible for enacting certain pieces of legislation, what sorts of policies are favored by candidates for office, the various roles played by different political agents and organizations, the identities of their representatives, and so on. If a well-informed population is central to the health of democracies, and if most voters are politically ignorant, then democracies face a serious problem. A necessary condition of a flourishing democracy is seemingly unsatisfied.

What, if anything, should be done to mitigate the harm caused by voter ignorance? Epistocrats think that we should allocate more formal political power to those citizens who possess more knowledge of politically relevant facts.¹⁰ In effect, the harm caused by voter ignorance cannot be mitigated without transitioning away from democratic institutions to some extent. Epistocratic reforms variously try to ensure that only politically knowledgeable people possess political power, or that the political power of politically knowledgeable people is

6 Mill, *Considerations on Representative Government*, ch. 2; Dewey, *Democracy and Education*; and Gutmann, *Democratic Education*.

7 Gutmann and Thomson, *Why Deliberative Democracy?* 35; and Guerrero, "Against Elections."

8 Landemore, *Democratic Reason*; and Goodin and Spiekermann, *An Epistemic Theory of Democracy*.

9 For helpful overviews of the relevant literature, see Caplan, *The Myth of the Rational Voter*; Oppenheimer and Edwards, *Democracy Despite Itself*; Somini, *Democracy and Political Ignorance*; and Brennan, *Against Democracy*.

10 Brennan, *Against Democracy*; Mulligan, "Plural Voting for the 21st Century"; Jones, *10% Less Democracy*; Malcolm, "Epistocracy and Public Interests." By "formal" I mean the sort of political power allocated to one by virtue of some law or institution. The political power granted to one by virtue of a legal right to vote is a paradigmatic example of such power, as are the sort of powers one gains upon occupying certain political offices (such as the various powers associated with being a member of the upper and lower houses of parliaments, the executive branch, and so forth).

amplified relative to less knowledgeable people. Ideally, such reforms would increase the degree to which political decision-making is based upon objective, politically relevant facts.

A seemingly unquestioned assumption underlying arguments for epistocracy is the claim that greater levels of political knowledge confer superior political judgment. But this might be mistaken. Suitable levels of political knowledge are not the only thing central to the health of a well-functioning polity. In addition to well-informed citizens, we also need citizens who are *rational*.¹¹ We need citizens who are willing to appropriately update their beliefs in response to new evidence. We need citizens who can set aside partisan loyalties when it becomes clear that these loyalties lead them astray. We need citizens who are open minded and undogmatic. Knowing what the available evidence indicates on politically relevant issues is, of course, deeply important. But a stubborn resistance to following new evidence can be quite damaging in its own right. Broadly speaking, such resistance might cause us to stick with the status quo even if the available evidence indicates that it is harmful. This has important upshots for the viability of epistocracy. If more knowledgeable citizens are less rational, then it becomes unclear whether epistocratic institutions will yield the benefits ascribed to them by proponents of epistocracy. Epistocratic institutions might solve one problem at the cost of another.¹²

Several critics of epistocracy contend that this is precisely what the available evidence from political psychology shows.¹³ They claim that since more knowledgeable citizens are more partisan and less rational than their less knowledgeable peers, epistocratic reforms may produce even worse outcomes than the existing democratic institutions they are intended to supplant. This is a power-

- 11 By “rational” here I have in mind *epistemic rationality*. We need citizens who are both well-informed and who behave in paradigmatically epistemically rational ways when reasoning about political matters. Epistemically irrational conduct may or may not be instrumentally rational for some citizens.
- 12 An independent problem facing epistocracy stems from the recognition that, in principle, the individually ignorant citizens empowered by democratic institutions might make collectively wise decisions, while the well-informed citizens empowered by epistocratic institutions might make collectively unwise decisions. In short, individual and collective intelligence sometimes pull apart. Since this issue has been extensively discussed elsewhere, I set it aside in this paper, though I note the complications it raises for discussions of the comparative epistemic merits of democracies and epistocracies (Surowiecki, *The Wisdom of Crowds*; Landemore, *Democratic Reason*; Goodin and Spiekermann, *An Epistemic Theory of Democracy*). I consider a variant of this problem, one involving individual and collective rationality, in section 4.
- 13 Friedman, “Democratic Competence in Normative and Positive Theory”; Gunn, “Against Epistocracy”; and Hannon, “Are Knowledgeable Voters Better Voters?”

ful challenge that proponents of epistocracy ought to take seriously, for if these critics are correct then the prospects for epistocracy are dim.

There is much evidence indicating that more knowledgeable citizens are also more *partisan* than their ill-informed peers.¹⁴ The more one knows about politics, the more likely one is to be strongly allied to certain views or politicians while strongly opposed to others. Indeed, the causal direction is often the reverse—people more interested in politics and more allied to certain views will generally tend to acquire more political information. But high levels of partisanship create problems. Michael Hannon puts it well when he writes:

The problem . . . is that the most politically partisan individuals (who are also the most knowledgeable, remember) are *also* the most likely to have their thinking corrupted by politics.¹⁵

In a similar vein, Paul Gunn, drawing on the seminal work of Philip Converse, notes that “the more political knowledge people possess, the more ‘constrained’ by ideology they tend to be.”¹⁶

It is worth understanding why more partisan individuals tend to have their thinking corrupted by politics more often (and more severely) than less partisan individuals. In general, heavily partisan individuals are more likely to possess a self-conception defined in part by their adherence to certain political beliefs.¹⁷ If one were to ask such partisan individuals to describe themselves, they might reply by saying that they are, for instance, a supporter of the second amendment, a fiscal conservative, or something to that effect. These beliefs are a core part of their identity (as they see it). Importantly, however, a body of research in political and cognitive psychology shows that such individuals are more prone to engaging in *motivated reasoning*.¹⁸ Partisans engaged in motivated reasoning are more likely to actively seek out evidence that confirms their preferred views, more likely to unquestioningly accept evidence that supports their views, more

14 Judd and Brauer, “Repetition and Evaluative Extremity”; Box-Steffensmeier and De Boef, “Macropartisanship and Macroideology in the Sophisticated Electorate”; Zaller, “Floating Voters in U.S. Presidential Elections, 1948–2000”; Abramowitz and Saunders, “Is Polarization a Myth?”; Feldman and Price, “Confusion or Enlightenment?”; Hetherington, “Putting Polarization in Perspective”; Joslyn and Haider-Markel, “Who Knows Best?”; and Kalmoe, “Uses and Abuses of Ideology in Political Psychology.”

15 Hannon, “Are Knowledgeable Voters Better Voters?” 3.

16 Gunn, “Against Epistocracy,” 35. Gunn here cites Converse, “The Nature of Belief Systems in Mass Publics,” 213. See also Friedman, *Power without Knowledge*.

17 Haidt, *The Righteous Mind*. Hannon calls such beliefs *identity-constitutive beliefs* (“Are Knowledgeable Voters Better Voters?” 3).

18 Kunda, “The Case for Motivated Reasoning.”

likely to avoid evidence that conflicts with their views, more likely to be disproportionately critical of opposing evidence, and so on.¹⁹ In short, partisans are biased toward conclusions they like and biased against conclusions they dislike. Rather than dispassionately assessing the relevant evidence, partisans press their cognitive abilities into the service of protecting their core political beliefs.

Partisans, then, engage in motivated reasoning more frequently than non-partisans because political beliefs form a central part of their self-conception, which they are strongly motivated to defend.²⁰ The problem for epistocracy becomes clear. Epistocratic institutions would enhance the political power of more knowledgeable citizens relative to their less knowledgeable peers. But these more knowledgeable citizens also happen to be much more partisan than others. Since they are partisan, they will engage in the sort of corrupted thinking outlined above—they will dismiss ideologically inconvenient data, they will rush to conclusions that fit their preconceptions, they will selectively focus their critical attention onto views they dislike, and so on.²¹ By attempting to empower the more knowledgeable among us, epistocracy will also empower those of us who most deeply exemplify a host of problematic epistemic vices. Political ignorance may indeed be a problem worth fixing, but the proposed epistocratic cure may end up worse than the illness, for it is not obvious that amplifying the political power of more partisan, less rational citizens is a reliable way to secure better policies and better outcomes.

2. HOW SERIOUS IS THE PROBLEM OF EPISTOCRATIC IRRATIONALITY?

The problem of epistocratic irrationality highlights a weakness in extant epistocratic proposals. By focusing too much on what voters know and not enough on how voters update their beliefs in light of new evidence, epistocratic reforms risk

19 Taber and Lodge, “Motivated Skepticism in the Evaluation of Political Beliefs”; Taber, Cann, and Kucsova, “The Motivated Processing of Political Arguments”; Hartman and Newmark, “Motivated Reasoning, Political Sophistication, and Associations between President Obama and Islam”; Lodge and Taber, *The Rationalizing Voter*; Bolsen, Druckman, and Cook, “The Influence of Partisan Motivated Reasoning on Public Opinion”; Kuru, Pasek, and Traugott, “Motivated Reasoning in the Perceived Credibility of Public Opinion Polls”; Guay and Johnston, “Ideological Asymmetries and the Determinants of Politically Motivated Reasoning”; and Vegetti and Mancosu, “The Impact of Political Sophistication and Motivated Reasoning on Misinformation.”

20 Kahan, “Misconceptions, Misinformation, and the Logic of Identity-Protective Cognition”; Kahan, Peters, Dawson, and Slovic, “Motivated Numeracy and Enlightened Self-Government”; and Mason, *Uncivil Agreement*.

21 Shani, “Knowing Your Colors”; and Bartels, *Unequal Democracy*.

placing too much power in the hands of citizens who, through their epistemic vices, would often make bad decisions. It is therefore unclear whether such reforms would constitute an improvement over existing democratic institutions.

However, it is important to bear in mind the following caveat regarding the problem of epistocratic irrationality: even if critics of epistocracy are right to claim that the data from political psychology shows that more knowledgeable citizens are typically less rational than their less knowledgeable peers, such data cannot by itself establish that democratic institutions will outperform epistocratic institutions along various dimensions. To establish the superiority of existing democratic institutions would require a serious empirical analysis of the performance of both sorts of institutions. This in turn would require investigating, *inter alia*, the impacts of ignorant voters on policy outcomes compared to the impacts of more knowledgeable but less rational voters. Then this would require an analysis of the importance of the issues that less knowledgeable voters are typically ignorant of, the sorts of issues more dogmatic voters are typically reluctant to change their minds about, the degree to which legislators and policymakers are responsive to both groups of voters, and more. In short, psychological data of the sort appealed to by critics of epistocracy cannot furnish us with evidence regarding the overall costs and benefits of both sets of institutions. Instead, we would need to measure the actual performance of democratic institutions against epistocratic ones.

In response, one might think that critics of epistocracy do not intend to show that the overall costs of epistocracy will outweigh its benefits, or that democratic institutions are superior to epistocratic ones. Instead, one might think that they intend only to *undercut* the claim that epistocratic institutions would be superior to democratic ones, and that they do so successfully.²² Perhaps it is true that critics of epistocracy intend to offer only undercutting evidence. Nonetheless, this is consistent with thinking that the ultimate test of epistocracy consists in implementing different epistocratic institutions and measuring their subsequent performance along various parameters, comparing such performance to the performance of democratic alternatives. For those willing to experiment with novel institutional arrangements, this is important. Additionally, in what follows I show that the critics have failed to successfully undercut the case for epistocracy. In this section I discuss some complications with the interpretation of the relevant psychological data critics defend, arguing that these complications should undermine our confidence in such an interpretation. In the following section I discuss certain classes of epistocratic institutions that can overcome the prob-

22 I thank an anonymous referee for highlighting this point.

lem of epistocratic irrationality *even if* critics of epistocracy have correctly interpreted the relevant psychological data.

Regarding the former, what sort of complications do I have in mind?

2.1. *Who Counts as Knowledgeable?*

First, critics of epistocracy frame the relevant studies as showing that more knowledgeable citizens are typically less rational than their ill-informed counterparts. But this description of the data is underspecified. With this information alone we do not know, for example, in what way the subjects of the experiments are knowledgeable, or whether such subjects would count as knowledgeable according to the standards of different epistocratic proposals. Without knowing whether the subjects of the experiments are the sort of people who would be disproportionately empowered by epistocratic mechanisms, we cannot know whether these studies in fact present a problem for epistocracy. In effect, we cannot know whether there really *is* a problem of epistocratic irrationality—at least, that is, for certain forms of epistocracy.

