

JOURNAL *of* ETHICS & SOCIAL PHILOSOPHY

VOLUME XIII · NUMBER 1

March 2018

ARTICLES

Intelligibility and the Guise of the Good

Paul Boswell

Ethical Reductionism

Neil Sinhababu

A View of Racism: 2016 and America's Original Sin

Benjamin Mitchell-Yellin

DISCUSSIONS

Does Contrary-Forming Predicate Negation

Solve the Negation Problem?

Robert Mabrito

Are Moral Error Theorists Intellectually Vicious?

Stephen Ingram

JOURNAL *of* ETHICS & SOCIAL PHILOSOPHY

VOLUME XIII · NUMBER 1

March 2018

ARTICLES

- 1 Intelligibility and the Guise of the Good
Paul Boswell
- 32 Ethical Reductionism
Neil Sinhababu
- 53 A View of Racism: 2016 and America's Original Sin
Benjamin Mitchell-Yellin

DISCUSSIONS

- 73 Does Contrary-Forming Predicate Negation
Solve the Negation Problem?
Robert Mabrito
- 80 Are Moral Error Theorists Intellectually Vicious?
Stephen Ingram

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.

Editor

Mark Schroeder

Associate Editors

James Dreier

Julia Driver

David Estlund

Andrei Marmor

Discussion Notes Editor

Douglas Portmore

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

Susan Wampler

Editorial Assistance

Renee Bolinger

Typesetting

Matthew Silverstein

INTELLIGIBILITY AND THE GUISE OF THE GOOD

Paul Boswell

DAVE CLAIMS he had never before considered suicide, and he does not recall being depressed or even sad at the time. But late one otherwise-ordinary night, while walking down a quiet street, Dave saw a car approaching, and it occurred to him that—as he later put it—“what I should do would be to kneel in the street and be hit by the car. It seemed to make perfect sense to me. And so that’s what I did.”¹ The car screeched to a halt just in front of him, and a man got out and demanded to know what he was doing. According to Dave, “I looked at him, and I said, ‘I don’t know.’ I had no idea. I had no explanation for him. . . . I still today—what, 25 years later—don’t understand what that was and why I did that.”²

In spite of the fact that in some way it made sense to him to kneel down, in another way it must have been utterly *unintelligible* to Dave to kneel. After all, he could find no reason for his action. Cases like Dave’s thus motivate what is often called the Intelligibility Constraint (IC) on action for a reason:

IC: If an agent ϕ s for a reason, then ϕ ing is intelligible to her (according to a certain sense of “intelligible”).

IC is not exactly a platitude, concerning as it does an intuitive yet unarticulated notion of intelligibility (about which more later). Moreover, it depends upon a particular notion of action for a reason that entails having a genuine reason *of one’s own* for acting. If instead, after months of psychotherapy, Dave were to unearth a repressed fear of appearing to be a failure to his father, and to discover that his genuflection so many years ago was a kind of prayer for forgiveness, then there would be another sense in which he knelt for a reason: he did it in the hope that he would be forgiven. (Call this case “Freudian Dave.”) But so long as we stipulate that at the time of action this motivation was fully repressed and that Freudian Dave had absolutely no idea why he was kneeling, then such a motiva-

1 Glass, “Devil on My Shoulder.”

2 Glass, “Devil on My Shoulder.”

tion does not speak to a reason *of his* for kneeling. To be a bit more explicit, I can grant that Freudian Dave has motivating reasons for action in his unconscious beliefs and drives. The sense in which he does not have a reason for action, however, is Anscombian: it is because there is no positive answer to the “Why [did you do that]?” question, at least as addressed to *him*.³ This latter, full-blooded sense of reasons for acting is the sense I am concerned with here.

IC, then, claims that there is a deep connection between this full-blooded sense of acting for a reason and a certain way in which actions may or may not be intelligible to their agents.

One traditional and perennially popular theory in the philosophy of action, the Guise of the Good (GG), is often held by its proponents to explain IC.⁴ According to GG, an action done for a reason must be seen *as good* by its agent, for a certain sense of “seen as good.” Furthermore, it holds that action for a reason must be intelligible to its agent *because* for an action to be intelligible in the relevant way just is for it to be seen as good, in that sense of “seen as good.”⁵ Now, it so happens that current GG theories can be divided almost neatly into two camps according to how they think of such appearances of the good. *Attitudinal* views hold that these appearances have a presenting-as-good character but that normative notions need *not* figure in the content of these states. *Assertoric* views hold that these appearances are better understood as representations with normative content that are presented with assertoric rational significance—that is, the rational significance possessed by belief and perception that consists in their power to defeasibly license further beliefs and inferences from that content.⁶

In this article, I present a dilemma for the attitudinal theorist who aims to ex-

3 Anscombe, *Intention*, 9.

4 See Anscombe, *Intention*; Quinn, “Putting Rationality in Its Place”; Raz, “Agency, Reason, and the Good” and “On the Guise of the Good”; Tenenbaum, *Appearances of the Good*; Sussman, “For Badness’ Sake”; Johnston, “The Authority of Affect,” 189–90; Schapiro, “The Nature of Inclination,” 251–53; and Boyle and Lavin, “Goodness and Desire,” 188–89, among others. Note that Johnston and (sometimes) Anscombe focus more on how a certain relation to evaluation can make *desiring*, as opposed to *acting*, intelligible. Especially in the case of Johnston, this leaves to some extent open the question of whether they accept the precise form of GG I am interested in here, one that concentrates on action for a reason. I leave this exegetical point aside in the rest of the article, however. For a concern about whether Quinn accepts IC, see note 32 below.

5 See Anscombe, *Intention*, 70–78; Tenenbaum, *Appearances of the Good*, 32–33; Wiland, *Reasons*, 49–52.

6 Schroeder (“How Does the Good Appear to Us?”), Schafer (“Perception and the Rational Force of Desire”), and Baker (“The Abductive Case for Humeanism over Quasi-Perceptual Theories of Desire,” 4–5) refer to similar distinctions between kinds of GG theories. The term “assertoric” was introduced to the debate by Schafer (“Perception and the Rational

plain IC. The dilemma arises out of an objection from Kieran Setiya, who argues that merely seeing an action as good does not suffice to render that action intelligible.⁷ (Indeed, as we will see, Dave's case is good evidence that he is right.) If that is so, it imperils GG's explanation of IC. I show that the very feature that GG theories need in order to answer this objection—the feature of motivation's presenting a *specific* form of the good—forces them to characterize their view in a way that either favors the assertoric model or fails to capture the intelligibility of action. The crux of the argument is that attitudinal views are fundamentally unable to account for agents' mental access to the good.

The immediate upshot of the article is that GG theorists should move to assertoric formulations of the view, at least insofar as they rely on the intelligibility motivation. But in another way, the impact of the argument is broader. Near the end of the article, the dilemma I present gives way to a general argument that attitudinalist accounts of the intelligibility of action, whether they are GG views or not, fail precisely because action is intelligible to an agent in virtue of the *content* of the agent's mental states.

A subsidiary aim of this article is to articulate an account of the intelligibility motivation for GG, a motivation frequently referenced in the literature but rarely set out at length. I give a limited defense of it here, focusing exclusively on defending it from Setiya's objection. This I do in section 2 after introducing GG more thoroughly in section 1. Section 3 then presents the dilemma against attitudinal views.

1. THE GUISE OF THE GOOD

GG is really a family of views that all hold that human action or motivation to act, of some special kind or another, is only possible insofar as the agent acts or is motivated to act because of the *good* she sees in so acting.⁸ According to these views, goodness, or apparent goodness, plays a primary role in the motivation and rationalization of action.⁹ Before setting out the attitudinal and assertoric versions of the view, I flag four key features of the GG family.

GG has been held by various authors to be a constraint on action for a rea-

Force of Desire”), though he uses it to qualify a kind of Fregean force that he holds explains the rational significance of perception and belief.

7 Setiya, “Sympathy for the Devil.” Derek Baker (“The Abductive Case for Humeanism over Quasi-Perceptual Theories of Desire”) has also pressed a version of this objection.

8 See Tenenbaum, “Guise of the Good,” and Orsi, “The Guise of the Good,” for recent overviews.

9 See for instance Raz, “Reasons,” on what he calls the “normative/explanatory nexus.”

son, intention, intentional action, will, or, chiefly, desire.¹⁰ Because my concern is with GG theories generally, and because it is reasonable to assume that GG constrains action for a reason if it constrains intentional action or desire, here I investigate the weakest form of the view, one holding only that *action for a reason* must appear good to its agent.

There is a related dispute over the nature of the appearance of the good that the view requires. Many authors hold that desires *are* appearances of the good or that they involve or express thoughts about or perceptions of the good, but others hold that these required appearances are distinct from desires, such as motivating beliefs about the good.¹¹ I am skeptical that desires are appearances of the good, but the argument in this paper does not require me to take a side on this issue. My focus is on the form and content of these allegedly required, motivating appearances of the good, whatever they may be. For that reason I speak generically of *motivation* as an appearance of the good. GG theorists can substitute for “motivation” their favored appearance-state, be it desire, evaluative belief, etc.

GG requires more than that when agents act for a reason they see that act as good. It further requires that they act *because* of the (apparent) good, or that they are guided by the good.¹² Similarly, GG requires more than that action for a reason is guided by the good in some way or another, for it entails that agents who act for a reason are guided by their mental access to the good. This point differentiates GG from *constitutivism*, which attributes necessary aims or standards to agency but which is not generally committed to epistemic claims about agents’ access to those standards.¹³ An implication of this same point is that GG is not

10 As a constraint on action for a reason, see Raz, “Agency, Reason, and the Good,” 22–23. As a constraint on intentional action, intention, or will, see Aquinas, *Summa Theologiae*, I–II 1.1, 94.2; Davidson, “How Is Weakness of the Will Possible?” 22; Raz, “Agency, Reason, and the Good”; Buss, “What Practical Reasoning Must Be If We Act for Our Own Reasons,” 400; Anscombe *Intention*, 75. As a constraint on desire, see Stampe, “The Authority of Desire”; Oddie, *Value, Reality, and Desire*; Tenenbaum, *Appearances of the Good*; and many others.

11 Tenenbaum, *Appearances of the Good*; Quinn, “Putting Rationality in Its Place,” 36; Scanlon, *What We Owe to Each Other*, 37–41; Buss, “What Practical Reasoning Must Be If We Act for Our Own Reasons”; Gregory, “The Guise of Reasons.”

12 This “because” in the formulation of GG is intimately connected with that of reason-explanations of actions, but it should not be construed as entailing that, when an agent ϕ s for a reason, her reason for ϕ ing is *that ϕ ing is good*, as such. This would otherwise tend to beg the question against attitudinal views. For there is independent pressure to think that an agent’s (nonpsychological) reasons for acting are all to be found among the contents of her beliefs, and as we will see below, attitudinalists believe that the good need not figure in the content of an agent’s mental states.

13 Paradigm constitutivists include Velleman, “The Possibility of Practical Reason”; Korsgaard,

equivalent to the claim that agents, in acting for a reason, are guided by a formal aim of acting or practical reasoning, even when conjoined with the claim that the good is that formal aim.¹⁴ Indeed, it is possible for an agent to be guided by something of which they have no intimation, as when unnoticed environmental cues affect an agent's choice.¹⁵ We can also easily imagine a constitutivist theory that holds that agents are guided by the good but not in virtue of their access to it: suppose it is a conceptual truth that agents are guided by certain dispositions, say, to realize their final desires and to know the world in which they live. These dispositions fix an intrinsic standard for agency, so that a perfect agent is one who exercises these dispositions perfectly. The theory could furthermore hold that states of affairs are good relative to an agent just in case, and because, that agent's perfect counterpart desires them.¹⁶ On this theory, it falls out of the concept of agency that agents are guided by what is good relative to them. However, this is not because of any epistemic access that agents may have to what is good relative to them, but merely because agents are guided by dispositions the perfect exercise of which *makes* something good relative to them. They are guided by the good but in a "blind" way, as a dynamical system is guided by an attractor.¹⁷

So let me reemphasize that GG is a double requirement: the agent who acts for a reason must be (a) *guided* by the good, and (b) so guided in virtue of her *access* to it. According to GG, the appearance of goodness is crucial to the generation of action for a reason. And were the agent to have no access to what guides them, the guidance would be blind, and not the agent's rational guidance, as action for a reason must manifest.¹⁸

Self-Constitution; and Smith, "A Constitutivist Theory of Reasons." Note that Velleman rejects GG while holding that agents are necessarily guided by representations of what *de re* justifies action, though they need not conceive them under that guise. See Velleman, "Introduction," 28n34. Silverstein, "Agency and Normative Self-Governance," gives an exceptionally lucid presentation of Velleman's view on this issue.

14 *Contra* Saemi, "Aiming at the Good."

15 The most well-known evidence of this is found in Nisbett and Wilson, "Telling More Than We Can Know."

16 The toy theory here is inspired by, but distinct from, the constitutivist theory in Smith, "A Constitutivist Theory of Reasons."

17 Some may take issue with describing such agents as guided by the good. Is it not more accurate to say that they are guided *toward* what is good by their dispositions? But the crucial point, I think, is that given the theory we could explain such an agent's action (say, an attempt to realize a final desire) by saying that they did it *because* it was good. This noncausal explanation already gives a deflationary sense in which an agent is guided by the good.

18 See Schapiro, "What Are Theories of Desire Theories Of?" 132–33, for an excellent articulation of this point.

Lastly, “good” here must not be presumed to refer to *value*, much less to good-for or to moral, public, or agent-neutral value; indeed, “Guise of Normative Force” or “Guise of Justificatory Force” would be more apt titles for the thesis, if less catchy. The GG family thus embraces “Guise of Ought” and “Guise of Reasons” theorists too.¹⁹

1.1. *The Attitudinal Model*

The attitudinal model is a long-popular way of understanding appearances of the good. Typically it is applied to desire as the motivation of choice. The point of departure for the attitudinal model is a supposed analogy between belief and truth on the one hand, and desire and the good on the other. The formal object of belief is truth, it is held, and furthermore to believe is necessarily to regard the propositional content of one’s belief *as true*. But truth cannot plausibly be held to necessarily be part of the propositional content of belief, as if regarding one’s belief as true were like regarding a project to be a success, since that would set off a problematic regress.²⁰ So, truth must relate to some aspect of belief apart from its content. Truth may, for instance, be held to figure in the intentional mode of the attitude of belief—that is, in the manner in which belief represents its propositional content: to believe that *P* is to represent-as-true that *P*.

According to the attitudinal model of GG, appearances of the good have the same relationship, at least in broad outline, with respect to the good. After all, when one is motivated to bring about that *P*, one is not motivated to bring about that *it be good that P*, but simply that *P*. From this it would appear that goodness need not be part of the content of motivation. So on the attitudinal model, motivation is held instead to be a primitive presenting-as-good attitude, or goodness is considered part of the form of motivation, or it is suggested that it is constitutive of a motivation to bring about that *P* that it aims to get it right as to whether *P* is good.²¹

1.2. *The Assertoric Model*

On the assertoric model of GG, an appearance of the good possesses goodness-including content that it presents with the rational significance of belief or

19 Schafer, “Perception and the Rational Force of Desire”; Massin, “Desires, Values and Norms”; Gregory, “The Guise of Reasons.”

20 I present the regress explicitly in section 3.1 below.

21 Saemi, “Aiming at the Good”; Stampe, “The Authority of Desire”; Tenenbaum, “Appearing Good”; Schafer, “Perception and the Rational Force of Desire”; Kriegel, *Brentano’s Philosophical System*; Velleman, “The Guise of the Good” (though Velleman goes on to reject GG).

perception.²² Two cognitive, and in a broad sense of the term, representational states can share content and yet differ with respect to rational significance. One's merely imagining that John Rawls is standing in the doorway would not even defeasibly rationalize the belief that John Rawls is standing in the doorway. However, a perceptual experience with the same content would. Thus perception and belief, but not imagination, share rational significance insofar as they both have the power to defeasibly license belief or inference. The thought is then that the motivations at the center of GG are to be understood in terms of beliefs or perceptions concerning good actions.