Consider, for example, the work of Lodge and Taber on motivated political reasoning.²³ Critics of epistocracy point to this work as evidence of the claim that more knowledgeable citizens are more prone to motivated political reasoning than less knowledgeable citizens.²⁴ Lodge and Taber do indeed find evidence of what they call a *sophistication effect*, wherein more knowledgeable citizens—*sophisticates*—are more prone to engaging in motivated reasoning.²⁵ More specifically, more knowledgeable citizens more frequently exhibit both *confirmation bias* (by willingly seeking out confirming arguments rather than disconfirming arguments) and *disconfirmation bias* (by spending more time and energy challenging attitudinally incongruent arguments than they do on attitudinally congruent arguments).²⁶ Additionally, sophisticates *polarize* in their beliefs to a greater degree than non-sophisticates.²⁷ However, it is worth noting that the subjects of their experiments were classed as sophisticates depending on how well they scored on a general political knowledge scale consisting of sev-

23 Taber and Lodge, “Motivated Skepticism in the Evaluation of Political Beliefs”; and Lodge and Taber, *The Rationalizing Voter*.

24 Gunn, “Against Epistocracy”; and Hannon, “Are Knowledgeable Voters Better Voters?”

25 They also find evidence for what they call an *attitude strength effect*, where those citizens with the strongest policy attitudes are most prone to politically motivated reasoning (Lodge and Taber, *The Rationalizing Voter*, 153).

26 For an overview of the evidence they adduce in support of these claims, see Lodge and Taber, *The Rationalizing Voter*, 158–67.

27 Lodge and Taber, *The Rationalizing Voter*, 168.

enteen items, including questions like “What proportion of Congress is needed to override a presidential veto?”²⁸ Notably, Lodge and Taber use a median split to differentiate sophisticates from non-sophisticates.

Similar measures of political knowledge are used in other studies purporting to show that more knowledgeable citizens engage in politically motivated reasoning more often than their less knowledgeable peers. Consider the work of Kuru, Pasek, and Traugott showing that more knowledgeable citizens are more prone to politically motivated reasoning about the results of public opinion polls, or the work of Vegetti and Mancosu showing that more knowledgeable citizens are prone to politically motivated reasoning about various news items.²⁹ The former measures political knowledge using *five* multiple-choice questions regarding recent news events.³⁰ The latter, using a sample of Italian citizens, measures political knowledge using *three* questions regarding the identity of the president of the Republic, the president of the Low Chamber, and the number of deputies in the Low Chamber.³¹

Are these sophisticates the sort of people who would possess more political power under epistocratic institutions? Naturally, the answer varies when different forms of epistocracy are considered. On some forms of epistocracy, the

28 Taber and Lodge, “Motivated Skepticism in the Evaluation of Political Beliefs,” 760; Lodge and Taber, *The Rationalizing Voter*, 84.

29 Kuru, Pasek, and Traugott, “Motivated Reasoning in the Perceived Credibility of Public Opinion Polls”; Vegetti and Mancosu, “The Impact of Political Sophistication and Motivated Reasoning on Misinformation.”

30 Kuru, Pasek, and Traugott, “Motivated Reasoning in the Perceived Credibility of Public Opinion Polls,” 431.

31 Vegetti and Mancosu, “The Impact of Political Sophistication and Motivated Reasoning on Misinformation,” 7–8. Indeed, such measures are pervasive in the literature on the relationship between levels of political knowledge, partisanship, and politically motivated reasoning. Virtually every study either uses their own general knowledge scale (Judd and Brauer, “Repetition and Evaluative Extremity”; Box-Steffensmeier and De Boef, “Macropartisanship and Macroideology in the Sophisticated Electorate”; Federico, “Predicting Attitude Extremity”; Feldman and Price, “Confusion or Enlightenment?”; Hetherington, “Putting Polarization in Perspective”; Taber, Cann, and Kucsova, “The Motivated Processing of Political Arguments”; Hartman and Newmark, “Motivated Reasoning, Political Sophistication, and Associations between President Obama and Islam”; Joslyn and Haider-Markel, “Who Knows Best?”; Bolsen, Druckman, and Cook, “The Influence of Partisan Motivated Reasoning on Public Opinion”; Guay and Johnston, “Ideological Asymmetries and the Determinants of Politically Motivated Reasoning”) or relies upon data from sources such as reports from American National Election Studies (Zaller, “Floating Voters in U.S. Presidential Elections, 1948–2000”; Abramowitz and Saunders, “Is Polarization a Myth?”; Bartels, *Unequal Democracy*; Kalmoe, “Uses and Abuses of Ideology in Political Psychology”). In either case, levels of political knowledge are gauged using relatively low numbers of general knowledge items.

class of sophisticates plausibly more or less aligns with the class of those citizens whose political power would be amplified. A scheme of restricted suffrage deploying minimally demanding voter-qualification exams that test for general political knowledge might be adversely impacted by the findings adduced by critics of epistocracy.³² The same applies to any form of epistocracy allocating more political power to those who possess more knowledge of these basic and general political facts alone.³³ But other forms of epistocracy with more demanding standards might escape the challenge of epistocratic irrationality altogether. For instance, under *veto council epistocracy*, a select group of extremely knowledgeable citizens would constitute an epistocratic council tasked with vetoing potentially harmful laws and policies.³⁴ Acquaintance with only basic questions of general political knowledge would not grant one access to such an institution. Instead, one must possess considerable levels of knowledge of—and perhaps even expertise in—one of several fields, such as economics, sociology, or political science. The experiments pointed to by critics of epistocracy do not establish that such citizens are more prone to engaging in motivated reasoning. Accordingly, forms of epistocracy using similarly demanding qualifications are not obviously subject to the problem of epistocratic irrationality.³⁵

One might grant this point while denying its significance. In the absence of evidence that the sort of citizens who would satisfy extremely demanding qualification requirements behave more rationally than either the citizens identified as sophisticates in the psychological literature or non-sophisticates, why grant such an assumption? Until such evidence is supplied, the importance of the relevant mismatch in standards of knowledge is unclear at best.

This point should be conceded. Still, it is important to recognize that the available data does not say much about citizens that we might call *extreme sophisticates*—that is, citizens who cannot only answer several basic general knowledge items, but who additionally possess considerable knowledge of the structure and function of existing political institutions, economics, political science, sociology, and more. It is at least possible that such citizens, outliers as they al-

32 On restricted suffrage, see Brennan, *Against Democracy*, 211–14.

33 For a discussion of minimal epistocracy—a form of epistocracy restricting itself only to uncontroversial, basic political facts—see Gibbons, “Political Disagreement and Minimal Epistocracy.”

34 Brennan, *Against Democracy*, 215–18.

35 This is not to deny that such epistocracies *might* be harmed by epistocratic irrationality. Whether they are or not is clearly an empirical question. I only claim that the studies adverted to by critics of epistocracy do not establish that they are in fact so harmed.

ready are in the degree to which they are well-informed about politically relevant matters, may differ in how they reason about politics compared to other citizens.

Moreover, there is *some* suggestive evidence that the most well-informed citizens (as measured by the typical scales deployed by political psychologists) can resist partisan influences better than their less knowledgeable peers.³⁶ Consequently, the fact that different forms of epistocracy deploy different measures of knowledge may indeed be important as far as the problem of epistocratic irrationality is concerned.

2.2. *What Is the Relationship between Political Knowledge and Political Rationality?*

Second, critics of epistocracy sometimes characterize the relevant data as showing that higher levels of knowledge positively correlate with a greater propensity to engage in motivated reasoning, adoption of more extreme views, more close-mindedness, and so on. In contrast to more knowledgeable citizens, less knowledgeable citizens are more open minded, less prone to engaging in motivated reasoning, and the like. For instance, in describing the relevant psychological findings, Gunn writes that “these effects tend to occur more frequently and stubbornly among citizens who are relatively *well* informed than among those citizens who are not.”³⁷ This description of the data is not inaccurate, but it omits crucial details, oversimplifying the relationship between levels of political knowledge, political irrationality, and other independent traits. Indeed, the relevant data is oversimplified in at least two important ways.

On the one hand, such a presentation of the psychological findings overlooks the possibility that other factors might be driving both the acquisition of political knowledge and motivated reasoning. Indeed, for many individuals, the key point is that they are *motivated* (perhaps because they are independently partisan) and not that they know more.³⁸ As Taber and Lodge write:

[Our] theory predicts less bias for unsophisticated and uncommitted respondents not because they possess a greater sense of evenhandedness, but rather because they lack the motivation and ability to engage in attitude defense.³⁹

36 Achen and Bartels, “It Feels Like We’re Thinking,” 16, 21. Interestingly, Hannon also acknowledges the fact that the most well-informed citizens may be more rational than merely moderately well-informed citizens (“Are Knowledgeable Voters Better Voters?” 14).

37 Gunn, “Against Epistocracy,” 42.

38 To his credit, Hannon acknowledges this point, writing that “it may be that increasing political knowledge is counterproductive only when it occurs in partisan individuals” (“Are Knowledgeable Voters Better Voters?” 5).

39 Taber and Lodge, “Motivated Skepticism in the Evaluation of Political Beliefs,” 767.

In short, ill-informed citizens are less epistemically virtuous than they are unmotivated, and it is the motivation of knowledgeable citizens that drives politically motivated reasoning, not the acquisition of facts per se. Similarly, some studies find that other traits that sometimes (though not always) correlate with high levels of knowledge are what drive attitude extremity. For instance, Federico argues that a high need to evaluate—a felt need to assess things for positive and negative qualities—is what drives attitude extremity, not high levels of knowledge.⁴⁰

Of course, this by itself does not show that critics of epistocracy are mistaken to highlight the importance of the correlation between high levels of political knowledge and political irrationality. If a tendency to engage in politically motivated reasoning correlates with levels of political knowledge, then epistocratic institutions may encounter a problem of irrationality even if the relevant relationship is not causal. For the problem of epistocratic irrationality to emerge, it is enough that epistocratic institutions disproportionately empower epistemically vicious citizens. Still, as we will see in section 3, the absence of a causal relationship paves the way for more refined epistocratic selection mechanisms that can identify (and subsequently empower) well-informed individuals without thereby empowering irrational individuals.

On the other hand, it should be noted that such a presentation of the data masks the considerable variety one can find among both well-informed and ill-informed citizens. To be sure, the data does indeed indicate that moderately knowledgeable citizens are *in general* less rational than less knowledgeable citizens. However, the gap between these two groups vanishes regarding certain issues.⁴¹ Further, Vegetti and Mancosu argue that more knowledgeable citizens are *less* susceptible to character-related misinformation than knowledgeable citizens.⁴² Again, Taber and Lodge put it best when they explicitly caution against taking their findings to show that less knowledgeable citizens are free from epistemic vice:

Provocative though it may be, this interpretation does not stand up to normative, theoretical, or empirical scrutiny. . . . We find no empirical evidence of principled moderation among the bottom or middle thirds of our sample, whose extremity scores were statistically indistinguishable from those of the most sophisticated participants.⁴³

40 Federico, "Predicting Attitude Extremity," 1287. On the need to evaluate, see Jarvis and Petty, "The Need to Evaluate."

41 Taber and Lodge, "Motivated Skepticism in the Evaluation of Political Beliefs," 765.

42 Vegetti and Mancosu, "The Impact of Political Sophistication and Motivated Reasoning on Misinformation."

43 Taber and Lodge, "Motivated Skepticism in the Evaluation of Political Beliefs," 767.

Ill-informed citizens may be less prone to engaging in motivated reasoning on some (though not all) issues, but they are *somewhat* prone, and they frequently adopt extreme political positions at similar rates to their allegedly less rational, high-information peers. Critics of epistocracy, while not outright denying such facts, fail to recognize their importance. As suggested at the beginning of this section, establishing the superiority of either democratic or epistocratic institutions will involve measuring the differential harmful impact of generally ignorant, moderately rational citizens as against the impact of more highly informed but less rational citizens. Bearing in mind that both sets of citizens display a lot of variety in how they process political information should caution us against too quickly assuming that the psychological data clearly favors one set of institutions over the other.

2.3. *Open Minds Are Overstated*

Last, we should not assume that it is always good when citizens are open minded, nor always bad when they are closed minded.⁴⁴ Ideally, we want citizens to update their beliefs *appropriately* in light of new evidence. We do not want citizens to constantly modify their views in a haphazard fashion. More obviously, we do not want citizens to change their beliefs when their beliefs track the truth, or when the countervailing evidence they encounter is sufficiently weak or misleading. In such cases, we want citizens whose minds are closed to evidence that would only lead them astray. This is not to deny that open mindedness is never an attitude that we should wish to cultivate among the electorate, but it does suggest that our evaluation of the behavior of purportedly irrational, closed-minded, high-information citizens should be more sensitive to the specifics of the beliefs in question. For some beliefs in some settings, open mindedness and a willingness to change one's mind are virtues; for others, they are vices.

For example, drawing upon Zaller, Hannon notes that less knowledgeable citizens are more responsive to the content of individual elections.⁴⁵ Among other things, less informed voters are more likely to reward incumbent candidates presiding over a strong national economy, more responsive to ideological shifts on the part of candidates, and at least as likely as their better-informed peers to respond positively to presidential success in managing foreign affairs. He further concludes, drawing upon Achen and Bartels, that “the *more knowl-*

44 Kruglanski and Boyatzi, “The Psychology of Closed and Open Mindedness, Rationality, and Democracy.”

45 Zaller, “Floating Voters in U.S. Presidential Elections, 1948–2000,” 166; Hannon, “Are Knowledgeable Voters Better Voters?” 4.

edgeable voters tended to ignore or downplay the very considerations that are typically viewed as an appropriate basis for electoral choice.⁴⁶

However, a closer look at the manner in which such voters respond to changes in the economy reveals that their behavior may *not* be rational. Notably, Achen and Bartels are themselves critical of the claim that voters can “reliably form and act upon sensible retrospective judgements at election time.”⁴⁷ On their account, voters who reward presidents presiding over strong economies often exemplify a myopic perspective, neglecting to take into consideration their economic welfare over sufficiently long periods of time. Accordingly, the degree to which less informed voters reward (or punish) presidents for the state of the economy is frequently inappropriate. It does not follow from this, of course, that knowledgeable citizens behave appropriately in not responding to economic conditions during election cycles. Still, this merely reinforces the fact that how we evaluate the behavior of citizens who update their beliefs in response to evidence should be more sensitive to the specific details of any given case. Moreover, it highlights once more the occasionally substantial variation in how citizens with different levels of knowledge process political information.

Summing up, then, the significance of the psychological data adverted to by critics of epistocracy is unclear. The standards by which the psychologists judge that citizens are knowledgeable do not always match epistocratic standards. Rather than a simple relationship between the possession of political knowledge and political irrationality, the data in fact reveals a complex interrelationship between levels of political knowledge, attitude extremity, and political irrationality. Lastly, the virtues of a willingness to change one’s mind on political issues vary with the subject matter, the evidence prompting the update of beliefs, and more. Consequently, our confidence in the interpretation of the psychological data offered by critics of epistocracy should be undermined. If so, the severity of the problem of epistocratic irrationality is uncertain.

Before moving ahead, it is worth considering the following possible response to the claims advanced in this section. One might think that, given the controversial nature of many epistocratic proposals, the burden of justification lies squarely with epistocrats defending the epistemic superiority of their preferred institutions. Rebutting critical discussions of epistocracy by issuing purely negative critiques in return is not enough to discharge this burden. Instead, a positive defense of epistocratic institutions (as against democratic alternatives) is

46 Achen and Bartels, *Democracy for Realists*, 294; Hannon, “Are Knowledgeable Voters Better Voters?” 4.

47 Achen and Bartels, *Democracy for Realists*, 175; see also Achen and Bartels, “Blind Retrospection.”

required. Applied to the case at hand, while it may be true that the data from political psychology lacks the implications critics of epistocracy claim it has, establishing this does not constitute a positive argument for epistocracy.

However, even if this claim about the burden of justification is correct, this sort of response suffers from two problems. First, it effectively leaves worries grounded in the putative irrationality of politically knowledgeable citizens behind, conceding that current claims of epistocratic irrationality are unsuccessful. In short, one cannot make this concession while at the same time claiming that the critics have successfully undercut the case for epistocracy. To be sure, this too does not constitute a positive argument in favor of epistocracy. However, such a concession is indicative of the fact that, as far as the comparative epistemic performance of both democratic and epistocratic institutions is concerned, it is an open question as to which is superior. As alluded to earlier, for those willing to experiment with novel institutional arrangements, this is important. Such individuals, noting the failure of current arguments to demonstrate the irrationality of epistocracy, ought to consider epistocratic institutions a live option—at least, that is, certain sorts of epistocratic institutions with the ability to overcome any *potential* problem with epistocratic irrationality.

Second, as we shall see in the next section, the burden of justification has been met for a certain class of epistocratic institutions. Indeed, since there exists much evidence in favor of the superior epistemic performance of certain epistocratic institutions (relative to more democratic alternatives), the burden of justification plausibly lies with critics of these institutions rather than their supporters.

3. OVERCOMING EPISTOCRATIC IRRATIONALITY

The previous section provided some reasons to doubt the interpretation of the psychological data appealed to by critics of epistocracy. Suppose, though, that this interpretation is accurate. What follows from this? In this section, I argue that even if their interpretation is accurate it still does not follow that we should reject epistocratic institutions in favor of democratic institutions.

The problem of epistocratic irrationality is more severe when the costs of epistocratic irrationality outweigh the benefits of amplifying the political power of more knowledgeable citizens.⁴⁸ However, some epistocratic institutions will plausibly be able to mitigate the harmful effects of epistocratic irrationality bet-

48 To be clear, such a distribution of costs and benefits would still not establish the superiority of democratic institutions. Even if the problem of epistocratic irrationality is severe, it might be that the problem of political ignorance in democracies is even more severe.

ter than others. If they can mitigate these effects such that the costs of irrationality are outweighed by the benefits of empowering more knowledgeable citizens, such institutions will still be viable.

I focus on two strategies for epistocrats. First, I explore whether refined selection mechanisms that avoid placing too much power in the hands of irrational citizens could help to ameliorate epistocratic irrationality. I then discuss implementing only those epistocratic institutions that have a reliable track record of outperforming their more democratic counterparts.

3.1. Refined Selection Mechanisms

A presupposition that seemingly underlies the problem of epistocratic irrationality is that epistocracies are committed to uncritically using mechanisms that empower more knowledgeable citizens. Since, let us suppose, more knowledgeable citizens are less rational, the virtues of such mechanisms are questionable at best. This presupposition is not unreasonable since proponents of epistocracy often place heavy emphasis on knowledge of politically important factual matters. Still, there is no reason why epistocrats cannot expand their focus to include the sort of epistemic virtues rightly stressed by critics of epistocracy.⁴⁹

How might epistocrats appropriately expand their focus? Recall that the psychological data adduced by critics of epistocracy does not show that *all* well-informed citizens are irrational, even setting aside the problems discussed in the previous section. It is perhaps true that, on average, better-informed citizens are less rational than ill-informed citizens, but there are exceptions to this general trend. Most notably, some well-informed citizens are also politically rational. Such citizens can form the target for suitably refined epistocratic *selection mechanisms*. If we could devise ways to identify citizens who are both well-informed *and* less prone to various forms of political irrationality, we could amplify their political power rather than the political power of well-informed citizens *tout court*.

The use of such refined selection mechanisms faces two important obstacles. First, one must devise sufficiently precise tests to distinguish between appropriately rational, knowledgeable citizens and their equally knowledgeable but less rational peers. Second, use of the relevant tests must be feasible given the overall

49 Indeed, the importance of such virtues is already recognized by epistocrats, even if they are discussed less often than knowledge of important factual matters. For example, Brennan defends veto council epistocracy at least partly due to the prospect of such a council providing a forum for its members to engage in careful deliberation (*Against Democracy*, 215–18). Presumably, such careful deliberation is not desirable for its own sake, but for its ability to allow council members to learn from one another, change their minds on certain issues (if necessary), and the like. The veto council, then, provides institutional settings within which attitudes and behaviors associated with political rationality can flourish.

epistocratic arrangement in question. Questions of feasibility here primarily revolve around the degree to which the tests can be used cheaply and effectively, especially given large numbers of citizens. These two obstacles are not independent. For instance, the most precise tests may be prohibitively costly to use when large numbers of people are involved, and this might count against certain forms of epistocracy. Correlatively, cost-effective tests may not reliably distinguish between the relevant groups of citizens. Epistocrats thus face potential trade-offs between accuracy and cost-effectiveness, with such trade-offs needing to be factored into any overall cost-benefit assessment of epistocratic proposals.

Let us consider an example that helps to highlight the difficulties that epistocrats may face in using refined selection mechanisms. A natural strategy for epistocrats looking to overcome the problem of epistocratic irrationality is to implement a *plural voting* scheme wherein the most epistemically virtuous citizens receive comparatively more votes than either their ignorant and ill-informed peers or their well-informed but irrational peers.⁵⁰ A nice feature of this form of epistocracy is that, in principle, it could accommodate the claim that the latter group of citizens more harmfully impact political outcomes than the former group. Roughly speaking, we have four groups of citizens to consider: (i) ignorant and irrational citizens, (ii) ignorant but more rational citizens, (iii) well-informed but irrational citizens, and (iv) well-informed and rational citizens. Depending on the magnitude of harm ascribed to empowering each group (or the magnitude of expected benefits in the case of the last group), we could allocate numbers of votes accordingly. If well-informed but irrational citizens make worse decisions than ignorant but more rational citizens, we could modify the numbers to diminish the influence of the former. However, the central aim of such a scheme of plural voting would be to amplify the power of the most epistemically virtuous citizens.

However, empowering the right citizens is easier said than done. We must first identify the relevant people. One option to consider is the use of *indirect* tests of political rationality such as standard measures of cognitive reflection and rationality quotients.⁵¹ If scoring well on such measures reliably correlates with a propensity to engage with politically contentious issues in a suitably rational manner, such tests may work well as proxies for political rationality. However, if no correlation exists (or if there is an inverse correlation between the relevant traits), use of these indirect tests risks bringing about the very distribution of

50 For more on plural voting, see Brennan, *Against Democracy*, 211–14; and Mulligan, “Plural Voting for the 21st Century.”

51 Frederick, “Cognitive Reflection and Decision Making”; and Stanovich, “The Comprehensive Assessment of Rational Thinking.”

power that critics of epistocracy worry about—namely, one where politically irrational citizens find their political power unduly increased.⁵²

Forgoing indirect measures, one might rely upon measures that *directly* test for political rationality. For instance, we can measure levels of political rationality in the same way that the political psychologists cited by critics of epistocracy measure it. In other words, in addition to examining the degree to which citizens possess knowledge of politically relevant facts, we can assess, among other things, the degree to which they are capable of mitigating confirmation and disconfirmation biases (as well as other cognitive biases), the degree to which they are willing and able to update their beliefs in response to evidence in ways that are incongruent with their prior positions, the degree to which they can fairly and accurately state the strongest case for views they are ideologically opposed to, and so on. Discussing the precise details of such measures will involve empirical concerns that are beyond the scope of this paper. The important point for our purposes is that while such direct measures are far more accurate than indirect measures, their use is likely not feasible for an arrangement such as plural voting, for subjecting large numbers of citizens to these measures is likely too expensive and too time consuming.

Generalizing the preceding discussion, it becomes clear that some forms of epistocracy are more affected by the problem of epistocratic irrationality than others. To a first approximation, forms of epistocracy that involve very large numbers of citizens—up to and including the entire electorate—are most negatively impacted. For example, schemes of restricted suffrage and, as noted, plural voting, are likely not feasible given the problem of epistocratic irrationality and a lack of accurate, cost-effective indirect measures for refined selection mechanisms. At the same time, though, since the cost of utilizing more precise, direct measures of political rationality varies with institutional background, some forms of epistocracy *will* be able to feasibly make use of them. Roughly put, as we reduce the number of people to which we seek to apply *individually* costly measures of political rationality, the *aggregate* cost of using such measures decreases to such an extent that certain epistocratic institutions become correspondingly more feasible.

What forms of epistocracy can avail of these direct measures? To take an example, consider again veto council epistocracy. Recall that under this form of epistocracy a select group of highly competent individuals would constitute an epistocratic council tasked with overseeing the legislative activities of other in-

52 Kahan suggests that there is indeed an inverse correlation between the relevant traits (“Misconceptions, Misinformation, and the Logic of Identity-Protective Cognition”). If this is correct, the use of indirect measures of political rationality may not be feasible whatsoever.

stitutions. Extremely demanding qualification requirements would be in place to ensure that only the most competent and knowledgeable individuals are admitted. Adding further tests with the aim of preventing politically irrational agents from joining the council would yield higher feasibility costs compared to a veto council without these tests. Importantly, though, the relatively low numbers involved ensures that veto council epistocracy has much lower *aggregate* feasibility costs than arrangements like plural voting or restricted suffrage epistocracy. While the latter arrangements employ significantly simpler qualification requirements, the massive number of tests required drives the feasibility costs up. In contrast, the veto council has lower *total* feasibility costs even though it uses relatively sophisticated qualification requirements. For veto council epistocracy, then, further tests to distinguish between prospective council members of varying levels of rationality are not the drawback they were for epistocracies with higher feasibility costs. Accordingly, the use of refined selection mechanisms provides the veto council with a plausible tool to mitigate the problem of epistocratic irrationality.