In the assertoric camp belong Raz, Buss, and Gregory, who all hold that action for a reason requires a belief about the good or about one's normative reasons.²³ Davidson holds that desires express evaluative judgments.²⁴ Oddie and Hawkins take desires to be perception-like experiences of the good.²⁵

The assertoric model is frequently rejected as over-intellectual, as it is alleged that it requires animals and small children, who clearly have desires and who arguably can act for reasons, to possess evaluative concepts and exercise them whenever they act for a reason.²⁶ I am optimistic that the problem is more apparent than real since motivation should be thought to have nonconceptual evaluative content. However, the present paper is not the place to pursue this idea.²⁷

There is also support for a hybrid model in the literature. This model agrees with the assertoric model that motivation can be associated with evaluative content, but disagrees with its assimilation of motivation to an evaluative perception or belief. A version of this model might hold that agents desire the good as such, or aim to do what is good. Talbot Brewer could be interpreted as falling into this camp.²⁸ At any rate, since the hybrid view appeals to evaluative content, in

22 Here I use "presentation" and its cognates in Brentano's sense as a merely contentful state that of itself entails no commitment to the truth, appropriateness, etc., of its content, and which can be taken up into an attitude that does entail some such commitment. See Brentano, *Psychology from an Empirical Standpoint*, 61–63.

23 Raz, "Agency, Reason, and the Good"; Buss, "What Practical Reasoning Must Be If We Act for Our Own Reasons"; Gregory, "The Guise of Reasons."

24 Davidson, "Intending," 86.

25 Oddie, *Value, Reality, and Desire*; Hawkins, "Desiring the Bad under the Guise of the Good."

26 See for instance Velleman, "The Guise of the Good"; Schroeder, "How Does the Good Appear to Us?"; Saemi, "The Guise of the Good and the Problem of Over-Intellectualism."

27 I investigate just such a view in Boswell, "Affect, Representation, and the Standards of Practical Reason," ch. 4. See also Oddie, *Value, Reality, and Desire*, 80; Hawkins, "Desiring the Bad under the Guise of the Good."

28 Brewer, "Three Dogmas of Desire," esp. 271–72.

principle it possesses the same content-based resources as the assertoric view to account for the intelligibility of action, so I will not discuss it separately here.

2. INTELLIGIBILITY

In this section I first present a brief tour of the notion of intelligibility with the aim of thereby rendering IC, as well as GG's explanation of it, *prima facie* plausible. I then argue that a proper understanding of intelligibility defuses an objection from Kieran Setiya.

I should note that both supporters and critics of the intelligibility motivation have, at times, targeted other notions of intelligibility than the one I sketch below. Most notably, although the notion of intelligibility at issue here owes a great deal to Anscombe, it owes yet more to Warren Quinn. Anscombe holds that statements or agents under a certain description (typically as *wanting* something) are intelligible or not absolutely, and the notion of intelligibility she has in mind seems to be analogous to that according to which "Colorless green ideas sleep furiously" is not intelligible.²⁹ Regrettably, considerations of space prevent me from giving this interpretation its due in this paper; suffice it to say that I think the more compelling version of the intelligibility motivation relies on intuitions not about what is interpretable as an agent, as Anscombe's does, but about our relationship to our own actions.

2.1. Intelligibility of an Action to an Agent

The notion of intelligibility at play is *specific* in that an action is intelligible to its agent only if the agent has some consideration in mind, generally a property thought to be instantiated by the action, in virtue of which it is intelligible to him. Take for instance Quinn's infamous Radio Man, who finds himself in a bizarre functional state that causes him to turn on any radio at hand—though he does not turn them on in order to hear anything, or indeed in order that anything else happen. It is just something that he is disposed to do.³⁰ What is most conspicuously lacking in Radio Man is an idea of what is to be accomplished by his

29 See, for example, Anscombe, *Intention*, 14, 19, and 26; cf. MacIntyre, "The Intelligibility of Action," 64. Vogler (*Reasonably Vicious*, 49, 51) and Raz ("The Guise of the Bad," 11–13) seem to rely on Anscombe's notion while Raz ("Incommensurability and Agency," 118) and Tenenbaum (*Appearances of the Good*, 33–35) seem to avail themselves of both Quinn's and Anscombe's notions of intelligibility.

30 Quinn, "Putting Rationality in Its Place," 32.

turning on radios—what the point of it is.³¹ But what reason for acting could he have if he has *no idea* what the *point* of the action might be?³²

Because it concerns the sense the agent made of her action, the notion of intelligibility is also *agent-relative* and *actual*: IC requires of any action for a reason that *its agent actually saw*—that is, actually had mental access to—some point to the action. In this sense, the action is not intelligible in virtue of falling under some type that would be intelligible to someone else in other circumstances, nor in virtue of being intelligible in some absolute sense.

To be sure, an intelligibility requirement built solely on this idea—that in order to ϕ for a reason there must be some consideration the agent takes to be the

- 31 GG insiders may recall that Quinn (“Putting Rationality in Its Place”) focuses a little less on the *intelligibility* of action than on the thesis that desires must be evaluative if they are to *rationalize* action. Some philosophers have also taken desire’s rationalization of action to be a main explanatory payoff of GG. One might wonder why I focus on intelligibility, then. One difficulty is that “rationalization” is used in different ways: Is a rationalization (1) an explanation of an action in terms of the agent’s own reason for it (Davidson, “Actions, Reasons, and Causes,” 3)? Or does something rationalize action just in case it (2) makes that action rational, or alternatively, (3) grounds or “gives” the existence of a *pro tanto* or *prima facie* reason for it (Quinn, “Putting Rationality in Its Place,” 40)? All these interpretations pose problems for GG’s potential to explain how motivation rationalizes action. In brief, whether it is plausible to claim that only evaluatively construed motivations contribute to the rationality of action, as on (2), depends on the correct theory of rationality, and as for (3), many GG theorists hold that desires *track* reasons but do not *ground* them. For (1), GG can indeed explain why the agent does not act on a reason of her own when her action does not appear good to her, but it does this precisely by using GG to explain the *unintelligibility* of such action, as in Quinn’s article. So I regard the intelligibility requirement, and GG’s explanation of it, as more fundamental than and key to GG’s explanation of how desire (or certain motivations) rationalizes action.
- 32 A referee pointed out that it is not at all clear that Quinn accepts IC. After all, does Radio Man not walk across the room for a reason—so as to turn on the radio—in spite of the unintelligibility of doing so? In a sense the criticism is fair. Quinn is interested not in the conditions for an action’s being done for a reason but for its having a *rationalization*, where “rationalization” is interpreted as (3) and not (1) (see previous note), and so there are not sufficient grounds for counting Quinn as endorsing IC. On the other hand, nor are there sufficient grounds for counting Quinn as rejecting it. That would require Quinn to hold that one can act for a reason without having an attitude that gives one’s reason for acting, since for Quinn an attitude’s giving a reason for acting entails the existence of that reason, and Quinn also holds that there is no reason for Radio Man to take *means* to turn on radios (“Putting Rationality in Its Place,” 32–33). This in turn undercuts the necessary connection between practical thought and the good that Quinn is trying to draw (“Putting Rationality in Its Place,” 29–30). A more coherent position for Quinn would be that all action for a reason is for a (at least *prima facie*) reason that rationalizes that action, but that mere goal-directed action is not necessarily done for a reason. Regardless, this exegetical point is not crucial to the argument of the paper; the goal of this section is to draw out the intelligibility motivation as it has generally been understood.

case and in light of which ϕ ing is intelligible to them—is neither controversial nor interesting.³³ Quite plausibly it falls out of the very concept of acting for a reason that one have at least *some reason* in mind in acting. But the special interest in the GG theorist's conception of intelligibility is that not just any feature, consequence, or aspect thought by the agent to be instantiated by the action suffices to make that action intelligible to the agent:

Suppose, for example, that you notice me spray painting my shoe. You ask why I am doing that, and I reply that this way my left shoe will weigh a little more than my right. You ask why I want the left shoe to weigh a little more. Now suppose I just look at you blankly and say, "That's it." I seem not to understand your puzzlement. You grasp for straws. "Is this some sort of performance art, on the theme of asymmetry?" "No." "Is someone going to weigh your shoes as part of some game?" "No. Why do you ask?"³⁴

Here it is clear that there is an oddness about the Shoe Painter's explanation even though they have correctly identified a plausible consequence of spray painting their shoe.

Thus there appear to be substantive constraints on intelligible-making properties. *Being instrumental to something one is compelled to do* does not make an action intelligible. Nor does thinking of one's action as possessing a thin normative property, such as being what one *should* do as such, render it intelligible, as Dave's example shows.³⁵ So if indeed action for a reason must be intelligible to its agent, what explains why such action can occur under the guise of its instantiating certain properties and not others? GG proposes just such an explanation: the property must be, as Anscombe puts it, "one of the many forms of good," i.e., a particular or substantive evaluative property or kind of good.³⁶ Thus we arrive at GG's account of the intelligibility of action: an action is intelligible to its agent

33 See Schapiro, "The Nature of Inclination," 235–59, for a longer, nuanced discussion on this point. Schapiro takes Quinn's argument to ultimately show that thinking of desires as dispositions neglects their crucial role as proposing courses of action.

34 Clark, "Aspects, Guises, Species, and Knowing Something to Be Good," 234–35.

35 I consider an objection to this interpretation in section 2.2 below.

36 Anscombe, *Intention*, 77. For echoes of Anscombe on this point, see Quinn; "Putting Rationality in Its Place," 41–42; Johnston, "The Authority of Affect," 189–90; and Vogler, *Reasonably Vicious*, 51. For the sake of simplicity, I write throughout of specific *evaluative* properties or kinds of *good*. Guise of Reasons and Guise of Ought theorists need to find equivalent notions. However, these are ready at hand. Just as it is plausible that there are distinct kinds of value, it is also plausible that there are distinct kinds of reasons and oughts, such as moral, prudential, and aesthetic.

just in case, and because, it seems to the agent to instantiate or promote a substantive value-property. Together with GG's claim that an action done for a reason must be seen by its agent as good, this account of intelligibility can explain why IC is true—that is, why action for a reason must be intelligible to its agent.

Before moving on to Setiya's objection, let me address a worry about this understanding of the relation between GG and the intelligibility motivation. There are two parts to this explanation: the first is the account of the intelligibility of action in terms of seeing that action as good, and the second is the claim that this account and GG together can explain why IC is true. One could imagine both steps going differently, however. After all, GG is compatible with constructivist approaches that explain the good in terms of features of rational agency, or indeed with a no-priority view about the relation between rational faculties and the good.³⁷ Since it is furthermore plausible to hold that finding an action intelligible belongs to rational agency and is an aspect of our rational faculties, it would seem possible that a GG theorist could attempt to explain what it is for an agent to see her action as good in terms of what she finds intelligible, or to hold that necessarily an agent sees as good just what she finds intelligible, without assigning explanatory priority to either notion. Such a theorist could then deploy such an account, together with IC, to explain GG. On this way of understanding the problem, we begin with the question, "Why must actions done for a reason be seen as good by their agents?" and the response is that to see an action as good just is to find it intelligible, and action for a reason must be intelligible to its agent.³⁸

But the affinity between constructivism and the seeing-good-to-intelligibility direction of explanation is more apparent than real. It is one thing to explain *the good* in terms of the possibility of finding an action an intelligible object of choice, and another to explain *seeing* something *as good* in terms of its being made an intelligible object of choice. At any rate, whatever the GG theorist's account of the relation between rationality and the good, it seems to me that she faces considerable pressure to understand the intelligibility of action in terms of seeing good in the action. First, an explanation in the other direction threatens to reduce away the normative guise that is essential to GG: if the GG theorist tells us that agents must see as good all actions they do for a reason, but then clarifies that by "see as good" she means "find intelligible" or "make sense of," then it is not clear that she has told us anything particularly controversial. Even as

37 See Tenenbaum, *Appearances of the Good*, ch. 3, for a Kantian view that is at least consistent with constructivism, and Raz, "Reason, Rationality, and Normativity," 86, for the latter, no-priority view.

38 Thanks to a referee for pressing me to address a possibility along these lines.

staunch a GG opponent as Setiya could in principle adopt this theory by in turn understanding the intelligibility of action in terms of the agent's having a causal-historical explanation for it.³⁹ Thus, this direction of explanation threatens GG with triviality.⁴⁰ And second, the no-priority view seems to me an option of last resort, insofar as it posits an unexplained necessary connection between seeing good and finding intelligible.⁴¹

2.2. *Setiya's Objection*

Once we clarify that seeing an action as good can render it intelligible only when an agent has a *particular* kind of good in mind, we can neatly sidestep an otherwise worrisome objection Kieran Setiya first raised to GG's explanation of IC. The objection centers around the fact that merely noting that someone's action was performed under the guise of the good does not suffice to make the action intelligible. Setiya's version of the objection alludes to a famous line in Raz, which is intended to illustrate the idea that rules of rationality constrain choice, and that one of these constraints is that one can only choose for what one takes to be a good reason: "I cannot choose to have coffee because I love Sophocles," he writes.⁴²

To this Setiya responds that merely conceiving one's options explicitly under the guise of the good need not make them any more intelligible, and he illustrates the point with the following dialogue:

"She is drinking coffee because she loves Sophocles."

"What? That makes no sense at all."

"Oh yes it does! She thinks it is a reason to drink coffee."⁴³

I think Setiya's dialogue is best taken as illustrating just what Dave's story does: it sure seems possible—indeed, actual, given Dave—for someone to think that an action would be good, or what she ought to do, or what she has reason to do, etc., and yet for this not to make the action any more intelligible. But, given section 2.1, the GG theorist can plausibly reply that what has gone wrong here is that we

39 Setiya, *Reasons without Rationalism*.

40 I discuss a related triviality worry for GG in greater detail in Boswell, "Affect, Representation, and the Standards of Practical Reason," 31–37.

41 Once we have established that the GG theorist should explain intelligibility in terms of seeing good, we still face the question of whether we might instead see this account together with IC as supporting GG. However, the choice between the two is a matter of rhetorical strategy and would not affect the central conclusions of the paper.

42 Raz, "When We Are Ourselves," 8.

43 Setiya, "Sympathy for the Devil," 97.

are not given any kind of good under whose guise the coffee drinker drinks. If we were told that she is drinking coffee because she thinks this *honors* Sophocles, we may be confused as to why she would think that. But we would then appreciate her goal of honoring a literary figure, and thus find it intelligible as she does.⁴⁴ Thus, Setiya's objection does not undermine the intelligibility motivation for GG, once this motivation is properly understood.

Derek Baker has recently pressed a version of the same objection to views of desire as evaluative perception-like seemings, and it fails for the same reason.⁴⁵ Taking up a Radio Man-like example from Johnston, he asks us to imagine someone who acts on a blind urge to turn a knob counterclockwise.⁴⁶ Baker grants that the action will strike its agent as unintelligible, especially if she believes it not good to turn it, but he then supposes the same agent now acts on a perception-like feeling that there is a reason to turn the knob, even though she knows this feeling to be inaccurate. Baker objects that this addition of a seeming of a reason does not make the turning any more intelligible. But of course, we can now see that this is not an effective test of the intelligibility motivation. For that we would need to suppose that the agent feels there to be a specific *kind* of reason to turn the knob, even though she disbelieves this feeling. So let us suppose instead that the agent desperately feels that by turning the knob she will be *free* of the dungeon in which she is trapped, though she knows the door is locked. Does *this* feeling, a misleading anticipation of freedom, make sense of her turning the knob? The answer is clearly yes. She can see a point to her turning, though the point goes beyond her own hope of fulfilling it.

So it appears that there are guises of the good such that acting under them does not suffice for action for a reason. Why might this be so? By returning to the groundwork of GG laid out in section 1, I think we can glimpse why acting merely under the guise of normative notions like "ought," "reason," "good," or "desirable" will not suffice. According to GG, there is a constitutive, formal aim of whatever motivations it holds are appearances of justificatory force: they function to register or reflect ways in which actions may be justified. Because GG holds that these appearances also guide one to act because of the justificatory

44 One might think that Setiya's objection fails for a different reason, that, quite generally, possessing the testimony that *P* need not make *P* intelligible to one. I may, for instance, accept that some mathematical theorem I do not understand is true simply because Professor Conway told me it was. Quite plausibly, this general truth relies on a distinct kind of intelligibility, that of simply not fully grasping a proposition. But in Setiya's dialogue we are to imagine that the interlocutor does understand the last line, and finds the purported reason unintelligible nonetheless.

45 Baker, "The Abductive Case for Humeanism over Quasi-Perceptual Theories of Desire," 20.

46 Johnston, "The Authority of Affect," 189–90.

force registered or reflected in them, it follows that action itself is constitutively regulated by justificatory force, at least when it is done for a reason. But one's reason for acting cannot reflect merely that the action meets this constitutive aim, any more than one's reason for believing a proposition can be merely that it is true. (Supposing, of course, that belief is constitutively regulated by the norm of truth.) A genuine reason must offer intelligible support, and intelligible support cannot be so directly question-begging. A reason for acting must reflect a particular way in which the constitutive aim of action for a reason is met, just as a genuine reason for believing must reflect a particular way in which the constitutive aim of belief is met.⁴⁷ Now, as we also saw in section 1, the notions of what one *ought* to do, what one has *reason* to do, and what is *good* to do, are simply different candidates for this formal aim of action. Their suitability to playing this role of the formal aim of action is thus precisely what prevents action merely under their guise from rendering action for a reason intelligible.

Let me turn to three objections to this strategy of response to Setiya. First, there is an alternative way of understanding the import of cases like Dave, the Shoe Painter, and drinking coffee for love of Sophocles. One might think that these cases do not show thin normative concepts themselves incapable of rendering action intelligible, and instead show only that their ability to render action intelligible is conditional on these concepts being *intelligibly applied* in a given case. This view holds that the Shoe Painter's action is unintelligible not because she saw no specific kind of good that painting her shoe would accomplish (taking for granted, of course, that making her shoe a little heavier is not of itself good), but because she could not intelligibly take making her shoe a little heavier as itself a reason for painting the shoe. The idea is that taking anything to be a reason for an action suffices to make that action intelligible, *so long as* the thing taken to be a reason can intelligibly be so taken. Indeed, there is one way of reading Raz's discussion of drinking coffee for the love of Sophocles on which this is the very moral that he offers: his point may be that there are substantive constraints on *what* could be taken to be a reason for choice or on *how* an action

47 One might object that there are many cases in which a good reason for believing *P* is simply that it seems to one that *P* is true; perhaps, indeed, we cannot rebut skeptical arguments without a like entitlement. Is this not a question-begging reason for belief? (Thanks to a referee for this objection.) But there are two ways to read the offered reason, that "it seems to one that *P* is true." On the first, less plausible reading, one's reason is simply that *P* is true, as it seems to one to be the case. That is directly question-begging, and not an intelligible reason for believing *P*. On the second, one's reason is that one has a seeming as of *P* being true. That is an intelligible reason for believing *P*, but it is also not directly question-begging insofar as it appeals to a state distinct from belief in *P* that may well confer justification on it.

may be seen to be good if it is to be done for a reason, and that on his view these constraints derive from a theory of value.⁴⁸

However, I fail to see how *making one's shoe a little bit heavier* could simply be conceptually ineligible to be taken as a reason for painting one's shoe. Indeed, the interlocutor in the Shoe Painter example suggests two contexts that could make this a perfectly intelligible reason for that action—and it bears noting that both contexts (a sports competition, an artistic piece) plausibly render the action intelligible by connecting the reason to a specific kind of good (winning a game, aesthetic value). The situation is even worse for the objector once we turn to other thin normative concepts. Is it really plausible to suppose that the unintelligibility of Dave's kneeling in the street in front of the oncoming car is due to the fact that it is simply unintelligible to think that one *should* kneel down in the street in front of an oncoming car? On the contrary, it is clear that we can make such a thought intelligible by connecting it to a kind of good, for instance by supposing that Dave thought kneeling before the car the only way to save the lives of five children tied down to the road farther ahead. The objector may after all be right that seeing a thin good in an action is conditionally sufficient for rendering the action intelligible, but the condition involves the agent's seeing a specific kind of good in the action.

Second, one might try to extend Setiya's objection by finding cases in which an agent does see some substantive kind of good in acting but also in which that action is not intelligible to them. To see the idea, suppose that a moderate form of judgment externalism is true—in particular, that moral judgments do not necessitate motivation. One might then think that *caring* about something, and not merely judging it good in some specific way, is what makes action intelligible.⁴⁹ Imagine a miser who always judges it *charitable* to give bonuses to his employees, but who is left entirely cold by the idea, and so—the objection runs—does not find the prospect of actually giving out a bonus intelligible. If that is possible, GG fails to give a sufficient condition for intelligibility.

To this the GG theorist should respond that GG is not committed to holding that every way of cognizing an action as good amounts to seeing the action as good in the sense relevant to intelligibility, just as GG is not committed to holding that any sort of motivation or action whatsoever occurs under the guise of the good.⁵⁰ And indeed, there seems to be an important difference between abstractly assenting to the proposition that something would be the charitable thing to do and really *getting, feeling, recognizing, or having a practical intuition of*

48 Thanks to a referee for suggesting a response along these lines.

49 See Stocker, "Desiring the Bad" and "Raz on the Intelligibility of Bad Acts."

50 Tenenbaum, *Appearances of the Good*, 73–75, effectively makes the same point.

its value, though different GG theories will have different ways of cashing this thought out.⁵¹ So a GG theorist could plausibly respond that the miser did not really *see* a substantive kind of value in bonus-giving in the relevant sense of “see,” and that is why it was not intelligible to him.⁵²

Third, one might object that finding oneself with the inexplicable intuition that drinking coffee somehow *honors* Sophocles hardly makes more sense of drinking coffee than the bare intuition that one *ought* to drink it. What does the kind of good add?⁵³

I agree that thinking that drinking coffee honors Sophocles, absent further beliefs that rationalize this belief, is odd. But the oddness is that of having an odd theory about honor, and for the resolution of our puzzlement we require only some story of the agent’s history that shows her to have arrived at this theory through familiar psychological mechanisms. This puzzlement contrasts with that about someone who thinks that some action just *ought* to be done, without any further idea as to its import—even supposing that they point to the promotion of some other nonnormatively described property as their claimed reason for acting. (“I am placing this book here because that way I will have seventeen green books placed horizontally on my roof, which seems like the number I ought to have.”) Any historical explanation of this state of mind would still leave us saying, “That alone can’t be his reason for acting, if he really has one.” We cannot see how his action could be intelligible *to him*, absent some further story of how that action appeared to him to be related to some further, specific kind of good.

3. AGAINST THE ATTITUDINAL MODEL

Note that the assertoric model is well-positioned to account for the sense in which the intelligibility of action requires an appeal to specific evaluative properties. This is simply because it can hold that whatever motivations are appear-

51 See Johnston, “The Authority of Affect,” 192–93, for an illustration of how what he calls “affective collapse” can lead an ongoing activity to resemble mere habit or automatic action, even if one still believes, in an abstract, non-affective way, that the activity is valuable.

52 Of course, other responses to this challenge are possible. One could deny that the miser ever fails to find bonus-giving intelligible, and that what changes is merely his motivation to do it. Tenenbaum (*Appearances of the Good*) argues against judgment externalism (as part of what he calls the “separatist thesis”), and so can deny the coherence of the supposition that the miser really believed bonus-giving good but did not find bonus-giving intelligible at all. Roughly, Tenenbaum’s view implies that the miser finds bonus-giving *only conditionally* intelligible (*Appearances of the Good*, ch. 8).

53 Thanks to Stephen White for impressing on me the importance of this objection.

ances of the good must have some such evaluative property as part of their representational content, with no expectation that the same good is part of the content of every motivation. Take, for example, the view that emotions are appearances of the good with evaluative content. This view can hold that fear presents approaching a growling dog as dangerous while anger presents retaliation as payback. Dangerousness and payback are just different values, and they make different sorts of actions intelligible.

In contrast, according to the attitudinal model it is constitutive of motivation (of whatever sort GG is concerned with) to ϕ that it presents-as-good ϕ ing, in much the same way that belief presents-as-true its content. Generally speaking, proponents of this version of GG hold that the motivation in question is *desire*, so for ease of exposition in discussing their view, below I frequently replace talk of motivation with talk of desire.

But it would appear that this reliance on a feature of presenting-as-good that is common to all desire is also what prevents the attitude model from accounting for the intelligibility of action. As we saw above, if GG is to explain the intelligibility of action, it must explain it by reference to the appearance to the agent of a *particular* kind of good or evaluative property. That was precisely what we needed to respond to Setiya's objection, above. Yet it is hard to see how the attitudinal model can deliver this. The attitudinal view cannot simply hold that we act under the guise of a *thin* notion of good, as in the case of a view that holds that any desire presents its content as choiceworthy, desirable, or what one should bring about, as such. For recall that if this were so then the attitudinal model would secure the intelligibility of desired action only if Dave's action is intelligible to him. After all, it appeared to Dave that what he *should* do is kneel down and be hit by the car, and nevertheless his action was not intelligible to him. Clearly it makes the action no more intelligible if we substitute "is desirable for me to do" or "is the choiceworthy thing to do" for "should do."

The attitudinalist thus needs to thicken up her notion of good. There are a number of conceivable ways to accomplish this, but one attitudinalist already appears to possess the resources to deal with this problem. According to Sergio Tenenbaum, desires not only present their contents as meeting the formal aim of desire, but they present it *under a certain perspective*.⁵⁴ The perspective under which one desires is then to explain the particular way in which one's desiring is intelligible. The bare fact that Sue wants to damage Ms. S's boat does not intelligibly explain why Sue is throwing stones at it. But if we are told that Sue wants to damage Ms. S's boat *out of envy*, it appears we have been given an explana-

54 Tenenbaum, *Appearances of the Good*, sec. 1.5.

tion that shows her stone casting to be intelligible.⁵⁵ Other perspectives that can make certain actions intelligible include honesty, or being a cinephile, a gourmand, or a good parent—all of which can be plausibly understood as organized around specific kinds of good.

Here is a general version of the strategy at work. The attitudinal theorist is weakening the proposed analogy with truth in a certain respect. According to the story told in section 1, the connection between belief and truth is constitutive of and common to any belief, for all beliefs present-as-true their content. But perhaps token *desires* may present different substantive evaluative properties. Perhaps it is constitutive of desire that it present *some* good or other, even though no particular good may be required. Truth is one and good many, it might be said. Some such necessary connection between desires and specific evaluative properties is needed to overcome Setiya's objection, and it also seems to be what Tenenbaum is trying to secure.

Unfortunately, this connection is not ultimately available to the attitudinal theorist. I consider two formulations of this view. The first characterizes the connection between motivation and a specific evaluative property in terms of an adjective embedded within a verb: a token appetite *presents-as-tasty* a treat, say. The second, Tenenbaum's way, characterizes it adverbially: Sue desires *enviously*. Against the first I argue there is no obvious, good reason to think that it characterizes an aspect of an attitude as opposed to its content, and that there is reason to think the opposite. Against the second, I argue that it cannot secure the needed mental *access* to the specific good that GG requires. This leads me to advance in the final section a general argument against attitudinal views.

3.1. *The Adjectival Formulation*

What is an attitude? Orthodoxy has it that it is a relation between a subject and a content, usually a propositional content.⁵⁶ Most of the disagreement on this question is over the deeper nature of this relation. Staying at the level of commonly accepted platitudes, we could say that the content of a mental state gives *what* the mental state presents to the mind, what it is about, and the attitude provides *how* that content is presented, or the way in which the subject takes that content or that presentation.

So when the attitudinal theorist tells us that the desire for a treat presents-as-tasty the treat, should we take her at her word that she has characterized an aspect of the *attitude* of desire? Against this, note that nearly all the substantive, intelligible-making goods, such as health, success, beauty, and tastiness, are

55 Tenenbaum, *Appearances of the Good*, 43 (Tenenbaum's example).

56 E.g., Fodor, "Propositional Attitudes," generalizing slightly.

properties of the things, states of affairs, and people that our motivations are concerned *about*. They are not properties of representational entities like propositions. If your appetite for the treat presents-as-tasty the treat, then it seems your attitude can also be expressed by saying that you take it as true that the treat is tasty. You in some way attribute tastiness to the treat—that is why it is intelligible to you to eat it. And of course, we would naturally think that “tasty” figures in specifying the content of the attitude when expressed *this way*. But it is distinctive of attitudinal views that the good is not supposed to play a role in motivation by figuring in the content of a mental state.

These reflections point to a sufficient condition for construing a property *P* as figuring in the content of an attitude-type *A*:

AC: If attitude *A* presents-*X*-as-*P*, where (in a token case) *X* is what *A* is about and *P* is a property that *X* could instantiate, and *P* is not exclusively a property of a representational entity like a proposition, then *A*'s content (in that token case) is that *X* is *P*.

This principle has the great benefit of not misconstruing the contents of beliefs and perceptions, contrary to an accusation that Karl Schafer has brought against assertoric views.⁵⁷ Here is one way to see the worry, and how AC avoids it. As noted above (section 1.1), beliefs present or regard their contents as true, but the mere fact that a belief that *X* presents *X* as true cannot suffice for the content of that belief being (*X* is true). If that were true, the belief that *X* would also be the belief that (*X* is true), since the content of a belief is what follows the complementizer “that” in “the belief that.” But as a belief, it would also present (*X* is true) as true. Thus a regress begins. So it cannot be a general truth about attitudes that, if an attitude presents-*X*-as-*P*, then *P* is part of its content. From this it would appear that appealing to the fact that motivation presents an action as good in support of an assertoric view is simply special pleading.

AC does not entail that truth is part of the content of belief, however, for the simple reason that if belief presents *X* as true, then *X* must be a proposition, and a proposition is a representational entity. Nor are propositions what most beliefs are about; they are about taxes, fish, and all other things under the sun, most of which are not truth-apt.

AC also captures why the formal objects of a number of attitudes are best construed as part of their content. Fear of a dog presents the dog as dangerous, and some find it natural to say that the content of this fear is *that the dog is dangerous*; AC explains why.⁵⁸ Or, take the class of beliefs about people's hairiness,

57 Schafer, “Perception and the Rational Force of Desire,” 269.

58 Those most inclined to find this intuitive are emotional perceptualists like Tappolet (*Emo-*

and suppose that this class forms a type of attitude. We may then note that my belief that Steve is hairy presents-as-hairy Steve. But AC explains why we should not construe attitudes in this class in terms of a primitive presenting-as-hairy attitude that takes people as contents. “Hairy” qualifies Steve, not the manner in which he is represented to me.⁵⁹

AC has the effect of limiting what properties the attitudinalist may substitute for *P* in an attitude *A* that presents-*X*-as-*P* to those that either cannot be instantiated by whatever the attitude is about, or those that, like accuracy and the alethic modes, are exclusively properties of representational entities. This might seem unnaturally disjunctive. It may also seem that the latter restriction in particular hides the bone of contention, since what is at issue is whether normative properties such as the good—which surely do not exclusively apply to representational entities—can ever be substituted for *P* without figuring in the content of *A*. But there is, I believe, an attractive, general, and altogether orthodox picture of the relation between attitudes and their content that supports AC. According to this picture, the content of an attitude “gives” what the attitude is about by *representing* what the attitude is about, and when *A* presents *X* as *P*, *P* quite generally figures in *A*’s content—*except* where *X* is the content of *A*, and *P* is taken as modifying the manner in which that content, *qua* representation, is held before the mind. But where *P* functions to modify what the attitude is about (taxes, fish, etc.), it is difficult to see how *P* could do so without figuring in the content of the attitude itself, at least so long as we hold on to our platitudes about the distinction between attitude and content. Alternatively, the attitudinal theorist could reject the assumption that contents of attitudes are representational, but this seems to me to be extraordinarily costly.⁶⁰ Thus, to reject AC the attitudinalist would need to construct a different conception of the relation between attitudes and their content.