Something similar is true of other epistocratic institutions aiming to increase the political influence of relatively low numbers of individuals. For instance, the *enfranchisement lottery*, wherein a descriptively representative random sample of the population is selected in order to engage in competence-building exercises so that they become better-informed about ballot options, could easily be tweaked to accommodate concerns regarding epistocratic irrationality.⁵³ Rather than competence-building exercises alone, descriptively random samples of the population could be subjected to additional screening to ensure that only appropriately rational citizens are furnished with voting rights. For another, consider *rule by simulated oracle*.⁵⁴ Under this arrangement, we simulate what the electorate would prefer if they were fully informed about important politically relevant facts. Oversimplifying somewhat, we achieve this by surveying citizens' political preferences together with their demographic information, as well as testing their political knowledge. We can then simulate what their preferences would be if we simulated full knowledge while holding the rest of their demographics fixed. A natural modification to such an institution, then, is to use the same method to simulate political preferences given full information *and* high levels of political rationality.

Summing up, then, while the problem of epistocratic irrationality threatens some epistocratic institutions, others are much less threatened. Restricted suffrage and plural voting may not be feasible, but institutions such as the veto

53 For more on the enfranchisement lottery, see López-Guerra, *Democracy and Disenfranchisement*, 4; and Brennan, *Against Democracy*, 214–15.

54 Brennan, *Against Democracy*, 220–22; Ahlstrom-Vij, “The Case for Modelled Democracy.”

council, enfranchisement lottery, and even rule by simulated oracle can be readily modified to better avoid increasing the power of knowledgeable but irrational citizens.⁵⁵

3.2. *Conservative Epistocratic Institutions*

Refined selection mechanisms provide a potentially useful tool for epistocrats seeking to safeguard more radical epistocratic proposals against the harms of widespread epistocratic irrationality. But there are easier ways to increase the likelihood that the costs of epistocratic irrationality are outweighed by the benefits of empowering more knowledgeable people. Instead of opting for controversial institutional reforms, we could adjust existing institutions in ways that have a good track record of improving performance. In the same vein, we could protect currently successful institutions where disproportionate levels of political power are already placed in the hands of more knowledgeable people. These sorts of *conservative epistocratic proposals* either modify existing institutions in an epistocratic direction or prevent the modification of existing institutions in less epistocratic directions.

By “conservative,” I mean to stress the sense in which these proposals are either manifestations of institutions we already have experience with or are significantly influenced by such institutions. Specifically, these are institutions that have a track record of good performance (relative to some non-epistocratic alternative). Consider the difference between a plural voting scheme utilizing refined selection mechanisms and, say, requiring that civil servants possess certain qualifications. Transitioning from democratic institutions with universal and equal suffrage to plural voting is *risky*, at least in part owing to uncertainty surrounding the efficacy of such an institution. We may have suggestive indirect evidence bearing on the prospective performance of plural voting, but we can point to very few concrete exemplars with an actual track record we can examine.⁵⁶ In contrast, requiring that civil servants possess certain qualifications relevant to the role they occupy is conservative in the sense that we have concrete evidence bearing on the importance of qualifications. If one were to measure the performance of civil engineers with legitimate engineering credentials against

55 Additionally, forms of limited epistocracy empowering a relatively low number of experts in narrowly circumscribed roles could feasibly avail of precise and costly measures of political rationality. On limited epistocracy, see Jeffrey, “Limited Epistocracy and Political Inclusion.”

56 Plural voting schemes of different sorts have historically been adopted by several countries, including the United Kingdom, New Zealand, and the Republic of Ireland. But it is important to note that these forms of plural voting were not identical to the sorts of plural voting defended by contemporary epistocrats, let alone epistocracy with refined selection mechanisms. Their evidential import is thus unclear.

the performance of civil engineers without such credentials, one would quickly see the virtues of qualification requirements that are, strictly speaking, epistocratic constraints on access to civil service positions.

Conservative epistocratic institutions of this sort have recently been defended by Garret Jones.⁵⁷ For example, drawing on a wide range of data from political science, he argues that independent central banks typically outperform central banks more tightly constrained by democratic politics. Independent central banks are more reliably correlated with low rates of inflation, low and stable rates of unemployment, steady economic growth, fewer financial crises, and more.⁵⁸ Maintaining the independence of central banks is epistocratic to the extent that insulation from democratic politics allows members of the central bank to draw upon their expertise in a more consistent fashion than would be possible if they were subject to pressure from the electorate, representatives of the electorate, and so on. Epistocrats, then, may wish to protect currently independent central banks from modification in less epistocratic directions. Alternatively, they may urge states without independent central banks to move in an epistocratic direction.⁵⁹

Jones defends other institutions on similar grounds. For example, he argues that appointed, epistocratic city treasurers typically outperform elected treasurers, doing a better job of managing their city's finances.⁶⁰ Moreover, he suggests that the virtues of such institutions provide us with a blueprint for creating novel epistocratic institutions. Drawing on work from Maskin and Tirole, he writes that "when it is crucial to get the technical details right and when the policy debate is less about values and more about facts and competent execution, that is a likely a good opportunity to delegate power to unelected bureaucrats."⁶¹ For instance, extending the rationale behind maintaining an independent central bank, we could implement a Federal Tax Board.⁶² A broad division of labor could exist between Congress (which would decide the broad contours of policy) and the Federal Tax Board (which would focus on the precise details). Implementation of such a novel epistocratic institution is certainly less conservative than main-

57 Jones, *10% Less Democracy*.

58 Jones, *10% Less Democracy*, 41–62.

59 One may think that having no central bank whatsoever is better than having either an independent central bank or a central bank constrained by democratic politics (Rothbard, *The Case Against the Fed*). But this is consistent with thinking that if we are to have a central bank, then we should have an independent central bank. In such a case, the epistocratic institution is still preferable to the democratic institution on instrumental grounds, even if there is an instrumental justification for abolishing the institution entirely.

60 Jones, *10% Less Democracy*, 76–80.

61 Maskin and Tirole, "The Politician and the Judge"; and Jones, *10% Less Democracy*, 91–92.

62 Jones, *10% Less Democracy*, 93–94. See also Blinder, *Advice and Dissent*.

taining the independence of an already independent central bank. But it is much less radical than, say, transitioning to a political arrangement characterized by use of the enfranchisement lottery since there is already much evidence that institutions of the former sort can outperform democratic alternatives.

For the purposes of this paper, we can remain silent on whether Jones correctly assesses the relevant data. The important point is that opting only for conservative epistocratic institutions offers a simple way for epistocrats to overcome the problem of political irrationality. This claim immediately raises two separate questions. First, what sorts of conservative institutions count as epistocratic? Second, and more important, in what way do these institutions help overcome the problem of epistocratic irrationality?

Regarding the first question, let us count as epistocratic any institution that makes the possession of knowledge or expertise an official requirement of occupying certain roles. According to this account, many existing institutions are epistocratic even if not widely recognized as such. For instance, limits to universal suffrage withholding the right to vote from children are epistocratic to the extent that they are defended on the grounds that children (especially young children) lack sufficient knowledge or judgment to vote competently.⁶³ Similarly, members of the judiciary are subject to epistocratic constraints given that jurisprudential expertise is a legal requirement for attaining the relevant positions. Conservative epistocratic reforms, then, might often take the form of implementing these sorts of relatively uncontroversial institutions (or, alternatively, safeguarding them against those who would wish to remove epistocratic constraints).

Regarding the second question, conservative epistocratic institutions undermine the problem of epistocratic irrationality because there is much evidence suggesting that the relevant institutions outweigh the costs associated with epistocratic irrationality. In a sense, reliance upon conservative institutions builds in a response to the problem of epistocratic irrationality from the outset, since this problem trades on the prospect of the costs outweighing the benefits. This is not to say that conservative institutions will always be superior to more radical ones such as plural voting, the veto council, and so on. But conservative institutions can be defended in a less speculative fashion. In addition to epistocracies using refined selection mechanisms with low feasibility costs, then, conservative epistocratic institutions provide a relatively clear way for epistocrats to rebut worries revolving around epistocratic irrationality—they simply outperform their democratic counterparts, even if more knowledgeable citizens tend to be less rational than less knowledgeable citizens.

63 Brennan, "The Right to a Competent Electorate," 701.

Before moving on, two further clarifications are in order. First, the class of conservative epistocratic institutions can shift over time as more evidence is gathered regarding the efficacy of different institutions. If epistocratic councils were to be formed in many different countries, and if such councils were to perform well, they could eventually be classified as conservative in the relevant sense. As evidence of their solid performance mounts, it would become correspondingly less risky to consider implementing them.

Second, conservative proposals are not necessarily proposals to modify institutions *incrementally* (or proposals to prevent the incremental modification of existing institutions in less epistocratic directions). In general, institutional reform might take place in increments—that is, by slight adjustments along the margins of existing institutions. Whether a given adjustment counts as incremental is vague, but one can point to clear instances of incremental adjustments as against non-incremental adjustments. Intensifying the qualification requirements for access to certain civil service positions is an incremental adjustment, while the abolition of universal suffrage is not. In practice, conservative epistocratic reform will often be incremental. Incremental adjustments may be easier to attempt, and it may subsequently be easier to develop a body of evidence about their performance. But in principle we could have solid evidence about wide-ranging epistocratic reforms. If so, these non-incremental reforms would properly be described as conservative in the sense outlined earlier. If conservative epistocratic reforms tend to be incremental, this at most reflects a contingent fact about what sorts of reforms people are typically willing to attempt.

4. OBJECTIONS AND REPLIES

In this section, I address some objections to the claims defended in previous sections. First, one might worry that empowering collections of individually rational agents may not translate to collectively rational group decision-making. If groups constituted by individually rational agents can behave irrationally, then even epistocracy with refined selection mechanisms might not solve the problem of epistocratic irrationality. Second, one might worry that the problem of epistocratic irrationality can only be solved by epistocratic institutions that would create or exacerbate other more serious problems. If so, epistocratic institutions should not be implemented.

4.1. *Epistocracy and the Independence Thesis*

Individual and group rationality can pull apart. Individually irrational agents can form rational groups, and individually rational agents can form irrational groups.

Call this the *Independence Thesis*.⁶⁴ For example, individual scientists who dogmatically defend certain theories can help ensure that good theories remain within the wider scientific community.⁶⁵ This, in turn, might help the scientific community ultimately converge on the truth. By analogy, perhaps a political decision-making body constituted by individually irrational agents can somehow outperform one constituted by individually rational agents.

The Independence Thesis complicates the process of creating epistemically well-functioning groups, in politics and elsewhere. One cannot simply gather a collection of individually rational agents and subsequently guarantee collectively rational decision-making. One must also pay attention to the group's internal structure, its dynamics, and more. Consequently, the case for using refined epistocratic selection mechanisms becomes considerably more complex since we cannot be sure that the eventual group of individually rational (and well-informed) agents will behave in collectively rational ways. If that is right, then we should be less confident in the ability of refined selection mechanisms to overcome the problem of epistocratic irrationality.

Though programmatic, this sort of objection is undeniably important. If epistocrats are serious about designing epistemically well-functioning institutions, they must consider the internal structure and group dynamics of the relevant decision-making bodies they seek to implement. With that said, there are at least three important qualifications one must bear in mind regarding such an appeal to the Independence Thesis.

First, the Independence Thesis does not say that no decision-making body constituted by individually rational agents can behave in collectively rational ways. Nor does it say that no epistocratic institution can outperform any democratic institution. Instead, it tells us that individual and group rationality pull apart. It cautions us against thinking that when we have a collection of individually rational agents, we *thereby* have a group of agents that will together behave rationally. But whether any given group is in fact organized in epistemically optimal ways is an empirical question. We should not assume that refined selection mechanisms will automatically yield rational groups, to be sure, but we also should not assume that we cannot successfully use refined selection mechanisms to mitigate the problem of epistocratic irrationality.

64 Mayo-Wilson, Zollman, and Danks, "The Independence Thesis." There are in fact several different formulations of the underlying insight that individual and group rationality can diverge. Strictly speaking, then, there is no single independence thesis (Mayo-Wilson, Zollman, and Danks, "The Independence Thesis," 655). I set aside this complication moving forward, writing of *the* independence thesis for convenience.

65 Zollman, "The Epistemic Benefit of Transient Diversity."

Second, an epistocratic *tu quoque*—much as the independence thesis complicates the case for epistocracy with refined selection mechanisms, it complicates the case against epistocracy grounded in the problem of epistocratic irrationality. The force of the problem of epistocratic irrationality lies in the fact that much psychological research seemingly indicates that the sort of well-informed citizens who would be empowered by epistocratic institutions are also typically less rational than their ill-informed peers. However, per the Independence Thesis, groups of individually irrational agents can constitute collectively rational groups. Individually irrational and knowledgeable citizens might constitute epistemically well-functioning groups despite their individual epistemic vices. Critics of epistocracy appealing to the Independence Thesis, then, undermine the case for refined selection mechanisms at the cost of undermining their initial critique.

Third, and more positively, conservative epistocratic institutions are not subject to this worry. At the very least, this sort of worry is much less serious for conservative epistocratic institutions. By stipulation, we can already be reasonably confident that the relevant institutions strike an appropriate balance between individual and group rationality. At the very least, we can be reasonably confident that conservative epistocratic institutions do a better job of striking such a balance than their democratic counterparts. If no such institutions existed, then the Independence Thesis would be much more troublesome for epistocrats. But since we can find examples of conservative epistocratic institutions, then at least some epistocratic institutions remain viable.

Ultimately, then, the appeal to the Independence Thesis fails. For use of refined selection mechanisms, it is inconclusive, at most suggesting that we ought to be cautious in assuming that group rationality will emerge from the interaction of individually rational agents. For conservative epistocratic institutions, we can already be confident that group rationality does indeed emerge from the interactions of the agents constituting the relevant group, at least to a certain extent.

4.2. *The Problem of Unintended Consequences*

To be fully successful, political institutions designed to solve certain problems need to avoid creating or exacerbating comparably severe (or even worse) problems. An anti-corruption agency with exorbitantly high operational costs, even if it functions as intended, might be a net cost if the funds allocated to its operation could have been better used elsewhere.⁶⁶ Along the same lines, epistocratic institutions with the means to overcome the problem of epistocratic irrationality

66 In general, the optimal level of corruption in any given society might be nonzero. Cf. Klitgaard, *Controlling Corruption*.

might solve that problem only while creating or exacerbating others. If the costs of these other problems exceed the benefits of the relevant epistocratic institutions, we should deem these institutions failures, even if they successfully tackle the problem they were designed to tackle.

For instance, suppose that a veto council can successfully overcome the problem of epistocratic irrationality, with its low feasibility costs allowing it to deploy highly multifaceted selection mechanisms. But suppose further that this council would increase risks of abuse and corruption that, in expectation, outweigh the expected gains.⁶⁷ If so, we ought not implement the veto council. Taken in isolation, the institution is a success, performing its function as intended. But when its overall impact on the larger political arrangement of which it is a part is considered, it is a failure. It worsens other problems, even if unintentionally.

Like the previous appeal to the Independence Thesis, this objection is at best inconclusive. Whether potential complications would arise, even with epistocratic institutions well-equipped to overcome the problem of epistocratic irrationality, is an empirical question. At most, it again counsels us to adopt a cautious approach to the implementation of novel political institutions. However, this is—or ought to be—a perfectly general point. The expected complications of novel institutions need to be factored into their overall assessment of whether they are epistocratic or democratic, and we should proceed with caution if uncertain about the downstream negative consequences of implementing them.

Moreover, and again like the previous objection, this worry is most serious for nonconservative epistocratic proposals for which there is uncertainty regarding their efficacy. But conservative epistocratic institutions are precisely those institutions for which we have some evidence of their merits. Since we have evidence of the expected performance of these institutions, in some cases we can be confident that they will not create or exacerbate problems to such an extent that the gains from mitigating epistocratic irrationality are outweighed by the costs. There may indeed be costs associated with civil service qualification requirements, independent central banks, and the like, but the evidence suggests that these costs are outweighed by various gains.

The problem of unintended consequences does not show that epistocratic

67 Vandamme, “What’s Wrong with an Epistocratic Council?” Two complications are omitted here as they are beyond the scope of this paper. First, that the veto council might increase the risk of various political abuses is simply taken for granted to illustrate a wider point, namely, that institutions may have unintended consequences that militate against their implementation. Second, I ignore the possibility that epistocratic institutions could be modified to avoid political abuse. Though I cannot defend the claim here, I think that worries about the potential abuse of epistocratic institutions are often overstated, especially given the possibility of modifications that could be made to such institutions to prevent abuse.

institutions are infeasible. For nonconservative proposals, it shows at most that we should be cautious about implementing them. Conservative epistocratic institutions, as before, bypass this problem. We can already be confident that they do not create negative unintended consequences that outweigh their expected benefits.

5. CONCLUSION

Epistocrats perhaps reasonably worry about high levels of political ignorance among voters. But their critics reasonably observe in turn that increasing the political power of those citizens who possess more knowledge is not guaranteed to constitute an improvement over the status quo. If these citizens also happen to be much less rational than their ignorant peers, epistocratic reforms might just make things worse.

It is important in that regard to get clear on what the data from political psychology shows. As argued in section 2, critics of epistocracy overstate the degree to which the relevant findings establish that different epistocratic institutions would empower irrational citizens in harmful ways. More crucially, though, critics of epistocracy have overlooked the possibility that certain incarnations of epistocracy could overcome the problem of epistocratic irrationality entirely, even if the relevant psychological data has the implications that critics claim it has. Perhaps more refined epistocratic selection mechanisms could allow us to screen for irrational citizens, provided the feasibility costs of using such mechanisms are sufficiently low. Perhaps we could play it safe, opting to implement or preserve conservative institutions with a track record of solid performance. Perhaps both could be pursued in tandem. Whatever the case, epistocrats have viable strategies for mitigating the problem of epistocratic irrationality.

On balance, then, while critics of epistocracy have alerted us to potential complications that could arise upon transitioning to some form of epistocracy, they have not shown that all forms of epistocracy are equally suspect. The problem of epistocratic irrationality, if it is indeed a problem, can be overcome.⁶⁸

Rutgers University
adam.gibbons@rutgers.edu

68 Many thanks to Alex Guerrero, Michael Hannon, Oliver Traldi, audiences at Rutgers University and the Forethought Foundation, and two anonymous referees for their feedback on this paper.

REFERENCES

- Abramowitz, Alan I., and Kyle L. Saunders. "Is Polarization a Myth?" *Journal of Politics* 70, no. 2 (April 2008): 542–55.
- Achen, Christopher H., and Larry Bartels. "Blind Retrospection: Electoral Responses to Drought, Flu, and Shark Attacks." Presentation at the Annual Meeting of the American Political Science Association, Boston, 2002.
- . *Democracy for Realists: Why Elections Do Not Produce Responsive Government*. Princeton: Princeton University Press, 2016.
- . "It Feels Like We're Thinking: The Rationalizing Voter and Electoral Democracy." Presentation at the Annual Meeting of the American Political Science Association, Philadelphia, 2006.
- Ahlstrom-Vij, Kristoffer. "The Case for Modelled Democracy." *Episteme* (forthcoming). Published ahead of print, May 6, 2020. <https://doi.org/10.1017/epi.2020.10>.
- Bartels, Larry. *Unequal Democracy: The Political Economy of the New Gilded Age*. Princeton: Princeton University Press, 2008.
- Blinder, Alan. *Advice and Dissent: Why America Suffers When Economics and Politics Collide*. New York: Basic Books, 2018.
- 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.
- Box-Steffensmeier, Janet M., and Suzanna De Boef. "Macropartisanship and Macroideology in the Sophisticated Electorate." *Journal of Politics* 63, no. 1 (February 2001): 232–48.
- Brennan, Jason. *Against Democracy*. Princeton: Princeton University Press, 2016.
- . "The Right to a Competent Electorate." *Philosophical Quarterly* 61, no. 245 (October 2011): 700–24.
- Caplan, Bryan. *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*. Princeton: Princeton University Press, 2007.
- Converse, Philip. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, edited by David Apter, 206–61. New York: Free Press, 1964.
- Dewey, John. *Democracy and Education*. New York: Macmillan, 1916.
- Federico, Christopher M. "Predicting Attitude Extremity: The Interactive Effects of Schema Development and the Need to Evaluate and Their Mediation by Evaluative Integration." *Personality and Social Psychology Bulletin* 30, no. 10 (October 2004): 1281–94.
- Feldman, Lauren, and Vincent Price. "Confusion or Enlightenment? How Exposure to Disagreement Moderates the Effects of Political Discussion and

- Media Use on Candidate Knowledge." *Communication Research* 35, no. 1 (February 2008): 61–87.
- Frederick, Shane. "Cognitive Reflection and Decision Making." *Journal of Economic Perspectives* 19, no. 4 (Fall 2005): 25–42.
- Friedman, Jeffrey. "Democratic Competence in Normative and Positive Theory: Neglected Implications of 'The Nature of Belief Systems in Mass Publics.'" *Critical Review* 18, nos. 1–3 (2006): 1–43.
- . *Power without Knowledge: A Critique of Technocracy*. Oxford: Oxford University Press, 2019.
- Gibbons, Adam F. "Political Disagreement and Minimal Epistocracy." *Journal of Ethics and Social Philosophy* 19, no. 2 (February 2021): 192–201.
- Goodin, Robert E., and Kai Spiekermann. *An Epistemic Theory of Democracy*. Oxford: Oxford University Press, 2018.
- Guay, Brian, and Christopher Johnston. "Ideological Asymmetries and the Determinants of Politically Motivated Reasoning." *American Journal of Political Science* (forthcoming). Published ahead of print, June 23, 2021. <https://doi.org/10.1111/ajps.12624>.
- Guerrero, Alexander A. "Against Elections: The Lottocratic Alternative." *Philosophy and Public Affairs* 42, no. 2 (Spring 2014): 135–78.
- Gunn, Paul. "Against Epistocracy." *Critical Review* 31, no. 1 (2019): 26–82.
- Gutmann, Amy. *Democratic Education*. Princeton: Princeton University Press, 1999.
- Gutmann, Amy, and Dennis F. Thomson. *Why Deliberative Democracy?* Princeton: Princeton University Press, 2004.
- Haidt, Jonathan. *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. London: Penguin Books, 2012.
- Hannon, Michael. "Are Knowledgeable Voters Better Voters?" *Politics, Philosophy and Economics* (forthcoming). Published ahead of print, January 10, 2022. <https://doi.org/10.1177/1470594x211065080>.
- Hartman, Todd K., and Adam J. Newmark. "Motivated Reasoning, Political Sophistication, and Associations between President Obama and Islam." *Political Science and Politics* 45, no. 3 (July 2012): 449–55.
- Hetherington, Marc J. "Putting Polarization in Perspective." *British Journal of Political Science* 39, no. 2 (April 2009): 413–48.
- Jarvis, W. Blair G., and Richard E. Petty. "The Need to Evaluate." *Journal of Personality and Social Psychology* 70, no. 1 (1996): 172–94.
- Jeffrey, Anne. "Limited Epistocracy and Political Inclusion." *Episteme* 15, no. 4 (December 2018): 412–32.