The attitudinalist may at this point object that there is a difference between

tions, Values, and Agency) and cognitivists like Solomon (“Emotions and Choice”) and Nussbaum (*Upheavals of Thought*). Emotional attitudinalists deny this, however; see Deonna and Teroni, *The Emotions*.

- 59 It must be granted that AC’s implications on other attitudes, particularly attitudes of confidence and doubt, are less clear. A suspicion that *X* presents *X* as somewhat likely. But does it present an *event* as likely, or does it present the proposition that *X* as likely true? I am inclined toward the latter, since the relation between probability and credence is broadly similar to that between outright belief and truth. But there is no space to pursue the issue here.
- 60 In brief, it seems to me that such views either cannot explain the failure of existential generalization from attitudinal contents or must rely on a mysterious relation to mere possibilia. But obviously, the debate on this question is long and intricate, and cannot be settled here.

an attitude that presents-*X-as-P* and the attitude of taking it to be true that *X* is *P*, for the former can be possessed by a creature without the concept of *P* while the latter cannot.⁶¹ But the principle this objection implicitly relies on—that one can only possess an attitude with content *C* if one possesses all the concepts required for a specification of *C*—is a rather contentious one insofar as it rejects the possibility of nonconceptual content.⁶² Moreover, the attitudinalist would owe us a justification for thinking that we *can* say that the creature presents-*X-as-P* even though she objects to saying that it takes it to be true that *X* is *P*. What could the difference be?

Now, in the foregoing I assumed that the “presents-as-good” locution is to be understood in terms of an *attribution* of a specific kind of goodness. With this assumption in place, AC is particularly intuitive. Some may seek to object to this assumption once it is seen that it leads to a rejection of the attitudinal view. However, if we do *not* understand the locution this way, it is hard to make sense of it, especially as an articulation of a GG view. What could it mean to take-as-beautiful a painting, if not to attribute beauty to it? What sort of manner of response to a content would that be? The attitudinalist cannot say that it means that one takes a certain kind of pleasure in the painting. For, to avoid circularity, she would then need to provide a non-attitudinal *evaluative* account of pleasure in order to secure GG. It is not clear to me how she could do this without ascribing evaluative content to pleasure, in which case the view would be a convoluted version of the assertoric strategy. Perhaps then it means responding to the painting in whatever manner is appropriate to a beautiful object.⁶³ But why would that manner necessarily imply that the subject has *mental access* to the relevant good, as GG requires?⁶⁴ Lastly, if the adjectival locution is not understood as implying an attribution of goodness, it does not seem to allow the possibility of motivation’s being mistaken about the good. Yet the ability of GG to secure the

61 Thanks to Christine Tappolet for this objection.

62 Nonconceptual content is frequently attributed to perceptual experiences and emotions. For an overview, see Bermúdez, “The Distinction between Conceptual and Nonconceptual Content.”

63 As noted above, many attitudinal theorists take the presenting-as-good character of desire to be primitive. The responses just considered are unavailable to these theorists, who are unable to explain this character in more basic terms and thus face extraordinary difficulty explaining why “good” should be thought to characterize an aspect of an attitude apart from its content.

64 Note the contrast with belief here: it is not at all clear that belief involves mental access to the *truth* of a proposition over and above access to the proposition itself. In believing, one is generally guided by defeasibly justified inferences, but this guidance need not take the form of tracking appearances of truth.

sense in which one's motivations can be mistaken is often considered one of the explanatory benefits of GG.⁶⁵

Thus this first version of the attitudinal view of GG could indeed make sense of how the appearance of specific kinds of good makes actions intelligible—but it can do this only by surreptitiously dropping the attitudinal view and taking up the assertoric view. Briefly put, if a mental state of presenting-*X*-as-good is to make action intelligible, it needs to be understood as attributing goodness to *X*, and it can only do this if goodness figures in the content of the state.

As we see in the next section, the second version of the attitudinal view faces a different but complementary problem: it cannot account for the way in which acting under the guise of the good involves the agent's mental access to the good.

3.2. *The Adverbial Formulation*

On Tenenbaum's view, Sue's desire to throw rocks at Ms. S's boat has an adverbial characterization: she desires *enviously* to throw rocks at it. Judging by Tenenbaum's development of the view, it is clear that this is a refinement of the attitude strategy. In order to act out of envy Sue need not have the explicit aim of acting out of envy; envy need not enter into the content of her mental states. She does not necessarily desire *to be envious*. Rather the envy is held to figure in the *manner* in which the world and practical possibilities appear to her. She *enviously desires* the destruction of the boat, and this manner of appearance is made manifest in the options for action she takes seriously, her irritability toward praise of the boat, the comments she makes about Ms. S, etc.

However, it turns out that this maneuver does less good for the attitudinal theorist than she might have hoped. Clearly there are such modes of acting and desiring, and we can appeal to them in order to give third-personal rationalizing explanations of agents' actions. But to say that an agent acted from a certain perspective is not always to explain the action in terms of the agent's point for that action. Sue *might* be self-consciously envious, but it is also possible that she is completely ignorant of her own enviousness. Her conscious motivations may have been limited to thoughts that Ms. S was fundamentally at fault, and that she needed to be taken down a peg. One can easily imagine a friend confronting Sue about her enviousness and Sue, upon realizing the truth about herself, coming to terms for the first time with her own envy. This reckoning would be rather like Freudian Dave's discovery in psychotherapy that he was motivated to lie down in the street by his repressed fear of failure: both agents may be said to have learned about their reasons for acting, but not in the sense of "reasons

65 See Tenenbaum, *Appearances of the Good*, as well as Baker, "The Abductive Case for Humeanism over Quasi-Perceptual Theories of Desire," for critical discussions.

for acting” we are looking for, since in both cases the motivation (envy or fear) was opaque to the agent.⁶⁶ Thus, while it may be that when one acts for a reason one acts under a perspective on the good, this latter cannot by itself explain the intelligibility of action for a reason.

Tenenbaum’s best reply might be to make a tactical retreat. Conceivably he could admit that desiring or acting under an *unconscious* perspective like Freudian Dave’s fear of failure does not make an action intelligible to its agent, but *conscious* perspectives do. The view is tempting, but it only reveals the fatal flaw with the attitudinal approach. For what could consciousness provide to intelligibility except further content? To hold otherwise would amount to an extreme kind of adverbialist treatment of perspectives, in the sense derived from the philosophy of perception.⁶⁷ Perhaps to be motivated consciously under the perspective of courage is, in being motivated, to be presented-to-courageously, where *being presented-to-courageously* is a non-intentional modification of a subject. But the absence of any relation to an intentional object or content of experience—say, to a desired action that might be courageous if one were to perform it—is precisely what renders this view incapable of capturing how awareness of specific evaluative properties renders action intelligible.⁶⁸

3.3. Content Makes Actions Intelligible

We have come at last to the fundamental reason attitudinal theories fail to capture the intelligibility of action, which is that intelligibility requires relation to a

66 Indeed, Tenenbaum seems to acknowledge that the perspective under which one acts can sometimes be opaque to the agent; see Tenenbaum, *Appearances of the Good*, 50.

67 Ducasse, “Moore’s Refutation of Idealism,” 252–53; Chisholm, *Perceiving*.

68 Dokic and Lemaire (“Are Emotions Evaluative Modes?”) raise a similar problem for the view that emotions are evaluative modes. This point against the attitudinal theory mirrors an influential critique of perceptual adverbialism, the many properties problem, which holds that this latter cannot account for the intentional structure of experience (Jackson, *Perception*). For a somewhat closer analogy, see Martin (“Setting Things before the Mind”), who argues that adverbialism cannot account for the fact that we are aware of our experiences through awareness of their intentional objects. However, there is at least one major reason to think that the present problem is much worse for the attitudinalist than these problems are for the adverbialist. Perceptual adverbialism is motivated by a desire to avoid reference to objectionable entities like sense data. But this motivation is not incompatible with attributing genuine intentional *content* to experience, for adverbialism can be interpreted as an alternative account of intentionality that still attributes content to experience, albeit in a deflationary way. (See Kriegel, “The Dispensability of (Merely) Intentional Objects,” for an example of this strategy.) The attitudinal version of GG, however, needs to avoid reliance on content in explaining the good’s contribution to intelligibility.

content, which in turn is precisely what attitudinal theories reject. We can make this thought explicit in the following argument:

1. IC: If an agent ϕ s for a reason, then ϕ ing is intelligible to her.
2. If an agent ϕ s for a reason, then ϕ ing is intelligible to her in virtue of a state of mental access to a (apparent) kind of good possessed by ϕ ing.
3. If ϕ ing is intelligible to an agent in virtue of a state S of mental access to something, then S is a state of consciousness of that thing.
4. If S is a state of consciousness of X , then X figures in the content of S .
5. So, some kind of good figures in the content of the state that makes action for a reason intelligible.

Premise 2 captures how GG explains IC: intelligibility is to be explained by a state of mental access to a kind of good. Freudian Dave's case is extremely good evidence that unless this state gives conscious access to a kind of good, an agent does not have the right kind of mental access to the (apparent) good of his action, and thus this state cannot explain the intelligibility of action. This justifies premise 3. Premise 4 is supported from what is quite generally taken as a starting point in the philosophy of mind: the content of consciousness just is whatever we are conscious of. From this we need to go only a little further and hold that the content of a state of consciousness of something is just that of which it is a state of consciousness.⁶⁹

From these premises, the conclusion in 5 follows. But 5 contradicts attitudinal views that aim also to account for the intelligibility of acting, according to which goodness need not figure in the content of desire. (Importantly, note that we could replace "kind of good" in this argument with a term denoting any property that is held to make action intelligible, and the resulting conclusion would still follow. Hence, the argument shows that attitudinalist-construed accounts of the intelligibility of action *in general* fail.)

The substantially new premises in this argument are 3 and 4. Because premise 4 is the more innocuous of the two, it will be best to concentrate on objections to premise 3.⁷⁰

69 It must be admitted that premise 4 is more contentious when certain varieties of *self-consciousness* are at issue. When I am struck with an awareness of a leg that is *mine*, or that my visual point of view differs from yours, must I figure in the content of such awareness? Fortunately, what is at issue for the present argument is consciousness of actions and their properties, not the self directly, nor even whose actions they are. With that granted, premise 4 can be understood as tacitly excepting self-consciousness and the argument will go through. (Thanks to Jens Gillissen for raising this issue.)

70 Of course, attitudinal theorists are free to define a notion of content on which premise 4 is false. However, this still leaves a very intuitive notion of content, expressed in that premise,

One thing to note about premise 3 is that at this point in the dialectic, the attitudinal theorist appears bound to accept it, for they need it to avoid predicting that Freudian Dan's kneeling is intelligible to him. Nevertheless, the attitudinal theorist might try one of two ways out of this corner.

The first is to object that premise 3 takes the perceptual analogy that GG invites far too seriously. Perhaps to say that a motivation presents its action as good just is to say that it has an action- or intention-rationalizing Fregean force.⁷¹ This implies nothing about the content of that motivation. Furthermore, to capture the way in which intelligibility is specific, we could postulate as above (section 3) that different motivations or perspectives rationalize different kinds of actions. Desiring under the perspective of courage rationalizes bold actions, say, while desiring under the perspective of envy rationalizes spiteful actions.

However, whatever the attractions of this view, it is not at all clear it is a GG view. GG explains the rationalizing power of motivation in terms of motivation's relation to an appearance of the good, but this alternate view characterizes appearances of the good in terms of rationalizing motivations. This threatens GG with triviality, since it is uncontroversial to claim that when one acts for a reason one acts on a motivation that rationalizes that action.⁷² Furthermore, the notion of having mental access to the good is purely honorific on this view, amounting to no more than being in a state that in fact rationalizes action. It is hard to see how this would entail an epistemic state—unless, of course, we then explained rationalization in terms of mental access to the good.

The second strategy is to search for less than fully conscious attitudes that nevertheless contribute non-instrumentally, and not in virtue of their content, to the intelligibility of action. The clause “non-instrumentally” is crucial, since the above argument (and specifically premise 2) is consistent with unconscious states being causally responsible for the intelligibility of action, for GG offers a theory of what intelligibility *consists in*. Suppose that my name is Woolstonecraft, and that at work I overhear two low voices plotting against Woolstonecraft. To save my skin, I immediately decide to foil their plans. My knowledge that I am Woolstonecraft is crucial to the intelligibility of foiling their plans, and I do not need to consciously rehearse it for it to have this effect. But it is not what *makes* foiling their plans intelligible to me, for that function is served by the rec-

for which it will be true that intelligibility requires goodness to figure in mental content. So for an attitudinalist who is determined to reject this premise, the burden will be on her to develop and defend the importance of this different notion.

71 Schafer, “Perception and the Rational Force of Desire,” 275–77.

72 Recall that we uncovered a similar triviality problem above (section 2.1) when discussing the possibility of understanding seeing an action as good in terms of finding it intelligible.

ognition that self-preservation demands that I foil their plans. My unconscious knowledge is simply a causal antecedent of that recognition.

So what kinds of non-conscious states could contribute non-instrumentally to the intelligibility of an action? Here are two salient possibilities. First, sometimes we explain what made sense to an agent by appealing to an entire worldview, a general *way* of conceiving things. It is hard to explain why suicide would seem the only honorable response to career failure to an ancient patrician Roman without appealing to moral assumptions that might have passed unnoticed in Roman times, but which come to light when we contrast them with our own moral assumptions. These assumptions may figure mentally not as content-bearing elements of one's belief box but in individuals' patterns of thought, motivation, and intention. However, these assumptions or patterns of response only *instrumentally* explain the intelligibility of suicide. They explain it by explaining why the Roman in dire straits would come to see suicide as honorable. This conscious recognition, in turn, would non-instrumentally explain the intelligibility of suicide.

The second possibility is to hold that while completely unconscious motivations do not make actions intelligible, except perhaps instrumentally, consciously accessible motivations do. After all, in executing a prior intention to ϕ we need not consciously rehearse our reasons for ϕ ing. When the time comes, we simply ϕ . Still we retain our reasons for ϕ ing, and so ϕ ing must be intelligible to us.

But if this is right, it does not show that the intelligibility of ϕ ing is not to be explained in terms of the *content* of a state. What enables ϕ ing in this case to be done for a reason is that one's reasons for ϕ ing are *consciously accessible*, either through being actually conscious, or being available to consciousness through attention or recall. After all, if one ϕ s for a reason in the full-blooded sense at issue in this paper, then one must have an answer to the question, "Why [did you do that]?" which gives one's reasons for acting, and it seems the only way one has such an answer is if one's reasons for acting are consciously accessible. Now, according to GG, motivation makes action intelligible by making it appear good in some respect. If it does so in virtue of the content of some state of appearing good when that state is conscious, then it seems that, if action can also be intelligible in virtue of such a state being consciously accessible but not accessed, then it is still the goodness-bearing content of that state that enables its contribution to action's intelligibility. For mental states do not massively change their content just for being consciously accessed.

In other words, the attitudinal theorist may complain that the sense in which GG requires mental access to the good requires only that the good of one's action be consciously accessible, not that it be actually accessed. But amending prem-

ise 2 to reflect this possibility would not help her, for what is at issue is only the accessibility of a state with goodness-bearing content. When an action is intelligible to an agent, it is still in virtue of the content of some of their mental states, contrary to what the attitudinal model holds.

4. CONCLUSION

The intelligibility constraint provides an attractive motivation for the Guise of the Good. Warren Quinn pointed out that without seeing something good about acting, action is not intelligible. It “fails for want of a point”—fails, that is, to have a reason that rationalizes it.⁷³ GG theorists are inclined to agree. But an important class of GG theorists, the attitudinalists, are unable to keep up with the commitments that ultimately derive from this basic insight. In order to reply to Setiya’s objection from apparent thin goods, these theorists need to appeal to a specific kind of good that makes action intelligible. But this leaves them with a dilemma: either they characterize the connection to the specific evaluative property in a way that makes an assertoric view far more plausible, or they fail to explain the agent’s mental access to the good, as GG requires.⁷⁴

Centre de recherche en éthique
Université de Montréal
paul.s.boswell@gmail.com

REFERENCES

- Anscombe, G. E. M. *Intention*. 2nd ed. Oxford: Blackwell, 1963.
- Baker, Derek. “The Abductive Case for Humeanism over Quasi-Perceptual Theories of Desire.” *Journal of Ethics and Social Philosophy* 8, no. 2 (October 2014): 1–29.
- Bermúdez, José Luis. “The Distinction between Conceptual and Nonconceptual Content.” In *The Oxford Handbook of Philosophy of Mind*, edited by Ansgar Beckermann, Brian P. McLaughlin, and Sven Walter, 457–73. New York: Oxford University Press, 2009.

73 Quinn, “Putting Rationality in Its Place,” 34.

74 Thanks to Samuel Asarnow, Daniel Drucker, Mauro Rossi, Christine Tappolet, Valéry Giroux, Kieran Setiya, Uriah Kriegel, Nandi Theunissen, Jens Gillissen, Angie Pepper, Sarah Buss, Stephen White, two referees at the *Journal of Ethics and Social Philosophy*, and audiences in Paris, Montreal, and Evanston for comments that greatly improved the paper.

- Boswell, Paul. "Affect, Representation, and the Standards of Practical Reason." PhD diss., University of Michigan, 2016.
- Boyle, Matthew, and Douglas Lavin. "Goodness and Desire." In Tenenbaum, *Desire, Practical Reason, and the Good*, 161–201.
- Brentano, Franz. *Psychology from an Empirical Standpoint*. 1874. Edited by Oskar Kraus and Linda L. McAlister. Translated by Antos C. Rancurello, D. B. Terrell, and Linda L. McAlister. London: Routledge, 1995.
- Brewer, Talbot. "Three Dogmas of Desire." In *Values and Virtues: Aristotelianism in Contemporary Ethics*, edited by Timothy Chappell, 257–84. Oxford: Oxford University Press, 2006.
- Buss, Sarah. "What Practical Reasoning Must Be If We Act for Our Own Reasons." *Australasian Journal of Philosophy* 77, no. 4 (December 1999): 399–421.
- Chisholm, Roderick M. *Perceiving: A Philosophical Study*. Ithaca, NY: Cornell University Press, 1957.
- Clark, Philip. "Aspects, Guises, Species, and Knowing Something to Be Good." In Tenenbaum, *Desire, Practical Reason, and the Good*, 234–44.
- Davidson, Donald. "Actions, Reasons, and Causes." In *Essays on Actions and Events*, 3–20.
- . *Essays on Actions and Events*. 2nd ed. Oxford: Oxford University Press, 2001.
- . "How Is Weakness of the Will Possible?" In *Essays on Action and Events*, 21–42.
- . "Intending." In *Essays on Actions and Events*, 83–102.
- Deonna, Julien A., and Fabrice Teroni. *The Emotions: A Philosophical Introduction*. London: Routledge, 2012.
- Dokic, Jérôme, and Stéphane Lemaire. "Are Emotions Evaluative Modes?" *Dialectica* 69, no. 3 (September 2015): 271–92.
- Ducasse, C. J. "Moore's Refutation of Idealism." In *The Philosophy of G. E. Moore*, edited by Paul A. Schilpp, 223–53. La Salle, IL: Open Court, 1942.
- Fodor, Jerry A. "Propositional Attitudes." *The Monist* 61 (October 1978): 501–23.
- Glass, Ira. "Devil on My Shoulder." *This American Life*. Podcast audio, May 24, 2002. <http://thisamericanlife.org/radio-archives/episode/213/transcript>.
- Gregory, Alex. "The Guise of Reasons." *American Philosophical Quarterly* 50, no. 1 (January 2013): 63–72.
- Hawkins, Jennifer. "Desiring the Bad under the Guise of the Good." *The Philosophical Quarterly* 58, no. 231 (April 2008): 244–64.
- Jackson, Frank. *Perception: A Representative Theory*. Cambridge: Cambridge University Press, 1977.

- Johnston, Mark. "The Authority of Affect." *Philosophy and Phenomenological Research* 63, no. 1 (July 2001): 181–214.
- Korsgaard, Christine M. *Self-Constitution: Agency, Identity, and Integrity*. New York: Oxford University Press, 2009.
- Kriegel, Uriah. *Brentano's Philosophical System: Mind, Being, Value*. Oxford: Oxford University Press, forthcoming.
- . "The Dispensability of (Merely) Intentional Objects." *Philosophical Studies* 141, no. 1 (October 2008): 79–95.
- MacIntyre, Alasdair. "The Intelligibility of Action." In *Rationality, Relativism, and the Human Sciences*, edited by Michael Krausz, Richard M. Burian, and Joseph Margolis, 63–80. Dordrecht: M. Nijhoff, 1986.
- Martin, Michael G. F. "Setting Things before the Mind." In *Contemporary Issues in the Philosophy of Mind*, edited by Anthony O'Hear, 157–80. Cambridge: Cambridge University Press, 1998.
- Massin, Olivier. "Desires, Values and Norms." In *The Nature of Desire*, edited by Federico Lauria and Julien Deonna, 165–200. Oxford: Oxford University Press, 2017.
- Nisbett, Richard E., and Timothy D. Wilson. "Telling More Than We Can Know: Verbal Reports on Mental Processes." *Psychological Review* 84, no. 3 (May 1977): 231–59.
- Nussbaum, Martha. *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press, 2001.
- Oddie, Graham. *Value, Reality, and Desire*. Oxford: Oxford University Press, 2005.
- Orsi, Francesco. "The Guise of the Good." *Philosophy Compass* 10, no. 10 (October 2015): 714–24.
- Quinn, Warren S. "Putting Rationality in Its Place." In *Value, Welfare, and Morality*, edited by Christopher W. Morris and R. G. Frey, 26–49. Cambridge: Cambridge University Press, 1993.
- Raz, Joseph. "Agency, Reason, and the Good." In *Engaging Reason*, 22–45.
- . *Engaging Reason: On the Theory of Value and Action*. New York: Oxford University Press, 1999.
- . "The Guise of the Bad." *Journal of Ethics and Social Philosophy* 10, no. 3 (July 2016): 1–14.
- . "Incommensurability and Agency." In *Incommensurability, Incomparability, and Practical Reason*, edited by Ruth Chang, 110–28. Cambridge, MA: Harvard University Press, 1997.
- . "On the Guise of the Good." In Tenenbaum, *Desire, Practical Reason, and the Good*, 111–37.

- . “Reason, Rationality, and Normativity.” In *From Normativity to Responsibility*, 85–101. Oxford: Oxford University Press, 2011.
- . “Reasons: Explanatory and Normative.” In *New Essays on the Explanation of Action*, edited by Constantine Sandis, 184–202. New York: Palgrave Macmillan, 2009.
- . “When We Are Ourselves: The Active and the Passive.” In *Engaging Reason*, 5–22.
- Saemi, Amir. “Aiming at the Good.” *Canadian Journal of Philosophy* 45, no. 2 (2015): 197–219.
- . “The Guise of the Good and the Problem of Over-Intellectualism.” *Journal of Value Inquiry* 48, no. 3 (September 2014): 489–501.
- Scanlon, Thomas M. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1998.
- Schafer, Karl. “Perception and the Rational Force of Desire.” *Journal of Philosophy* 110, no. 5 (May 2013): 258–81.
- Schapiro, Tamar. “The Nature of Inclination.” *Ethics* 119, no. 2 (January 2009): 229–56.
- . “What Are Theories of Desire Theories Of?” *Analytic Philosophy* 55, no. 2 (June 2009): 131–50.
- Schroeder, Mark. “How Does the Good Appear to Us?” *Social Theory and Practice* 34, no. 1 (January 2008): 119–30.
- Setiya, Kieran. *Reasons without Rationalism*. Princeton: Princeton University Press, 2007.
- . “Sympathy for the Devil.” In Tenenbaum, *Desire, Practical Reason, and the Good*, 82–110.
- Silverstein, Matthew. “Agency and Normative Self-Governance.” *Australasian Journal of Philosophy* 95, no. 3 (2016): 517–28.
- Smith, Michael. “A Constitutivist Theory of Reasons: Its Promise and Parts.” *Law, Ethics and Philosophy* 1 (2013): 9–30.
- Solomon, Robert C. “Emotions and Choice.” *Review of Metaphysics* 27, no. 1 (September 1973): 20–41.
- Stampe, Dennis W. “The Authority of Desire.” *Philosophical Review* 96, no. 3 (July 1987): 335–81.
- Stocker, Michael. “Desiring the Bad.” *Journal of Philosophy* 76, no. 2 (December 1979): 738–53.
- . “Raz on the Intelligibility of Bad Acts.” In *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, edited by R. Jay Wallace, Philip Pettit, Samuel Scheffler, and Michael Smith, 303–32. New York: Oxford University Press, 2004.

- Sussman, David. "For Badness' Sake." *Journal of Philosophy* 106, no. 11 (November 2009): 613–28.
- Tappolet, Christine. *Emotions, Values, and Agency*. Oxford: Oxford University Press, 2016.
- Tenenbaum, Sergio. *Appearances of the Good*. Cambridge: Cambridge University Press, 2007.
- . "Appearing Good: A Reply to Schroeder." *Social Theory and Practice* 34, no. 1 (January 2008): 131–38.
- , ed. *Desire, Practical Reason, and the Good*. New York: Oxford University Press, 2010.
- . "Guise of the Good." In *The International Encyclopedia of Ethics*, edited by Hugh LaFollette, 2262–71. Hoboken, NJ: Wiley, 2013.
- Velleman, J. David. "The Guise of the Good." In *The Possibility of Practical Reason*, 99–122.
- . "Introduction." In *The Possibility of Practical Reason*, 1–31.
- . *The Possibility of Practical Reason*. Oxford: Oxford University Press, New York, 2000.
- . "The Possibility of Practical Reason." In *The Possibility of Practical Reason*, 170–99.
- Vogler, Candace. *Reasonably Vicious*. Cambridge, MA: Harvard University Press, 2002.
- Wiland, Eric. *Reasons*. London: Continuum, 2012.

ETHICAL REDUCTIONISM

Neil Sinhababu

NATURALISTIC MORAL REALISTS hold that moral properties are part of the natural world.¹ They can accept either reductionism or nonreductionism about how moral properties relate to properties invoked in the best natural and social scientific explanations, which I call “scientific properties.”² This article argues that reductionism is the best form of naturalistic moral realism.

Reductionism and nonreductionism differ about whether moral properties and scientific properties are identical.³ Reductionists see moral properties as identical to individual scientific properties or disjunctions of scientific properties. Supposing for illustration that hedonism is the true theory of moral value, reductionism treats goodness as identical to pleasure, just as water is identical to H₂O.⁴ Nonreductionists see moral properties as natural properties supervening on and constituted by scientific properties without being identical to them.⁵

- 1 My characterization of the positions follows Darwall, Gibbard, and Railton, “Toward *Fin de Siècle* Ethics”; Miller, *Contemporary Metaethics*; and Shafer-Landau, *Moral Realism*.
- 2 “Scientific properties” refers to properties of physics, biology, psychology, and other natural and social sciences, but not moral properties unless they are identical to these properties. I know no better term. “Nonmoral properties” makes reductionism sound contradictory in claiming that moral properties are nonmoral, “natural properties” makes nonreductionism sound nonnaturalist in denying that moral properties are natural, and “descriptive properties” erects a false contrast, as realists regard moral language as descriptive. “Natural kinds” may be an equivalent term, though I do not know how broadly it is used this way. These positions are stated in terms of the abundant view of properties. The sparse view will be discussed shortly.
- 3 While reduction in some contexts does not entail identity, I defend property identity, which is part of strong-reductive theses. Schroeder regards reduction as property analysis rather than property identity (*Slaves of the Passions*). Our views are compatible, since he allows identity claims like the one defended here to fit within a property analysis.
- 4 More technically, “goodness is identical to *being* pleasure.” Following much of the metaethics literature, I usually omit the “being” and talk of property existence rather than instantiation.
- 5 While nonnaturalists like Shafer-Landau (*Moral Realism*) and Huemer (*Ethical Intuition-*

Again assuming hedonism for illustration, nonreductionism treats goodness as supervening on pleasure without identity. The Cornell Realists liken this to how psychological properties supervene on neuroscientific properties in Jerry Fodor's influential view of the special sciences.⁶

Both views share many features. They address the conceptual is/ought gap by agreeing with G. E. Moore that normative ethical truths are synthetic and not analytic.⁷ They reject his view that moral properties are nonnatural. They answer John Mackie's argument that moral properties are unacceptably queer by denying that they produce categorical reasons.⁸

Today, nonreductionism is the dominant form of synthetic naturalistic moral realism.⁹ Russ Shafer-Landau describes the consensus, writing of Richard Boyd's moral semantics:

Boyd himself does not believe that application of his theory will yield a reductive view . . . no one has supplied any reason for thinking that he has

ism) also hold that moral properties supervene on nonmoral properties without identity, they deny that moral properties are natural properties. This makes an intuitionist moral epistemology suit their views better, and may have other consequences for their metaethical theorizing.

- 6 Brink (*Moral Realism and the Foundations of Ethics*) provides the most comprehensive defense of nonreductionism, likening his position to Fodor's metaphysics of mind ("Special Sciences"). Sayre-McCord similarly defends "belief in two kinds of properties: those which can be reductively identified with explanatorily potent properties and those we have independent reason to think supervene upon, without being strictly reducible to, explanatorily potent properties" ("Moral Theory and Explanatory Impotence," 274). This view is often attributed to Sturgeon, "Moral Explanations," and Boyd, "How to Be a Moral Realist." Their metaethical approach is called "Cornell Realism" because Sturgeon, Boyd, and Brink were affiliated with Cornell University (whose press published Sayre-McCord's anthology). Boyd's and Sturgeon's criticisms of reductionism are less explicitly focused on synthetic reductionism. Boyd criticizes "the conclusion that all natural properties must be definable in the vocabulary of physics" ("How to Be a Moral Realist," 194), and Sturgeon repeatedly criticizes "reductive definitions" ("Moral Explanations," 240–43). Harman ("Moral Explanations of Natural Facts") notes that these remarks can be read merely as criticisms of analytic reductionism, as definition is primarily a semantic notion rather than an ontological notion. So I focus on Brink and Sayre-McCord here.
- 7 Moore, *Principia Ethica*. If Moore is wrong about moral semantics, the door is open for analytic reductionists like Jackson, *From Metaphysics to Ethics*, and Finlay, *Confusion of Tongues*.
- 8 Mackie, *Ethics*. For arguments favoring naturalists' rejection of categorical reasons, see Foot, "Morality as a System of Hypothetical Imperatives," and Svavarsdóttir, "Moral Cognitivism and Motivation."
- 9 Railton ("Moral Realism," "Naturalism and Prescriptivity") and Jackson (*From Metaphysics to Ethics*, "In Defense of Reductionism in Ethics") offer sympathetic treatments of reductionism. But they respond to anti-realists and nonnaturalists rather than naturalistic nonreductionists.

erred in this regard. No one has done anything towards showing that his semantics, when well applied, would yield the surprising conclusion that goodness (and rightness, and forbiddenness, etc.) is identical to some specific natural property.¹⁰

This consensus should be overturned. Naturalistic moral realists should accept identities between moral properties and scientific properties, even if they believe that properties of special sciences like psychology are not identical to properties of neuroscience or physics. If reduction is easier than Fodor allows, ethical reductionists may have more resources at their disposal than I use here.¹¹ But even if reductionism fails in psychology, it succeeds in ethics.