- Jones, Garrett. *10% Less Democracy: Why You Should Trust the Elites a Little More and the Masses a Little Less*. Stanford: Stanford University Press, 2020.
- Joslyn, Mark R., and Donald P. Haider-Markel. "Who Knows Best? Education, Partisanship, and Contested Facts." *Politics and Policy* 42, no. 6 (December 2014): 919–47.
- Judd, Charles M., and Markus Brauer. "Repetition and Evaluative Extremity." In *Attitude Strength: Antecedent and Consequences*, edited by Richard E. Petty and Jon A. Krosnick, 43–72. Mahwah, NJ: Lawrence Erlbaum, 1995.
- Kahan, Dan M. "Misconceptions, Misinformation, and the Logic of Identity-Protective Cognition." Cultural Cognition Project Working Paper Series, no. 164. <http://ssrn.com/abstract=2973067>.
- Kahan, Dan M., Ellen Peters, Erica Dawson, and Paul Slovic. "Motivated Numeracy and Enlightened Self-Government." *Behavioral Public Policy* 1, no. 1 (2017): 54–86.
- Kalmoe, Nathan P. "Uses and Abuses of Ideology in Political Psychology." *Political Psychology* 41, no. 4 (August 2020): 771–93.
- Klitgaard, Robert. *Controlling Corruption*. Berkeley: University of California Press, 1988.
- Kruglanski, Arie W., and Lauren M. Boyatzi. "The Psychology of Closed and Open Mindedness, Rationality, and Democracy." *Critical Review* 24, no. 2 (2012): 217–32.
- Kunda, Ziva. "The Case for Motivated Reasoning." *Psychological Bulletin* 108, no. 3 (November 1990): 480–98.
- Kuru, Ozan, Josh Pasek, and Michael W. Traugott. "Motivated Reasoning in the Perceived Credibility of Public Opinion Polls." *Public Opinion Quarterly* 81, no. 2 (Summer 2017): 422–46.
- Landemore, Hélène. *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton: Princeton University Press, 2013.
- Lodge, Milton, and Charles S. Taber. *The Rationalizing Voter*. Cambridge: Cambridge University Press, 2013.
- López-Guerra, Claudio. *Democracy and Disenfranchisement: The Morality of Electoral Exclusions*. New York: Oxford University Press, 2014.
- Malcolm, Finlay. "Epistocracy and Public Interests." *Res Publica* (forthcoming). Published ahead of print, June 17, 2021. <https://doi.org/10.1007/s11158-021-09502-7>.
- Maskin, Eric, and Jean Tirole. "The Politician and the Judge: Accountability in Government." *American Economic Review* 94, no. 4 (September 2004): 1034–54.

- Mason, Lilliana. *Uncivil Agreement: How Politics Became Our Identity*. Chicago: University of Chicago Press, 2018.
- Mayo-Wilson, Conor, Kevin J.S. Zollman, and David Danks. "The Independence Thesis: When Individual and Social Epistemology Diverge." *Philosophy of Science* 78, no. 4 (October 2011): 653–77.
- Mill, John Stuart. *Considerations on Representative Government*. LaVergne, TN: Book Jungle, 2010.
- Mulligan, Thomas. "Plural Voting for the 21st Century." *Philosophical Quarterly* 68, no. 271 (April 2018): 286–306.
- Oppenheimer, Danny, and Mike Edwards. *Democracy Despite Itself: Why a System That Shouldn't Work at All Works So Well*. Cambridge, MA: MIT Press, 2012.
- Rothbard, Murray N. *The Case Against the Fed*. Auburn, AL: The Ludwig von Mises Institute, 1994.
- Shani, Danielle. "Knowing Your Colors: Can Knowledge Correct for Partisan Bias in Political Perceptions?" Annual Meeting of the Midwest Political Science Association, 2006.
- Somin, Ilya. *Democracy and Political Ignorance: Why Smaller Government Is Smarter*. Stanford: Stanford University Press, 2013.
- Stanovich, Keith E. "The Comprehensive Assessment of Rational Thinking." *Educational Psychologist* 51, no. 1 (2016): 23–34.
- Surowiecki, James. *The Wisdom of Crowds*. New York: Anchor Books, 2005.
- Taber, Charles S., Damon Cann, and Simona Kucsova. "The Motivated Processing of Political Arguments." *Political Behavior* 31, no. 2 (June 2009): 137–55.
- Taber, Charles S., and Milton Lodge. "Motivated Skepticism in the Evaluation of Political Beliefs." *American Journal of Political Science* 50, no. 3 (July 2006): 755–69.
- Vandamme, Pierre-Étienne. "What's Wrong with an Epistocratic Council?" *Politics* 40, no. 1 (2019): 90–105. <https://doi.org/10.1177/0263395719836348>.
- Vegetti, Federico, and Moreno Mancosu. "The Impact of Political Sophistication and Motivated Reasoning on Misinformation." *Political Communication* 37, no. 5 (2020): 678–95.
- Zaller, John. "Floating Voters in U.S. Presidential Elections, 1948–2000." In *Studies in Public Opinion: Attitudes, Nonattitudes, Measurement Error, and Change*, edited by Willem E. Saris and Paul M. Sniderman, 166–214. Princeton: Princeton University Press, 2004.
- Zollman, Kevin J.S. "The Epistemic Benefit of Transient Diversity." *Erkenntnis* 72, no. 1 (January 2010): 17–35.

ON THE NORMATIVE CONNECTION BETWEEN PATERNALISM AND RIGHTS

Stephanie Sheintul

SOME SCHOLARS working on the ethics of paternalism are interested in whether there is a systematic normative connection between hard paternalism and people's moral rights. This discussion understands hard paternalism as paternalism that takes place between competent adults, i.e., "human beings in the maturity of their faculties."¹ One affirmative view is that hard paternalism is *pro tanto* wrong inasmuch as it always involves a rights violation.² Daniel Groll defends this view on the grounds that hard paternalism always violates a competent adult's right to be the only one to act only (or overridingly) for his own good.³ I call this right the *right to self-beneficence*.

I argue that Groll misidentifies a right that competent adults have. Rather than the right to self-beneficence, I argue that if hard paternalism violates any particular right, it is a right that a competent adult has against others "taking over" matters that fall within his sphere of legitimate agency or, by extension, the legitimate exercise of his agency. Following George Tsai, "taking over" can consist of an agent interfering with, intruding on, circumventing, supplanting, replacing, or ignoring the legitimate exercise of another's agency.⁴ I call this right the *right against legitimate agency interference*.⁵ I am agnostic about whether competent adults actually have this right. But if there is a systematic normative connection between hard paternalism and people's rights, it is in virtue of violating the proposed right.

Three preliminary remarks are in order. First, understanding what paternalism involves is controversial.⁶ For present purposes, I understand paternalism

1 Mill, *On Liberty*, 9. Hereafter, all referenced individuals are competent adults.

2 Groll, "Paternalism and Rights," 124.

3 Groll, "Paternalism and Rights," 124. Hereafter, I use the term "primarily" in place of "overridingly."

4 Tsai, "Rational Persuasion as Paternalism," 87. I added "ignoring" to the above list.

5 Though I use the term "interference" in naming this right, any of the terms in the list above can stand in for it.

6 For an overview of this debate, see Begon, "Paternalism."

as follows. *A* acts paternalistically toward *B* if and only if *A* takes over a matter that falls within *B*'s sphere of legitimate agency or, by extension, the legitimate exercise of *B*'s agency, and is motivated by a distrust of *B*'s agency and a concern for *B*'s good.⁷

Second, while Groll has little to say about the nature of the right to self-beneficence, I find it most charitable to understand this right as a *claim right*. On Wesley Hohfeld's influential account of claim rights, a person's claim right always generates a correlative duty in at least one other person.⁸ Accordingly, the right to self-beneficence generates a duty that others refrain from acting only or primarily for another's good. Relatedly, the right to self-beneficence should be understood as a *pro tanto* claim right. According to Danny Frederick, "a right is *pro tanto* if and only if there are exceptional circumstances in which it is permissible to infringe it, even though it is impermissible to infringe it in normal circumstances."⁹ Conceptualizing the right to self-beneficence as *pro tanto* allows for the possibility that, under certain circumstances, it is permissible for an agent to act only or primarily for another's good.¹⁰

Third, in light of understanding this right as a *pro tanto* claim right, I jettison the language of "rights violation." The view that Groll defends should be reconceived as follows: hard paternalism is *pro tanto* wrong inasmuch as it always involves a *rights infringement*. For Groll, this view is true in virtue of hard paternalism always infringing a person's right to self-beneficence. I argue, however, that if this view is true, it is in virtue of hard paternalism always infringing a person's right against legitimate agency interference.

1. THE RIGHT TO SELF-BENEFACTENCE

According to John Stuart Mill, "in the part [of his conduct] which merely concerns himself, [a person's] independence is, of right, absolute."¹¹ Drawing from Mill, a person might be thought to have a claim right against others interfering in his self-regarding matters. Or, drawing from Joel Feinberg, a person might be thought to have a *right to autonomy*, understood as a right against others inter-

7 I draw from George Tsai and Seana Shiffrin in articulating this understanding of paternalism. See Tsai, "Rational Persuasion as Paternalism," 88; and Shiffrin, "Paternalism, Unconscionability Doctrine, and Accommodation," 218.

8 Hohfeld, *Fundamental Legal Conceptions*.

9 Frederick, "Pro-Tanto Versus Absolute Rights," 375.

10 Groll suggests this much in claiming that people are sometimes morally authorized to act only for another's good. See Groll, "Paternalism and Rights," 124–26.

11 Mill, *On Liberty*, 9.

fering with matters that fall within his private realm. Choices that fall within this realm must not directly violate the interests of others.¹²

As Groll discusses, both of these rights are difficult to conceptualize. A worry with the Millian right is that there are few actions that are self-regarding. Similarly, a worry with the Feinbergian right is that it is difficult to identify the boundaries between the private realm and the public realm.¹³ These difficulties might lead one to worry that any attempt to carve out a sphere of choices over which a competent adult is sovereign will result in conceptual complications.¹⁴

In light of this consideration, Groll suggests that a dialectic shift is in order. Rather than focus on *what* an agent is interfering with, one should instead focus on “*why* an agent is doing whatever he is doing.”¹⁵ According to Groll, a competent adult is sovereign when it comes to acting for a particular reason, namely, only or primarily for the sake of his good. If an agent acts only or primarily for the sake of his good, then she infringes his right to self-beneficence.

To help illustrate how hard paternalism infringes this right, Groll offers the following case.

Rob's Keys: Fatima loves listening to loud, live music. This has taken a toll on her hearing. Her doctors have told her that if she doesn't start wearing earplugs to shows, she may well lose her hearing altogether in the coming year. Fatima has ignored their warnings to this point. Tonight, Napalm Death is in town and Fatima is keen to go. Her friend, Rob, is very concerned about Fatima's hearing. The only way for Fatima to get to the concert is to borrow Rob's car. Rob knows that she'll take the car whether he gives her permission to or not, so he hides his keys, thereby preventing her from making it to the concert.¹⁶

According to Groll, Rob acts wrongly in virtue of infringing Fatima's right to self-beneficence. Rob infringes this right by hiding his keys from Fatima primarily because he wants to prevent her from ruining her hearing. If Rob's primary reason for hiding his keys was that he did not want Fatima to use the rest of the gas in his tank, then he would not have infringed her right to self-beneficence.

12 Feinberg, *The Moral Limits of the Criminal Law*.

13 Groll, “Paternalism and Rights,” 124.

14 Though carving out such a sphere may be difficult, doing so is not insurmountable. If hard paternalism infringes a particular right, then constructing a person's “sphere of legitimate agency” is required to properly understand this right.

15 Groll, “Paternalism and Rights,” 124.

16 Groll, “Paternalism and Rights,” 122.

But since his primary reason for acting is his concern for Fatima's hearing, Rob infringes this right.

2. THE GIFT-GIVING OBJECTION

To its credit, the right to self-beneficence is not as difficult to conceptualize as either the Millian right or the Feinbergian right discussed above. Nevertheless, it generates the counterintuitive implication that, in an ordinary instance of gift giving among competent adults the gift giver, in acting only or primarily to make the gift recipient happy, infringes his right to self-beneficence. I call this the *gift-giving objection*.

Groll acknowledges and responds to this objection. He writes:

The general answer to [the gift-giving objection] is that ... there is normally a presumption that our benevolent actions have been *implicitly authorized*. And to the extent that there is *not*, to the extent we believe or don't care that our benevolence is not authorized by the benefactee, then these actions *do* acquire a moral taint.¹⁷

He also writes:

If we act without concern for [whether our help is *welcome*], or our concern for a person's good overrides concerns for *whether the help is wanted*, then I think it is plausible that there is something morally problematic about our action. The same point holds for gift giving and surprise parties: we typically take ourselves to have implicit authorization to do these things and if we really didn't think we did ... the practices start to look morally problematic.¹⁸

As I interpret Groll, the gift giver's *attitude* toward or *belief* about her *implicit authorization* at least partly (if not entirely) determines whether giving a gift is *pro tanto* morally objectionable. What it means for the gift giver to be implicitly authorized, however, is underspecified. As I read Groll, he seems to use the terms "authorization" and "welcome" interchangeably. As such, the gift giver's being *implicitly authorized* to give another a gift just means that the gift giver is *implicitly welcome* to do so.

Supposing this is the correct way to read Groll, the following view presents itself.