Ethical nonreductionists borrow two arguments from Fodor's philosophy of mind, which the two main sections of this article answer. First, nonreductionists argue that the multiple realizability of moral properties defeats reductionism. I solve multiple realizability in ethics by identifying moral properties uniquely or disjunctively with special science properties. This eliminates the main purported disadvantage of reductionism. Second, nonreductionists argue that irreducible moral properties explain empirical phenomena, just as irreducible special science properties do. But since irreducible moral properties do not succeed in explaining additional regularities, error theorists can rightly say that they are pseudoscientific. Since reductionism entails the existence of moral properties when combined with the existence of the reduction bases, it is the more defensible form of naturalistic moral realism.

In recent years, the popularity of nonnaturalistic realism has exceeded that of naturalistic realism.¹² The dialectical situation makes this unsurprising. The best-known versions of naturalistic realism are the Cornell Realists' nonreductionism with its dubious moral explanations, and Jackson's reductionism with its ties to a very different philosophy of mind than the one that the Cornell Realists invoke against their opponents. I show that, even on the philosophy of the special sciences the Cornell Realists assumed, reductionism can accommodate all existing normative ethical theories while avoiding dubious empirical commitments, making it the best form of naturalistic realism.

10 Shafer-Landau, *Moral Realism*, 68–69.

11 Against Fodor are Kim, "Multiple Realization and the Metaphysics of Reduction," and Sober, "The Multiple Realizability Argument against Reductionism." To see how disjunctive properties might provide reduction bases, see Clapp, "Disjunctive Properties."

12 Especially influential are Enoch, *Taking Morality Seriously*; Huemer, *Ethical Intuitionism*; Parfit, *On What Matters*; and Scanlon, *What We Owe to Each Other*.

1. MULTIPLE REALIZABILITY DOES NOT
ENDANGER ETHICAL REDUCTIONISM

This section responds to the argument that multiple realizability defeats reductionism. David Brink presents this argument clearly and uses it to motivate nonreductionism. I compare his position to Fodor's nonreductionism in philosophy of mind, and respond with a solution for multiple realizability in ethics: identifying moral properties with natural kinds from sciences like psychology and sociology, uniquely or disjunctively. This makes reductionism as good as nonreductionism for accommodating all currently defended normative ethical theories.

Brink invokes multiple realizability to defend nonreductionism, which he calls "constitutional materialism," against reductionism, which he calls "identity materialism":

There are what should by now be familiar reasons to prefer constitutional to identity materialism. If materialism is only contingently true, then higher-order properties, though actually physical properties, could have been realized nonphysically. If so, these higher-order properties are not necessarily physical properties and so cannot be identical with physical properties. Moreover, higher-order properties and property instances could have been realized in a variety of different physical ways. If we deny that identity is a relation that can hold between relata that are indefinitely or infinitely disjunctive, the multiple realizability of these higher-order properties provides reason to deny that they are identical with physical properties.¹³

I summarize Brink's two multiple realizability arguments against reductionism.

First, moral properties like wrongness could be realized even in worlds with different fundamental properties. In a nonphysical world, it would be wrong for ghosts to torture other ghosts. And in a world whose fundamental physics differs from that of our world, slavery would still be just as wrong. Since things could be wrong in worlds that do not have our physics, wrongness cannot be identical to properties of our physics. This argument parallels arguments against reducing mental properties to neural properties: since robots and aliens who lack humanlike brains can have beliefs, belief cannot be identical to anything from neuroscience.

Second, many different actual physical structures realize wrongness. Slavery, gender discrimination, and torture have little in common at the level of physics

13 Brink, *Moral Realism and the Foundations of Ethics*, 178.

that distinguishes them from things that are not wrong. Wrongness and belief are equally unlikely to have unified or even finite characterizations in the language of physics. Brink rejects identifying higher-level properties with infinitely disjunctive lower-level properties. This argument parallels arguments against reducing mental properties to properties of physics: belief cannot be identical to anything from physics, since it is constituted in such disunified ways at that level.

Brink's arguments share the structure of those against reductionism in the philosophy of mind.¹⁴ On these reductionist views, mental states like belief and desire are identical to states from physics or neuroscience.¹⁵ The problem for reductionism is that one could have the same mental state by having any of many physical structures. For example, the belief that philosophy is fun need not be realized by neurons. Robots and aliens could have the same belief, realized by silicon chips or whatever is in aliens' heads. Believing that philosophy is fun cannot be identical to anything neurological, biological, or physical, because creatures can do it by having any among an infinite disjunction of different neurological, biological, and physical structures.¹⁶ I call this form of multiple realizability "infinite realizability."

Brink's theory of how moral properties relate to lower-level properties is explicitly built on Fodor's model, which accommodates infinite realizability while maintaining an attractive physicalist thesis.¹⁷ On Fodor's view, each science is a

- 14 Fodor, "Explanations in Psychology"; Putnam, "The Nature of Mental States." Here I do not focus on a powerful nonreductionist argument from the material constitution debate that concerns modal properties of ordinary objects: the statue cannot survive squashing, while its constituent clay survives squashing, so the statue is not identical to the clay. Nonreductionism about statues can plausibly claim to explain why their modal properties differ from those of their constituent clay, but the parallel argument for metaethical nonreductionism fails. Removing the goodness of any state of affairs requires removing whatever property of the state of affairs made it good. So you cannot destroy the goodness without destroying the underlying property specified by the right theory of goodness. Paul provides a helpful discussion of material constitution ("The Puzzles of Material Constitution").
- 15 Place, "Is Consciousness a Brain Process?"; Feigl, "The 'Mental' and the 'Physical'"; Smart, "Sensations and Brain Processes." Bennett, "Mental Causation," provides an overview.
- 16 Here I discuss ontology in terms of properties rather than facts, following Fodor.
- 17 Brink cites Fodor's "Special Sciences" four times, the last two specifically to support ethical nonreductionism with multiple realizability arguments (*Moral Realism and the Foundations of Ethics*, 166, 167, 180, 194). No existing work has shown that Fodor's arguments do not carry over to moral properties as Brink thinks they do. Jackson's discussion of Cornell Realism (*From Metaphysics to Ethics*) does not distinguish reductionists like Railton from naturalistic nonreductionists like Brink and Sayre-McCord. His article with Pettit and Smith ("Ethical Particularism and Patterns") and his later work ("In Defense of Reductionism in Ethics") respond to Dancy, Parfit, and other nonnaturalists who do not share Fodor's metaphysics of natural and causally effective but irreducible properties.

separate layer of irreducible properties. Physics is the bottom layer, and properties like “belief” from the special-science layer of psychology need not be identical to any properties of physics or other sciences. Still, every actual belief is a physical thing. The language of physics allows a full characterization of every individual human, robot, and alien belief, even if it does not give us a well-unified general characterization to cover all of them. As Fodor concludes “Special Sciences”: “If physics is to be basic science, then each of these things had better be a physical thing. But it is not further required that the taxonomies which the special sciences employ must themselves reduce to the taxonomy of physics. It is not required, and it is probably not true.”¹⁸

Having presented these multiple realizability arguments, I explain how ethical reductionism answers them. Moral properties are reducible either to individual special science properties as water is reducible to H_2O , or to disjunctions of them as jade is reducible to jadeite or nephrite. A finitely disjunctive reduction base provides the flexibility to accommodate the most complex existing moral theories. Locating the reduction base at the special-science level allows finite realizers across physical and nonphysical worlds.

Normative ethical theories typically give accounts of the moral in terms of the psychological, social, or biological. On a reductionist metaethical treatment, the moral and scientific terms refer to the same properties. The easiest cases for reductionism are monistic theories like hedonism about moral value, on which something is good iff it is pleasure. This is a full account of a moral property—goodness—in terms of a psychological property—pleasure. While many metaethical views are open to hedonists, a natural one is the reductionist view that goodness is pleasure.

Of course, many moral theories are more complex, and do not seek to unify all of morality under one principle. These pluralist theories provide type-reductions of moral properties to disjunctive reduction bases. Moore takes aesthetic appreciation and friendship to have moral value that goes beyond the pleasure experienced. Reductionists treat his normative ethics as describing the combinations of psychological and sociological properties that make up aesthetic appreciation and friendship, and to which goodness is reducible. Goodness, on this view, is reducible to pleasure or aesthetic appreciation or friendship. To accommodate more good things, one simply needs more disjuncts.

One might object that disjunctive properties cannot provide reduction bases for higher-level properties, so reductionism will not accommodate more complex moral theories. Fodor claims that belief cannot be reduced to a big, disunified disjunction of neural, silicon, and alien states because it is a natural kind

18 Fodor, “Special Sciences,” 114.

that should provide unified explanations. Jade is not a natural-kind term in the special sciences because it is disjunctively realized by jadeite and nephrite. So jade is not a natural kind. Jadeite explains some geological phenomena while nephrite explains others, but their disjunction does not explain things. If we similarly cannot reduce moral properties to disjunctions of scientific properties, reductionism will be an option only for monistic moral theories like hedonism.

Moral properties, however, can have disjunctive reduction bases, since moral concepts do not require unified roles for moral properties in scientific explanations. So moral properties can be disjunctions of natural kinds rather than natural kinds themselves. Reductionists can let goodness be like jade—a real thing in the natural world that is not a natural kind. Denying that jade is a natural kind is not denying that jade exists. If your brother wants some phlogiston and your sister wants a jade necklace, only one of them must be disappointed. Just as those wanting jade necklaces need not be disappointed, those who want to make the world a better place need not be disappointed if goodness is a disjunction of natural kinds.¹⁹ Objective features of the natural world then satisfy the predicate “good,” making naturalistic moral realism true.

This disjunctive solution gives naturalistic moral realists what they care about, even on a sparse view of properties.²⁰ For ease of exposition, most of this article assumes an abundant view of properties on which there is a (nonscientific) property of being jade, since there are (scientific) properties of being jadeite and being nephrite. This abundant view makes disjunctions of scientific properties identical to moral properties. To show that reductionism saves realism even without assuming the abundant view, we might consider sparse views, on which there is no property of being jade—only being jadeite and being nephrite. Then if pluralism makes the reduction base of goodness disjunctive, there is no property of goodness. This may sound like victory for the error theorist, but it is not. Goodness still exists, as a disjunction of properties that exists without deserving to be called a property. Beliefs about it will be true because of disjunctions of scientific properties, rather than an individual moral property that is an individual scientific property. Reductionists can be open-minded about whether goodness is a property as long as they can identify it with scientific properties, uniquely or disjunctively.

19 As Jackson writes, “Jade, it turned out, comes in two quite different forms (nephrite and jadeite), but this did not lead us to deny the existence of jade. It led us to say there are two kinds where we might have thought that there was only one” (*From Metaphysics to Ethics*, 112).

20 See Armstrong, *Nominalism and Realism*. The abundant versus sparse terminology comes from Lewis, *On the Plurality of Worlds*. Abundant views are better at tracking when predicates refer, while sparse views support useful metaphysical distinctions.

To be compatible with folk belief, the sparse view must allow disjunctions of properties to satisfy nondisjunctive predicates like “good” in this way. Most of the things we care about are not natural kinds with a one-to-one correspondence to sparse properties. They are instances of highly gerrymandered kinds like the sonnets that poets write, the whiskies that drinkers enjoy, and the jobs that academic philosophers seek. The sparse theorist can say that there is no property of being a job, but should not say that the predicate “job” fails to refer to anything real. Denying that “job” refers because a single predicate must refer to a single property would force the sparse theorist to give up on having a job! Sparse theorist David Armstrong avoided this bad result and got a job that let him write, “In the theory of *properties*, it is in general a mistake to look for a one-one correlation to hold between properties and predicates.”²¹ If it turns out that goodness is metaphysically like water, jade, or jobs, belief in goodness can be true and moral realism is vindicated. Whether goodness is a property does not matter to reductionists as long as scientific properties satisfy “good” in some way. So the important question is not whether goodness is a natural kind that belongs on the sparse theorist’s list of properties. It is whether belief in goodness is false like belief in phlogiston, or true like belief in water, jade, and jobs.

The example of jade shows why reductionist metaethical views are compatible with even the most complex existing normative ethical theories.²² It is an interesting question whether reductionism is compatible with a moral theory suggesting an infinitely long reduction base, or whether only nonreductionism could handle infinite realizability, as Brink suggests. No moral theory that I know of claims that the number of moral principles is infinite. Jonathan Dancy’s particularism comes closest. But his view concerns the role of principles in moral thought and judgment rather than whether moral properties are infinitely realizable.²³ So infinite realizability, which reductionism in the philosophy of mind

21 Armstrong, *Nominalism and Realism*, 6.

22 Many philosophers discuss whether a “reductionist” view makes sense of thick concepts like “cruel,” as Roberts discusses (“Thick Concepts”). But this reductionism is an analysis of these concepts into evaluative and non-evaluative components suggested by noncognitivists like Blackburn who defend a ontological distinction between fact and value (“Through Thick and Thin”). My reductionism proposes fact-value property identities, rejecting this distinction. The issue of thick concepts was initially raised by cognitivist and naturalist Philippa Foot (“Moral Arguments”). For a treatment of thick concepts that makes them compatible with naturalism, see Väyrynen, *The Lewd, the Rude and the Nasty*.

23 Dancy has confirmed this in personal communication. He defines particularism as the view that “the possibility of moral thought and judgement does not depend on the provision of a suitable supply of moral principles” (*Ethics without Principles*, 7). A finite number of principles too large for human moral thought to apply would support his arguments. There is little motivation for insisting that the number of principles is literally infinite. What sort

must face, seems to be absent from ethics. Even on the most complex theories, moral properties can be understood as identical to finite disjunctions of natural kinds, like jade. Reductionism does not force us to accept monistic normative ethical theories that are overly simplistic. Jade is the model of how reductionism handles complex theories that do not identify moral properties with unique natural kinds.

So far, I have argued that moral properties are identical to special science properties, uniquely like H_2O /water or in finite disjunctions like jade/jadeite-or-nephrite, but not that they are identical to anything from fundamental physics. They may not be, since their special-science realizers probably are infinitely realizable at the level of fundamental physics. Even if the case of jade convinces you that finite realizability permits reduction of properties that are not natural kinds, you might join Brink in denying that anything can be reduced to an infinite disjunction and reject ethical reductionism because moral properties are infinitely realizable at the level of physics. So I explain why ethical reductionism succeeds if moral properties are infinitely realizable at the level of physics, as long as they are finitely realized at levels like psychology or sociology.

Reducing moral properties merely requires their being identical to some scientific properties, not necessarily those of physics. As Thomas Polger explains, realizability is relational, holding between particular sciences rather than absolutely.²⁴ Consider the water/ H_2O -type identity. Suppose it surprisingly turned out that protons were realizable by an infinite range of different arrangements of quarks. H_2O and water would then be infinitely realizable at the fundamental physical level, since the protons in the atoms would be infinitely realizable at that level. But this would give us no reason to reject the water/ H_2O -type identity! It would still be a necessary truth that water is H_2O . Identity would ground this necessity. While type-reduction would fail between protons and quarks, it would hold between water and H_2O . This is how moral properties relate to the special-science properties invoked in moral theories. Moral properties are infinitely realizable at the level of physics only because their special-science reduction bases—pleasure, actions caused by a mental state specified by deontologists, or a disjunction specified in some more complex moral theory—are infinitely realizable at that level. Type-reduction can still hold between moral properties and special-science properties.

The relational nature of realizability answers Brink's multiple realizability objections. Reductionists can accept that the special-science reduction bases

of normative ethical data could only be explained with infinite principles? Simplicity principles may also help keep naturalists away from infinitely complex moral theories.

24 Polger, "Two Confusions concerning Multiple Realization."

of moral properties may themselves be infinitely realizable at the level of physics. Treating others as ends in themselves presumably is infinitely realizable in worlds sharing our fundamental physics, and especially across the space of metaphysical possibility. But moral properties may still be type-reducible to tidy sets of special-science properties. Reductionists who see the Formula of Humanity as the sole normative ethical principle can simply say that however treating others merely as means is physically or nonphysically realized, it is identical to wrongness. Lying will then be wrong for humans, aliens, and ghosts.²⁵ The infinite physical and nonphysical realizers of moral properties are already accommodated between the fundamental properties and psychology by nonreductionism about the special sciences. No more accommodation is needed between psychology and ethics, making nonreductionism unnecessary there.

2. NONREDUCTIONISM'S EXPLANATORY PROBLEM AND REDUCTIONISM'S SOLUTION

This section examines arguments that nonreductionism lets moral properties explain phenomena just as special-science properties do. I respond that irreducible moral properties do not add to our explanations of observed regularities, giving naturalists no reason to believe in them. Reductionism uniquely or disjunctively identifies moral properties with scientific properties that add to our explanations of regularities, justifying belief in them. While nonreductionism cannot answer error theorists' epistemological arguments, reductionism can.

Gilbert Harman famously argues that irreducible moral properties do not explain our observations.²⁶ This is a problem because naturalists are reluctant to believe in properties that do not explain our observations. Nonreductionists respond by noting the role of irreducible properties in special sciences like psychology and suggesting that moral properties play a similar role, explaining regularities unexplained by the properties that they supervene on.²⁷ Here oppo-

25 For ghosts' actions to be wrong, we need not include supernatural properties in the reduction base. Reductionists should instead identify wrongness with more familiar properties like treating others merely as means. Even if these familiar properties supervene on supernatural properties in ghostly worlds, they do so without identity, so instantiating them does not require instantiating supernatural properties. Otherwise, treating others merely as means could not happen in our world, which lacks the supernatural properties. The multiple realizability of the mental is inconsistent with reductive supernaturalism, just as with reductive physicalism.

26 Harman, "Ethics and Observation." Audi, "Ethical Naturalism and the Explanatory Power of Moral Concepts," and Thomson, "Reply to Critics," raise similar issues.

27 As Darwall, Gibbard, and Railton observe, "Nicholas Sturgeon, Richard Boyd, David Brink,

nents of nonreductionism can concede that irreducible properties of some special sciences explain phenomena, perhaps as higher-level causes or in program explanation.²⁸ This concession leaves open the empirical question of whether moral properties explain phenomena. It does not entail that they actually explain phenomena, just as it does not entail that irreducible astrological or alchemical properties explain phenomena. Higher-level causation and program explanation should not save astrology and alchemy along with morality! Simply allowing irreducible properties to participate in explanations does not tell us why irreducible moral properties succeed while irreducible astrological properties fail. Here we should recall why we posit irreducible special-science properties in the first place.

Irreducible special-science properties are worth positing because they provide unified explanations of observed regularities that more fundamental properties explain only in a disunified way. Fodor writes that the sciences “state such true, counterfactual supporting generalizations as there are to state.”²⁹ These generalizations are systematic and unified accounts of regularities, some of which resist unified characterization in the language of physics. Reducing psychology to physics would prevent its laws from being well-unified, as they would involve huge disjunctions of physical states producing huge disjunctions of other physical states.³⁰ Psychological laws also seem to hold under different fundamental laws, perhaps in possible worlds in which a different version of string theory is true, or where nonphysical ghosts have psychologies like ours. So psychological properties could be instantiated without actual physical properties. This is why we need irreducible special-science properties as well as those of fundamental physics. If the regularities psychology describes did not exist, or if physics or

and others have pursued analogies with natural and social science to argue that moral properties might be both irreducible and explanatorily efficacious,” *Toward Fin de Siècle Ethics*,” 26.

28 Majors argues that properties of special sciences like psychology are genuine causes, so moral properties could be causes as well (“Moral Explanation in the Special Sciences”). Nelson (“Moral Realism and Program Explanation”) argues that moral explanations can be “program explanations” in which higher-level properties “program for” the existence of lower-level properties that really explain things, as Miller (*Contemporary Metaethics*) considers.

29 Fodor, “Special Sciences,” 114.

30 Against disjunctive laws and natural kinds, Fodor writes: “I think, for example, that it is a law that the irradiation of green plants by sunlight causes carbohydrate synthesis, and I think that it is a law that friction causes heat, but I do not think that it is a law that (either the irradiation of green plants by sunlight or friction) causes (either carbohydrate synthesis or heat). Correspondingly, I doubt that ‘is either carbohydrate synthesis or heat’ is plausibly taken to be a natural kind predicate” (“Special Sciences,” 109). He sees reductionism about mental states as providing similar disunity, with its huge disjunctions in the reduction base.

some other science explained them with equal unity, accepting irreducible psychological properties would be ontologically extravagant.

While psychology passes this test, many empirical theories have failed. Even if astrological claims occasionally accord with data (some Capricorns are ambitious), they do not explain additional regularities. The ambition of these Capricorns will be explained by biological, developmental, and social factors that leave no regularities for irreducible astrological properties to explain. So we reject irreducible astrological properties. While alchemists discovered some regularities concerning the production of acids and ceramics, chemistry explained these phenomena and more.³¹ We reject irreducible alchemical properties because they do not explain any additional regularities that chemistry leaves behind. Irreducible moral properties explain phenomena if they succeed in providing unified explanations of regularities that more fundamental properties do not. This determines whether they can figure in higher-level causation or program explanation.

Brink and Sayre-McCord argue that moral properties explain additional regularities, making ethics like psychology rather than astrology or alchemy.³² Sayre-McCord claims that “certain regularities—for example, honesty’s engendering trust or justice’s commanding allegiance, or kindness’s encouraging friendship—are real regularities that are unidentifiable and inexpressible except in terms of moral properties.”³³ On Fodor’s view, we are justified in treating special-science properties as more than mere heuristics because they explain additional regularities. Brink and Sayre-McCord claim that irreducible moral properties do so too.

Why must Brink and Sayre-McCord argue that moral properties explain nonmoral phenomena, as with justice engendering allegiance? Terence Cuneo describes how moral properties might explain moral phenomena, as when having a virtue causes someone to act rightly.³⁴ Brink and Sayre-McCord must go further and explain nonmoral phenomena because naturalists will reject putative special sciences with closed loops of irreducible properties that only causally explain each other. This rules out realms of spirits interacting only with each oth-

31 Morris, *The Last Sorcerers*.

32 Brink defends “the causal and, hence, explanatory irreducibility of higher-order facts—including moral facts—to lower-order facts that constitute, but are not identical with, those higher-order facts” (*Moral Realism and the Foundations of Ethics*, 197). For example, protest against the South African government is better explained in terms of its injustice than by the particular laws it passed, because different unjust laws would have resulted in similar protest (*Moral Realism and the Foundations of Ethics*, 195).

33 Sayre-McCord, “Moral Theory and Explanatory Impotence,” 276.

34 Cuneo, “Moral Facts as Configuring Causes.” At the end, he suggests further application to nonmoral explanations.

er, but does not rule out psychological properties. When race car drivers desire to win races and believe that they can win by accelerating, this explains not only the psychological event of their intending to accelerate, but the nonpsychological event of cars accelerating. Similar cross-domain explanations are ubiquitous. Economic events like industrialization explain geophysical events like climate change, which explain biological events like extinctions. Moral properties must do the same, or they will be as eliminable from our ontology as realms of spirits.

This takes us back to familiar debates about the explanatory potency of irreducible moral properties, but with a clearer view of the central question: do they explain additional regularities? Let us consider Sayre-McCord's claim that justice engenders allegiance. This generalization also has a merely psychological explanation. People desire that they and others be treated justly, and believe that allegiance to the just makes just treatment more likely. For justice's engendering allegiance to support the irreducibility of moral properties, morality has to provide a unified explanation of regularities that psychology does not explain. If we consider the reasons for regarding psychology as irreducible to neuroscience, and for positing chemical properties that are not alchemical properties, we can see two ways for moral explanations to have such an advantage. Justice could systematically engender allegiance in creatures lacking beliefs and desires. Or it could systematically engender allegiance in creatures whose beliefs and desires do not support psychological explanations of allegiance. I consider both options, explain how they parallel good defenses of irreducible special-science properties, and argue that they fail.

First, justice might systematically engender allegiance even in creatures that lack human psychological states like beliefs and desires, paralleling how psychology explains even robot behavior.³⁵ Neuroscience cannot explain robot behavior, since robots have other structures instead of brains. Psychology does so, supporting its irreducibility to neuroscience. Moral explanations of creatures without humanlike psychology would justify ethics as an irreducible special science by explaining phenomena in which psychology does not apply. This would be the best case for nonreductionists, as it would make their position perfectly analogous to Fodor's.

Sadly, justice does not systematically engender allegiance in creatures without our psychological architecture. Being just to amoebas and bees does not engender their allegiance. Creatures without humanlike psychologies do not systematically respond to moral properties, except perhaps in ways that scientific properties already explain. Psychology explains regularities that neuroscience does not address and that physics handles with disunity. But ethics has no

35 Some robots, like the Mars Rover, have a belief-desire-intention architecture.

similar advantages over psychology and sociology. Nonreductionists do not explicitly defend moral explanations of amoeba and bee behavior. They probably have not recognized that their arguments require these bad explanations to succeed. But if psychology is not reducible to neuroscience because its laws apply to creatures without humanlike brains, parallel arguments against reducing moral properties to psychological properties require moral laws to apply to creatures without humanlike psychologies.

Second, justice might systematically engender allegiance in creatures psychologically like us, but whose beliefs and desires do not suggest a psychological explanation of justice's engendering allegiance. Then psychology would fail to explain some regularities. If moral theories filled this gap, systematically explaining regularities that other special sciences did not, that would justify belief in irreducible moral properties. Belief in chemical properties similarly is justified by their ability to systematically explain regularities that alchemy does not.

Empirical evidence suggests that ethics does not fill any such gaps left by other sciences. If it did so, social scientists would invoke irreducible moral properties to explain regularities that scientific properties did not explain. But as Brian Leiter writes, "moral facts appear to play no role in any developed explanatory theory. . . . While, for example, there are Marxist historians using broadly 'economic' facts to explain historical events, there is no school of 'moral historians' using moral facts to do any interesting or complex explanatory work."³⁶ Current practice in the social sciences suggests that irreducible moral properties play no useful role in explaining regularities. Social scientists instead use psychological or sociological explanations that invoke economic or cultural facts.³⁷ Histories

36 Leiter, "Moral Facts and Best Explanations." Majors misunderstands the problem Leiter raises, taking it to be "that no moral generalization will be exceptionless" ("Moral Explanation in the Special Sciences," 150). That indeed would not be a problem. Fodor writes, "Intentional psychology is a special (i.e., nonbasic) science, so its laws are *ceteris paribus* laws. And *ceteris paribus* laws tolerate exceptions, so long as the exceptions are unsystematic" (*The Elm and the Expert*, 39). Leiter's point is that moral generalizations fail systematically enough to make them useless, so that social scientists must invoke nonmoral natural facts instead. Cuneo notes that explanations of empirical phenomena are incomplete and sometimes "it is just not obvious what these natural facts are," ("Moral Facts as Configuring Causes," 154). But psychologists, sociologists, and economists are discovering these natural facts, and moral facts do not seem to be among them.

37 Might ethics develop in such a way that moral facts would explain psychological or sociological events, contrary to current psychological or sociological methodology? As Parfit notes, "Non-Religious Ethics is at a very early stage," and we should be open-minded about how it will develop (*Reasons and Persons*, 454). But there is plenty of room for open-mindedness without expecting ethics to overturn the methodology of better-understood empirical disciplines.

in which irreducible moral properties exert pressure on political events, systematically pushing toward better outcomes, are pejoratively labeled “Whig history” and rejected along with the nineteenth-century school of historiography that provides the name.³⁸

While psychology does not have systematic exceptions that moral generalizations explain, the moral generalizations suggested by nonreductionists have systematic exceptions that psychology explains. To return to Sayre-McCord’s example, those who profit from injustice often align themselves with the unjust rather than the just. It is unclear how moral explanations would explain this regularity. Psychology explains it—their desire to profit from injustice exceeds their desire for justice. In this and other cases, moral generalizations have systematic exceptions that psychology explains, but not vice versa.³⁹

Frederick Engels claimed that justice was “social phlogiston.”⁴⁰ In trying to make irreducible moral properties explain social regularities just as chemists before Lavoisier tried to make phlogiston explain combustion, nonreductionism fails just as phlogiston theories failed. Error theory then defeats naturalistic moral realism.

Reductionism saves ethics from this misfortune. It treats moral properties as identical to scientific properties that explain phenomena, answering the challenge that we have no reason to believe in moral properties because the best explanations of our observations do not entail their existence. Harman, who famously brought this challenge against the Cornell Realists, allows that reductionism answers it.⁴¹ After discussing an example in which Jane believes that Albert has done something wrong after seeing him beat his cat, Harman writes, “certain naturalistic reductions of wrongness might enable us to explain how the wrongness of Albert’s action could help to explain Jane’s disapproval of it.”⁴² If wrongness is identical to causing pain, and causing pain explains Jane’s disapproval, wrongness explains Jane’s disapproval.⁴³ While disjunctions of natural

38 In *The Whig Interpretation of History*, Butterfield provides a classic criticism of Whig history.

39 Consider Brink’s example of apartheid ending in South Africa. On a psychological explanation, other restrictions regarded as unjust would have generated indignation, causing protest and instability. Moral and psychological explanations differ about what would happen if everyone regarded South Africa’s injustices as just, perhaps because of racism among whites and internalized oppression among blacks. Moral explanations implausibly predict that there still would have been instability and protest.

40 Engels, “The Housing Question.”

41 Harman, “Moral Explanations of Natural Facts.”

42 Harman, “Moral Explanations of Natural Facts,” 63.

43 Railton proposes a “reduction basis” for moral value (“Moral Realism,” 142) and later sympathetically considers the view I accept—a goodness/pleasure property identity on the

kinds may not provide a unified explanation of any one regularity, each disjunct provides a unified explanation of some regularity. Then wrongness can be identical to a disjunction of natural kinds, as jade is. Even if wrongness does not explain any regularity, its disjuncts each explain regularities, entailing its existence.

Nonreductionists cannot construct an analogous position on which irreducible moral properties supervene on arbitrary disjunctions of realizers. The way irreducible properties can inherit the causal powers of their supervenience bases might seem to suggest such a position, as nonreductionist theories of mental causation typically involve higher-level properties exercising causal powers through lower-level realizers.⁴⁴ But these higher-level properties explain additional regularities, unlike irreducible moral properties.

How can we discover which scientific properties are identical to moral properties? Many answers are possible. Brink's favored method of reflective equilibrium works just as well for reductionists as for nonreductionists. The moral theory that results from reflective equilibrium can be treated as a synthetic identity.

Unsurprisingly, moral properties fare better when they do not have to explain regularities beyond those of scientific properties and can simply be identified with elements of existing scientific explanations, as reductionism allows. Our interest in ethics is not about providing new, unified explanations of natural phenomena. We care about rightness, virtue, and goodness whether or not they explain additional regularities. We want to act rightly, be virtuous, and make the world a better place. It would be neat if moral properties explained regularities that scientific properties did not, but that is not why we care about them. Moral concepts leave open which phenomena moral properties explain, or whether they explain any at all.

Then why do naturalistic moral realists care about explanations? It is because

water/H₂O model ("Naturalism and Prescriptivity," 157). He argues that moral properties can explain things that beliefs about the properties cannot explain. His reductionism helps his moral explanations succeed, since moral properties that reduce to scientific properties explain whatever the scientific properties explain. This goes beyond what moral or scientific beliefs about the properties explain.