17 Groll, "Paternalism and Rights," 125, first emphasis added.

18 Groll, "Paternalism and Rights," 125, emphasis added.

View 1: If the gift recipient's right to self-beneficence is impermissibly infringed, then in acting only or primarily for his good, the gift giver either acts without concern for whether the gift is welcome or she believes that the gift is unwelcome.

If the consequent does not obtain, then it follows that the gift giver does not impermissibly infringe the gift recipient's right to self-beneficence. In an ordinary instance of gift giving, the gift giver usually believes that her gift is welcome. By believing so, she does not impermissibly infringe the gift recipient's right to self-beneficence.

It is worth registering a worry about View 1. The gift giver's attitude toward or belief about her authorization does not correctly explain in virtue of *what* she impermissibly infringes the gift recipient's right to self-beneficence. At best, the gift giver's lack of concern or belief that her gift is unwelcome explains why her act of gift giving is vicious. That the gift giver is not, *as a matter of fact*, implicitly authorized to give the gift recipient a gift, better explains in virtue of what she impermissibly infringes his right to self-beneficence.

This worry, though, is not devastating. Groll could simply accept this point and revise his view as follows.

View 2: If the gift recipient's right to self-beneficence is impermissibly infringed then in acting only or primarily for the gift recipient's good, the gift giver either acts without concern for whether the gift is welcome or she believes that the gift is unwelcome, *and the gift is, in fact, unwelcome.*

A more pressing worry for Groll is that, even when the gift giver believes that she is implicitly authorized and *is* implicitly authorized, she still *permissibly infringes* the gift recipient's right to self-beneficence. When a right is permissibly infringed, the infringer still has "residual obligations," such as to apologize to the person whose right is infringed.¹⁹ But this is the wrong result in ordinary instances of gift giving. To illustrate, consider the following example.

Gift Giving 1: Sam and Amy are co-workers. Sam has told Amy that he enjoys reading for leisure. Sam's birthday is tomorrow and Amy is confident that Sam will enjoy receiving a new book. So, Amy buys Sam a book and gives it to him for his birthday.

In this case, an apology does not seem called for. The best way to make sense of this case is to argue that Sam has *waived* his right to self-beneficence, and in doing so, has removed Amy's *pro tanto* duty not to act only or primarily for his good.

19 See Thomson, *The Realm of Rights*, 85–98.

But this move is not convincing. Sam has not expressly relieved Amy of her *pro tanto* duty. Furthermore, even if this right is sometimes waived within certain relational contexts, Sam's and Amy's relationship (i.e., co-workers) hardly seems like the sort of relational context in which Sam's right is waived.

3. THE RIGHT AGAINST LEGITIMATE AGENCY INTERFERENCE

Nevertheless, I agree that there is something morally problematic about giving a gift to a person who has made it clear that the gift is unwelcome. For illustrative purposes, consider the following case.

Gift Giving 2: Sam and Amy are co-workers. Sam tells Amy that he dislikes reading for leisure and does not want to be given a book for his upcoming birthday. Amy believes that Sam is missing out on the pleasures of reading and that if she gives him a book for his birthday he may come to enjoy reading more. So, despite Sam's wishes, Amy gives him a book for his birthday.

According to Groll, in giving Sam a book for his birthday, Amy impermissibly infringes his right to self-beneficence. However, this is not quite right. While Amy does act primarily for Sam's good in giving him a gift, she also *takes over the legitimate exercise of his agency* by *ignoring* his request that others not give him a book. Presumably, whether one is comfortable with receiving a certain gift is a decision that falls within one's sphere of legitimate agency. By taking over this matter, Amy wrongs Sam.

An important observation is that, in impermissible instances of gift giving, acting primarily for the gift recipient's good is often *coextensive* with the gift giver taking over the legitimate exercise of the gift recipient's agency. I suggest that the latter feature, rather than the former, explains why the gift giver acts wrongly. If the gift giver infringes a particular right of the recipient's, it is more plausibly a right that the gift recipient has against others taking over the legitimate exercise of his agency. Hence, the *right against legitimate agency interference*.

For present purposes, a person's *sphere of legitimate agency* should be understood as a sphere in which he has legitimate control over the matters that fall within it. A competent adult has *legitimate control* over these matters insofar as he has a claim right against others taking over these matters. A person *legitimately exercises his agency* when he makes decisions pertaining to matters that fall within this sphere. Like the right to self-beneficence, the proposed right should be understood as a *pro tanto* claim right. A competent adult has a *pro tanto* duty not to take over the legitimate exercise of another's agency.

A full account of a competent adult's sphere of legitimate agency is beyond the scope of this note. However, following Jessica Begon, I suggest that Martha Nussbaum's influential list of central capabilities can provide us with a helpful starting point.²⁰ Matters pertaining to a person's mental and physical health, body, education, play/recreation, intimacy/intimate relationships, religious and political affiliation(s), conception of the good, identity, and property/personal possessions contribute to a person living a flourishing life.²¹ As such, decisions pertaining to each of these matters plausibly fall within a person's sphere of legitimate agency.

The right against legitimate agency interference more plausibly explains why hard paternalism is *pro tanto* wrong. Like instances of impermissible gift giving, in instances of hard paternalism, the paternalist acting primarily for the good of the recipient is coextensive with the paternalist taking over the legitimate exercise of his agency. This latter feature better explains why hard paternalism almost always involves a rights infringement and, consequently, why it is *pro tanto* wrong.

Additionally, the right against legitimate agency interference does not imply that the gift giver, in ordinary instances of gift giving, permissibly infringes the gift recipient's right. Given that these instances do not ordinarily involve the gift giver taking over the legitimate exercise of the gift recipient's agency, the gift giver neither permissibly nor impermissibly infringes his right.

4. CONCLUSION

Recall the case of Rob's Keys. On Groll's view, Rob acts wrongly in hiding his keys from Fatima in virtue of infringing her right to self-beneficence. This is not the correct explanation. Presumably choices involving Fatima's hearing are hers to make, and her choice is to ignore the recommendations of her doctors. By hiding his car keys from her, Rob interferes with the legitimate exercise of her agency, thereby infringing her right against legitimate agency interference. If hard paternalism is *pro tanto* wrong in virtue of always infringing a right, it is a competent adult's right against legitimate agency interference.²²

University of Wisconsin–Madison
sheintul@wisc.edu

20 See Begon, *Policy without Paternalism*, 66–70.

21 See Nussbaum, *Women and Human Development*, 78–80.

22 I am indebted to Steven Norris and two anonymous reviewers for invaluable feedback on this paper. I also want to thank Harry Brighthouse for encouraging me to write this paper, Lisa Tessman for her unwavering support of my work on paternalism, and several members of

REFERENCES

- Begon, Jessica. "Paternalism." *Analysis* 77, no. 3 (July 2016): 355–73.
- . "Policy without Paternalism: A Capability Approach to Legitimate State Action." PhD dissertation, 2013. <https://etheses.whiterose.ac.uk/5412/>.
- Feinberg, Joel. *The Moral Limits of the Criminal Law*. Vol. 7, *Harm to Self*. New York: Oxford University Press, 1986.
- Frederick, Danny. "Pro-Tanto Versus Absolute Rights." *Philosophical Forum* 45 (December 2014): 375–94.
- Groll, Daniel. "Paternalism and Rights." In *The Routledge Handbook of the Philosophy of Paternalism*, edited by Kalle Grill and Jason Hanna, 119–30. New York: Routledge, 2018.
- Hohfeld, Wesley. *Fundamental Legal Conceptions*. New Haven: Yale University Press, 1919.
- Mill, John Stuart. *On Liberty*. Indianapolis: Hackett, 1978.
- Nussbaum, Martha. *Women and Human Development: The Capabilities Approach*. Cambridge: Cambridge University Press, 2000.
- Shiffrin, Seana. "Paternalism, Unconscionability Doctrine, and Accommodation." *Philosophy and Public Affairs* 29, no. 3 (Summer 2000): 205–50.
- Thomson, Judith Jarvis. *The Realm of Rights*. Cambridge, MA: Harvard University Press, 1990.
- Tsai, George. "Rational Persuasion as Paternalism." *Philosophy and Public Affairs* 42, no. 1 (April 2014): 78–112.

OTHER-SACRIFICING OPTIONS

REPLY TO LANGE

Romy Eskens

IN A RECENT PAPER, Benjamin Lange argues that, when distributing benefits and burdens, we may discount the interests of the people to whom we stand in morally negative relationships relative to the interests of other people.¹ Morality permits not only “positive partiality” toward intimates, but also “negative partiality” toward adversaries. Lange’s case for negative partiality proceeds in two steps. First, he presents a hypothetical example that commonly elicits intuitions favorable to negative partiality. Second, he invokes symmetry considerations to reason from permissible positive partiality to permissible negative partiality. I will argue that neither the intuition elicited by Lange’s example nor the invoked symmetry considerations support a permission for negative partiality. This does not mean that negative partiality is unjustified. It means only that the justification, if there is one, must take a different form.

Here is the plan. Section 1 summarizes Lange’s arguments for permissible negative partiality. Section 2 argues that the intuition-based argument fails because the intuition elicited by Lange’s example is explained by factors other than negative partiality. Section 3 argues that the symmetry-based argument fails because there is asymmetry between partiality to intimates and partiality to adversaries at the level of justification. I end by suggesting an alternative way of justifying negative partiality, one that mirrors gratitude-based justifications of positive partiality rather than justifications based on intimacy.

1. OTHER-SACRIFICING OPTIONS

Lange begins by asking us to conceive of ourselves and our personal relationships as “existing in a kind of moral space.”² In this space, the distance between us and other people is a function of the personal relationship we have with them. The function includes variables such as the nature of the interactions, the inten-

1 Lange, “Other-Sacrificing Options.”

2 Lange, “Other-Sacrificing Options,” 614.

sity of the relationship, and whether the relationship is one-directional or reciprocal. Strangers represent “a morally neutral midpoint” in our moral space.³ Intimates are closer to us than strangers, given the history of positive interactions we have with them, and adversaries are further from us than strangers, given the history of negative interactions we have with them.

With this model in place, Lange begins the appeal to intuitions. He presents:

Callous Colleague: Imagine that Ann’s co-worker Beth has recently been very mean to her with no justified cause. Ann has tried talking to Beth about them having gotten off on the wrong foot, but Beth has not shown any willingness to change. Suppose now that Beth could really use Ann’s help with preparing a document for an upcoming meeting, a favour that Ann could grant Beth easily by having a chat with her, hence making her very happy. However, suppose that Ann could also instead help her new co-worker Chloe with preparing a document for her first weekly report. For this she would have to have an equally long chat with her. This would make Chloe happy, though not quite as much as Beth receiving Ann’s help. Further, suppose that Chloe herself has also had a rough start at her new workplace and has been mean to almost all of her new colleagues except Ann, who she has not met yet.⁴

Intuitively, Lange argues, Ann is permitted to give the smaller benefit to Chloe rather than the larger benefit to Beth. This is not because Beth is impartially less deserving of the benefit than Chloe. Although Chloe has not treated Ann badly, she has treated other people as badly as Beth has treated Ann. Ann’s permission to discount Beth’s interest relative to Chloe’s is instead explained by the fact that Ann has a negative history with Beth but not with Chloe.⁵

Regarding the symmetry argument, Lange first states that “if symmetry considerations lead us from independently plausible phenomenon *A* to some other phenomenon *B*, then this provides some evidence for the plausibility of the phenomenon.”⁶ He argues that a permission for positive partiality is an independently plausible phenomenon, and that negative partiality is the negative mirror image of positive partiality on the two leading pictures of positive partiality: the List Picture, which posits a list of moral permissions, and the Agent-Relative/Neutral Picture, which “understands options as a permission to take on

3 Lange, “Other-Sacrificing Options,” 614.

4 Lange, “Other-Sacrificing Options,” 615

5 Lange, “Other-Sacrificing Options,” 616.

6 Lange, “Other-Sacrificing Options,” 616.

and act from an agent-relative perspective that is sensitive to moral closeness facts.”⁷

2. THE INTUITIONS ARGUMENT

Lange’s intuitive argument for the permissibility of negative partiality relies on our judgments about Callous Colleague. I share the intuition that Ann is permitted to help Chloe rather than Beth in this case, despite that distribution being impartially suboptimal. However, this intuition does not support a permission for negative partiality. It would do so only if Ann is subject to a presumptive requirement to choose the impartially best option, which is then defeated by Ann’s having an adversarial relationship with Beth. But Ann is not subject to such a presumptive requirement: she would be permitted to choose the impartially suboptimal distribution even if her choice concerned two intimates or two strangers instead.

Note first that it is morally optional for Ann to help Beth or Chloe. Ann’s assistance is simply a favor: it would be nice of her to help Beth or Chloe, but withholding assistance would not be wrong. This does not yet mean that Ann is not under a presumptive requirement to help Beth rather than Chloe. One might think that it is impermissible to help in impartially suboptimal ways (at least when the impartially best option is not more costly for the agent), even when it is permissible not to help at all.⁸ If this is right, the fact that Ann seems permitted to help Chloe rather than Beth supports a permission for negative partiality after all. In the absence of Ann’s negative relationship with Beth, it is wrong for Ann to help Chloe rather than Beth.

However, even if it is normally impermissible to help in impartially suboptimal ways, that prohibition does not seem to apply to favors. Suppose you decide to bake a cake for either your brother or your friend, with whom you are equally close. You know that the cake will make your friend happier than it will your brother. Nonetheless, you seem permitted to bake the cake for your brother rather than your friend. That this way of distributing the favor would be impartially suboptimal does not affect your permissible options.

Why would choices about the distribution of favors be exempted from the prohibition on impartially suboptimal benefitting, if choices about the distribution of other types of benefits are not? To answer that question, it is helpful to consider Shelly Kagan’s well-known bird/child case.⁹ Suppose a building is

7 Lange, “Other-Sacrificing Options,” 617.

8 See, for instance, Pummer, “Whether and Where to Give.”

9 Kagan, *The Limits of Morality*, 16.

on fire and you can enter it to try to rescue those inside. Because the rescue is very risky, you are not required to attempt it. However, if you attempt the rescue, and you can save either a child or a bird (but not both), commonsense morality requires you to rescue the child. Suboptimal helping seems impermissible in this case because it involves the death of a child. But giving the cake to your brother rather than your friend involves no significant loss. The loss is merely the bit of extra happiness that your friend would have experienced had you given the cake to them instead of your brother. Similarly, the loss involved in Ann's suboptimal helping is merely the bit of extra happiness that Beth would have experienced had Ann helped Beth instead of Chloe.

Moreover, in the bird/child case, the beneficial act is morally optional because performing it is very costly for you. Had the rescue involved no cost for you, or only a minor cost compared to the child's dying, you would be required to attempt it. Favors are optional for a different reason: they are benefits that we lack claims to regardless of how costly they are for the agent. Your brother and friend each lack a claim that you bake them a cake even if we stipulate that doing so is costless or even beneficial for you (say, because you love baking). Likewise, I argue, Beth and Chloe each lack a claim that Ann help them even if we stipulate that helping is costless or even beneficial for Ann (say, because she enjoys assisting co-workers).

Lange could reply that he assumes consequentialism with options, and that consequentialism with options requires that even favors are distributed in the impartially best way. However, the stated ambition of the article is to defend permissible negative partiality on the basis of commonsense morality.¹⁰ As we have seen, commonsense morality denies that favors ought to be distributed in the impartially best way. Moreover, the intuition that Ann is permitted to choose the impartially suboptimal option in *Callous Colleague* seems to stand or fall with intuitions about the distribution of favors. When we move away from the domain of favors to the domain of uncontroversially required benefits, intuitive support for permissible negative partiality disappears.

Imagine that the boss of the company at which Ann, Beth, and Chloe are employed is known to get disproportionately angry at employees for their minor errors, which often results in unfair demotions. Beth and Chloe will each make a minor error if Ann does not help them, for which they will be unjustly demoted. However, the boss will inform Beth about her demotion in front of her co-workers, in a humiliating way, whereas Chloe will be told in private. May Ann still help Chloe rather than Beth in this scenario? I think not. It seems to me that Ann ought to help Beth rather than Chloe.

10 Lange, "Other-Sacrificing Options," 612, 613.

My intuitions become less clear when considering cases that involve serious wrongdoing. What if Beth has instead been bullying and abusing Ann for weeks, and Chloe has been doing the same to another colleague? Does it seem more plausible now that Ann may help Chloe rather than Beth? I think so, and this change in intuitive judgment suggests that negative partiality might be permissible at least when the wrongdoing is sufficiently severe. Still, there are factors other than negative partiality that could explain the change in intuitive judgment. For example, we might have an excuse for choosing the impartially sub-optimal distribution if the wrong we suffered is severe. Alternatively, our cooperation with the people who severely wronged us might damage our self-respect to the extent that cooperation becomes too costly to be obligatory. Before we can conclude that our intuitions support permissible negative partiality in cases involving severe wrongdoing, these possible alternative explanations need to be eliminated.

3. THE SYMMETRY ARGUMENT

Lange's second argument for permissible negative partiality is that, on the two leading pictures of partiality, negative partiality being permissible is the negative mirror image of positive partiality being permissible. Since permissible positive partiality is an independently plausible phenomenon, symmetry considerations support permissible negative partiality.

The two pictures of partiality that Lange considers—the List Picture and the Agent-Relative/Neutral Picture—are ways of *conceptualizing* partiality, rather than ways of *justifying* it. But even if we agree with Lange that “other-sacrificing options can be *understood* as the negative mirror image of the other-favoring and agent-relative-favoring option,” this does not mean that other-sacrificing is *permissible*.¹¹ Two behaviors can mirror each other without sharing a deontic status. For instance, we commonly think that creating good lives is permissible but ending bad lives is not (unless special conditions are met). So, behaviors that are symmetrical at the conceptual level need not be symmetrical at the deontic level.

Still, if two behaviors are symmetrical at the conceptual level, that normally gives us some reason for suspecting symmetry at the deontic level. But not in this case. In this case, we have good reason to suspect *asymmetry* between positive partiality to intimates and negative partiality to adversaries at the deontic level. Permissions to favor intimates' interests do not always derive from the moral quality of our relationships. If they did, it would be impermissible for a parent to favor their infant child's interests, or for a wife to favor her patriarchal

¹¹ Lange, “Other-Sacrificing Options,” 618, emphasis added.

husband's interests. By contrast, negative partiality seems clearly impermissible when it targets rivals or competitors who frustrate our interests without acting impermissibly, such as people competing with us for professional opportunities. If it is ever permissible to discount adversaries' interests, it must be because they wronged us. But if our aim is to only justify negative partiality toward people who wronged us, partiality toward intimates seems altogether the wrong analogue. A more apt analogue would be partiality toward people who benefitted us in morally significant ways. In that analogue, it is the moral quality of the person's past treatment of us that underpins the partiality, rather than our being intimate with them.

The asymmetry in justification between partiality toward intimates and partiality toward adversaries becomes even more evident when we consider leading justifications for favoring intimates' interests, for none of those justifications seem to point to a symmetrical justification for discounting adversaries' interests. To make this point, it will be helpful to follow Simon Keller's taxonomy of the justifications philosophers have offered in defense of partiality toward intimates, which categorizes them into three camps.¹²

First is the Projects View.¹³ The Projects View holds that permissions for positive partiality are grounded in the significance of ground projects and fundamental commitments. We all pursue certain ground projects and fundamental commitments throughout our lives, and this pursuit gives our lives meaning and defines us as practical agents. Participation in positively partial relationships with intimates is crucial to the pursuit of many of these projects and commitments, and thus for living a worthwhile life and exercising our practical identity. Since the projects to which these relationships are essential seem legitimate, the positive partiality is justified.

For there to be a symmetrical justification for negative partiality, negative partiality must be essential to the pursuit of our projects, and those projects must be legitimate. Neither seems plausible. Discounting the interests of adversaries is not normally crucial to pursuing our projects. Moreover, even if somebody would include engagement in negatively partial relationships under their description of what makes their life worth living and their actions worth undertaking, we would regard their projects as illegitimate. We would think the person deeply misguided about what is worth pursuing in life, and would rightly refuse to treat their projects as a source of moral justifications.

The second view is the Relationships View, which holds that permissions for

12 Keller, *Partiality*.

13 See Stroud, "Permissible Partiality, Projects, and Plural Agency"; Williams, "Persons, Character, and Morality"; and Wolf, "Morality and Partiality."

positive partiality derive from the final or noninstrumental value of intimate relationships.¹⁴ These valuable relationships can exist only if those party to them engage in positively partial behavior toward one another. Since the relationships are desirable for their own sake, the partiality they require to exist is morally justified.

A symmetrical justification for negative partiality would hold that adversarial relationships are finally or noninstrumentally valuable, and that the fact that these relationships can exist only if those party to them act negatively partial justifies the partiality. The idea that adversarial relationships are desirable for their own sake is clearly implausible. But if adversarial relationships are not desirable for their own sake, the fact that they require negatively partial behavior seems to do nothing to justify that behavior.

The third view is Keller's Individuals View.¹⁵ Keller argues that our permissions for positive partiality derive from the value of the individuals with whom we are intimate. These individuals are objectively as valuable as everyone else, but the relationship we have with them enables their value to be a source of reasons for us to favor their interests relative to the interests of others. It is the value our intimates have as individuals that justifies treating them preferentially, but this justification is enabled only when we stand in an intimate relationship with them.

It is plausible that our adversaries objectively have value, and that the reason this gives to treat them preferentially is disabled for us by the adversarial relationship we have with them. However, the absence of a justification for treating adversaries *better* than strangers does not entail a justification for treating adversaries *worse* than strangers. So, this third mirror justification fails to generate a permission for negative partiality as well.

4. CONCLUSION

I have argued that Lange's case for permissible negative partiality is unsuccessful. The argument from intuitions fails because the intuition elicited by Lange's example is explained by factors other than negative partiality. The symmetry argument fails because there is an asymmetry between positive partiality toward intimates and negative partiality toward adversaries at the level of justification. However, I suggested that a better analogue for negative partiality would be positive partiality toward people who benefitted us in morally significant ways (but

14 See Jeske, *Rationality and Moral Theory*; Kolodny, "Love as Valuing a Relationship"; and Scheffler, "Relationships and Responsibilities."

15 Keller, *Partiality*.

with whom we are not necessarily intimate). Reconceptualizing negative partiality in this way opens up a novel line of justification. Commonsense morality holds that we may (or should) favor our benefactors' interests relative to the interests of strangers, to express fitting gratitude. Perhaps we may also discount the interests of people who have wronged us, to express fitting resentment. This line of justification mirrors a gratitude-based justification for positive partiality rather than a justification based on intimate relationships.¹⁶

Stockholm University
romy.eskens@philosophy.su.se

REFERENCES

- Jeske, Diane. *Rationality and Moral Theory: How Intimacy Generates Reasons*. London: Routledge, 2008.
- Kagan, Shelly. *The Limits of Morality*. Oxford: Oxford University Press, 1989.
- Keller, Simon. *Partiality*. Princeton: Princeton University Press, 2013.
- Kolodny, Niko. "Love as Valuing a Relationship." *Philosophical Review* 112, no. 2 (April 2003): 135–89.
- Lange, Benjamin. "Other-Sacrificing Options." *Philosophical and Phenomenological Research* 101, no. 3 (November 2020): 612–29.
- Pummer, Theron. "Whether and Where to Give." *Philosophy and Public Affairs* 44, no. 1 (Winter 2016): 77–95.
- Scheffler, Samuel. "Relationships and Responsibilities." *Philosophy and Public Affairs* 26, no. 3 (July 1997): 189–209.
- Stroud, Sarah. "Permissible Partiality, Projects, and Plural Agency." In *Partiality and Impartiality: Morality, Special Relationships, and the Wider World*, edited by Brian Feltham and John Cottingham, 131–49. New York: Oxford University Press, 2010.
- Williams, Bernard. "Persons, Character, and Morality." In *Moral Luck*, 1–19. New York: Cambridge University Press, 1981.
- Wolf, Susan. "Morality and Partiality." *Philosophical Perspectives* 6 (1992): 243–59.

16 I am grateful to Gunnar Björnsson, Julia Driver, Helen Frowe, James Goodrich, Jonas Haeg, Benjamin Lange, and an anonymous reviewer for helpful comments and discussion. Work on this paper was supported by the Knut and Alice Wallenberg Foundation, grant no. 1521101.

JOURNAL of ETHICS & SOCIAL PHILOSOPHY
<http://www.jesp.org>
ISSN 1559-3061

The *Journal of Ethics and Social Philosophy* (JESP) 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. Articles are typically published under the CREATIVE COMMONS ATTRIBUTION-NONCOMMERCIAL-NODERIVATIVES 4.0 license, though authors can request a different Creative Commons license if one is required for funding purposes.



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.