44 See Bennett, "Mental Causation." This avoids problems Elizabeth Tropman discusses about knowing moral facts "via inferences from the best explanation of some observed phenomenon," "Can Cornell Moral Realism Adequately Account for Moral Knowledge?" 33. Perhaps nonreductionist ambitions of explaining additional regularities prevented Joseph Long from invoking reflective equilibrium in response ("In Defence of Cornell Realism"). Reductionists do not require goodness to provide unified explanations of additional regularities that scientific properties cannot provide. While Lei Zhong ("An Explanatory Challenge to Moral Reductionism") is right that error theory and reductionism are equally simple, deep, and unified, reductionism gains an advantage in reflective equilibrium through its coherence with our existing moral beliefs.

of their broader epistemological commitments, not because of anything specific to morality. They are happy to believe in whatever the best explanations invoke or entail, and reluctant to believe in anything else. While they may accept that it is conceptually possible for there to be moral properties that do not explain anything, they deny that there is reason to believe in them. By identifying goodness with scientific properties or their disjunctions, reductionism makes our explanations entail its existence so that naturalists can believe in it.

Readers may be wondering how my explanatory arguments fit with the previous ones concerning multiple realizability. To solve multiple realizability for pluralism, I gave goodness a disjunctive reduction base that could not provide unified explanations. Then I argued against irreducible goodness on grounds that it does not add to our unified explanations of regularities. But how does a disjunctive reduction base avoid this problem? Why accept reductions of goodness to disjunctive bases that do not provide unified explanations, while rejecting irreducible goodness for not providing unified explanations? The answer is that if pluralism prevents goodness from doing unified explanatory work, only reductionism allows our scientific and normative ethical theories to jointly entail that there is goodness in the world.

Our best explanations invoke some things, entail the existence of disjunctions of the things invoked, and cast doubt on the existence of things that do not fit the data. Suppose normative ethics treats pleasure and democracy as the two good things, and metaethics commits us to the objectivity of their value. Nonreductionism then suggests understanding goodness as being constituted by pleasure or democracy but not identical to them. The existence of pleasure or democracy does not entail that anything fits that description, and the suggested empirical effects of this irreducible property do not fit our social-scientific observations. So we should reject such an irreducible property. Our explanations invoke only scientific properties, and entail that there are disjunctions of these properties. On a reductionist construal, the above ethical theory suggests that goodness is identical to pleasure or democracy. Even if this disjunction is too disunified to explain anything and is not a property on the sparse view, the existence of either disjunct entails its existence. So the reductionist construal of this ethical theory and our scientific ontology jointly entail moral realism.

Jade is identical to jadeite or nephrite. Jade does no unified explanatory work. But wherever there is jadeite or nephrite, there is jade. If reductionism is true and the true normative ethical theory says pleasure and democracy are the two good things, goodness is identical to pleasure or democracy. If either exists, there is goodness in the world.

3. HAPPY ENDING

I conclude by explaining why Cornell Realists should be happy to accept reductionism.

One does not accept nonreductionism for its own sake. One accepts it to address multiple realizability. Contrast the reasons to accept the rest of Cornell Realism. We might accept that our moral concepts demand objectivity because we feel that nothing less would count as genuine moral value. We might accept moral realism because error theory is so unappealing. We might accept naturalism because nonnatural moral facts are epistemically dubious and ontologically extravagant. We might accept externalism about moral judgment because it is possible for Satan to be rational, fully understand evil, and wholeheartedly do evil for evil's sake. By contrast, nothing directly pushes us toward nonreductionism. We want wrongness to be instantiated across infinite disjunctions at the level of physics and when ghosts torture other ghosts. But if reductionism delivers these results, nonreductionism has no further appeal. Moreover, reductionism answers Harman's objections, making our scientific explanations entail the existence of things that make moral belief true.

The simplicity of reductionism should attract naturalistically minded philosophers like the Cornell Realists. Einstein writes, "It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience."⁴⁵ Sometimes we have to abandon simple theories because they fail to explain the phenomena. But when a view like reductionism explains everything, it delivers the supreme goal of all theory.⁴⁶

*National University of Singapore
neiladri@gmail.com*

45 Einstein, "On the Method of Theoretical Physics."

46 For helpful questions and comments, I thank audiences at the Slovak Academy of Sciences, the Zagreb Institute of Philosophy, the Creighton Club, the University of Western Australia, the University of California at San Diego, the University of Erfurt, the Naturalism and Normativity in the Social Sciences Conference, Nanyang Technological University, the University of Texas at Austin, Southern Methodist University, the University of Houston, the University of Florida, the UNC Metaethics Working Group, the University of Cincinnati, the University of Puget Sound, the University of Tennessee, and the Australasian Association of Philosophy. Dan Korman and Nicholas Laskowski also offered helpful feedback.

REFERENCES

- Armstrong, David M. *Nominalism and Realism: Vol. 1: Universals and Scientific Realism*. Cambridge: Cambridge University Press, 1978.
- Audi, Robert. "Ethical Naturalism and the Explanatory Power of Moral Concepts." In *Moral Knowledge and Ethical Character*, 112–28. Oxford: Oxford University Press, 1997.
- Bennett, Karen. "Mental Causation." *Philosophy Compass* 2, no. 2 (March 2007): 316–37.
- . "Why the Exclusion Problem Seems Intractable, and How, Just Maybe, to Tract It." *Noûs* 37, no. 3 (September 2003): 471–97.
- Blackburn, Simon. "Through Thick and Thin." *Proceedings of the Aristotelian Society, Supplementary Volume* 66 (1992): 285–99.
- Boyd, Richard M. "How to Be a Moral Realist." In Sayre-McCord, *Essays on Moral Realism*, 181–228.
- Brink, David O. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press, 1989.
- Butterfield, Herbert. *The Whig Interpretation of History*. London: W.W. Norton and Company, 1931.
- Clapp, Lenny. "Disjunctive Properties: Multiple Realizations." *Journal of Philosophy* 98, no. 3 (March 2001) 111–36.
- Cuneo, Terence. "Moral Facts as Configuring Causes." *Pacific Philosophical Quarterly* 87, no. 2 (June 2006): 141–62.
- Dancy, Jonathan. *Ethics without Principles*. Oxford: Oxford University Press, 2004.
- Darwall, Stephen, Allan Gibbard, and Peter Railton. "Toward *Fin de Siècle* Ethics: Some Trends." *Philosophical Review* 101, no. 1 (1992): 115–89.
- Einstein, Albert. "On the Method of Theoretical Physics." *Philosophy of Science* 1, no. 2 (April 1934): 163–69.
- Engels, Friedrich. "The Housing Question." *Volksstaat* (1872). <https://marxists.org/archive/marx/works/1872/housing-question/cho3.htm>.
- Enoch, David. *Taking Morality Seriously*. Oxford: Oxford University Press, 2011.
- Feigl, Herbert. "The 'Mental' and the 'Physical.'" *Minnesota Studies in the Philosophy of Science* 2 (1958): 370–497.
- Finlay, Stephen. *Confusion of Tongues*. New York: Oxford University Press, 2014.
- Fodor, J. A. *The Elm and the Expert: Mentalese and Its Semantics*. Bradford: MIT Press, 1994.
- . "Explanations in Psychology." In *Philosophy in America*, edited by Max Black, 161–79. Ithaca, NY: Cornell University Press, 1965.

- . “Special Sciences (or: The Disunity of Science as a Working Hypothesis).” *Synthese* 28, no. 2 (October 1974): 97–115.
- Foot, Philippa. “Moral Arguments.” *Mind* 67, no. 268 (October 1958): 502–13.
- . “Morality as a System of Hypothetical Imperatives.” *Philosophical Review* 81, no. 3 (July 1972): 305–16.
- Harman, Gilbert. “Ethics and Observation.” In *The Nature of Morality: An Introduction to Ethics*, 3–10. New York: Oxford University Press, 1977.
- . “Moral Explanations of Natural Facts—Can Moral Claims Be Tested Against Moral Reality?” *Southern Journal of Philosophy* 24, no. 51 (Spring 1986): 57–68.
- Huemer, Michael. *Ethical Intuitionism*. New York: Palgrave Macmillan, 2005.
- Jackson, Frank. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Oxford University Press, 1998.
- . “In Defense of Reductionism in Ethics.” In *Does Anything Really Matter?* edited by Peter Singer, 195–211. Oxford: Oxford University Press, 2016.
- Jackson, Frank, Philip Pettit, and Michael Smith. “Ethical Particularism and Patterns.” In *Moral Particularism*, edited by Brad Hooker and Margaret O. Little, 79–99. Oxford: Oxford University Press, 2000.
- Kim, Jaegwon. “Multiple Realization and the Metaphysics of Reduction.” *Philosophy and Phenomenological Research* 52, no. 1 (1992): 1–16.
- Leiter, Brian. “Moral Facts and Best Explanations.” *Social Philosophy and Policy* 18, no. 2 (Summer 2001): 79–101.
- Lewis, David. *On the Plurality of Worlds*. Oxford: Blackwell, 1986.
- Long, Joseph. “In Defence of Cornell Realism: A Reply to Elizabeth Tropman.” *Theoria* 80, no. 2 (May 2014): 174–83.
- Mackie, J. L. *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin, 1977.
- Majors, B. “Moral Explanation in the Special Sciences.” *Philosophical Studies* 113, no. 2 (March 2003): 121–52.
- Miller, Alex. *Contemporary Metaethics: An Introduction*. 2nd ed. Cambridge: Polity Press, 2013.
- Moore, G. E. *Principia Ethica*. Cambridge: Cambridge University Press, 1903.
- Morris, Richard. *The Last Sorcerers: The Path from Alchemy to the Periodic Table*. Washington, DC: Joseph Henry Press, 2003.
- Nelson, Mark T. “Moral Realism and Program Explanation.” *Australasian Journal of Philosophy* 84, no. 3 (2006): 417–28.
- Parfit, Derek. *On What Matters*. Oxford: Oxford University Press, 2011.
- . *Reasons and Persons*. Oxford: Oxford University Press, 1984.
- Paul, L. A. “The Puzzles of Material Constitution.” *Philosophy Compass* 5, no. 7 (July 2010): 579–90.

- Place, U. T. "Is Consciousness a Brain Process?" *British Journal of Psychology* 47, no. 1 (February 1956): 44–50.
- Polger, Thomas W. "Two Confusions concerning Multiple Realization." *Philosophy of Science* 75, no. 5 (December 2008): 537–47.
- Putnam, Hilary. "The Nature of Mental States." In *Philosophical Papers, Vol. 2: Mind, Language and Reality*, 429–40. New York: Cambridge University Press, 1975.
- Railton, Peter. "Moral Realism." *Philosophical Review* 95, no. 2 (April 1986): 163–207.
- . "Naturalism and Prescriptivity." *Social Philosophy and Policy* 7, no. 1 (Autumn 1989): 151–74.
- Roberts, Debbie. "Thick Concepts." *Philosophy Compass* 8, no. 8 (August 2013): 677–88.
- Sayre-McCord, Geoffrey, ed. *Essays on Moral Realism*. Ithaca, NY: Cornell University Press, 1988.
- . "Moral Theory and Explanatory Impotence." In Sayre-McCord, *Essays on Moral Realism*, 256–81.
- Scanlon, T. M. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1998.
- Schroeder, Mark A. *Slaves of the Passions*. Oxford: Oxford University Press, 2007.
- Shafer-Landau, Russ. *Moral Realism: A Defence*. Oxford: Oxford University Press, 2003.
- Smart, J. J. C. "Sensations and Brain Processes." *Philosophical Review* 68, no. 2 (April 1959): 141–56.
- Sober, Elliott. "The Multiple Realizability Argument against Reductionism." *Philosophy of Science* 66, no. 4 (December 1999): 542–64.
- Sturgeon, Nicholas. "Moral Explanations." In Sayre-McCord, *Essays on Moral Realism*, 229–55.
- Svavarsdóttir, Sigrún. "Moral Cognitivism and Motivation." *Philosophical Review* 108, no. 2 (April 1999) 161–219.
- Thomson, Judith Jarvis. "Reply to Critics." *Philosophy and Phenomenological Research* 58, no. 1 (March 1998): 215–22.
- Tropman, Elizabeth. "Can Cornell Moral Realism Adequately Account for Moral Knowledge?" *Theoria* 78, no. 1 (February 2012): 26–46.
- Väyrynen, Pekka. *The Lewd, the Rude and the Nasty: A Study of Thick Concepts in Ethics*. Oxford: Oxford University Press, 2013.
- Zhong, Lei. "An Explanatory Challenge to Moral Reductionism." *Theoria* 78, no. 4 (December 2012) 309–25.

A VIEW OF RACISM

2016 AND AMERICA'S ORIGINAL SIN

Benjamin Mitchell-Yellin

THE ELECTION OF DONALD J. TRUMP as the forty-fifth president of the United States has reinigorated the American left's interest in combating racism in a way not seen since perhaps the Civil Rights Movement of the 1950s and 1960s. Trump's campaign rhetoric full of dog whistles gave way to an administration constituted by troubling people enacting troubling policies. Some have seen this as a wake-up call, while others have seen it as the unfortunate but eminently foreseeable price of America's original sin. Whether or not one sees the current political moment as a troubling aberration or as the laying bare of America's racist underbelly, many have a sense of a renewed mission to eradicate or at least mitigate racism in this country.¹

Many obstacles lie in the way of progress. This article draws attention to one: neither of the two main philosophical views about racism is fully up to the task of combating it. One view holds that racism is primarily a matter of institutional and social structures that perpetuate and enshrine racially disparate and oppressive policies and outcomes. The other holds that racism is primarily a matter of individuals' attitudes, such as beliefs about inferiority, hatred, and other forms of ill will. Both views appear congenial to the aim of combating racism. The first calls our attention to the sources of racism's most impactful harms, while the second is committed to the impermissibility of racist conduct and attitudes. But, as examination of the case of Attorney General Jeff Sessions shows, they both

1 It is important that the fight against racism not be characterized only as aiming at eradication. Racial realism, as described by Bell, for instance, holds that racism is a permanent feature of society ("Racial Realism"). The nature of the fight against racism, according to the racial realist, should not aim at eradication but rather mitigation. The goal is to "make life bearable in a society where blacks are a permanent, subordinate class," to "better appreciate and cope with racial subordination," and to recognize that "the fight itself has meaning ... that the struggle for freedom is, at bottom, a manifestation of our humanity that survives and grows stronger through resistance to oppression, even if that oppression is never overcome" (377–78). The anti-racist mission here is to mitigate the harms racism inflicts on the oppressed, even while the goal of eradicating racism is deemed illusory.

fall short of providing proper guidance. We need to think differently about racism in order to effectively combat it.

There is a second reason to adopt a different account of racism. Both of the main philosophical views appear to suggest that justificatory appeals to the concept of race preceded the attitudes or structures that supposedly constitute racism. But this gets it backward. A careful reckoning with the past shows that the concept of race was invoked to justify racially disparate structures of domination and attitudes of superiority that were already in place. We want a view that properly attends to the unfolding of history.

In response, this article presents a third view of racism, one that adopts a genealogical as opposed to analytical approach. The main claims to be defended will be (1) that individuals' attitudes as well as institutional structures are essential to a proper account of racism and (2) that this account must be essentially historical, taking proper notice, in particular, of the application of the concept of race in relation to the oppressive structures, relations, and attitudes that have come to characterize racism as we know it. The account offered here posits a drive to dominate that works in concert with the fact that the powerful get to write history, including the justifications for social relations and the institutions that encode them. Three virtues of the account are (a) its ability to support the anti-racist mission, (b) its description of the moral psychology implicated in racism, both interpersonal and institutional, and (c) its aid in allowing us to make proper sense of what racism is by attending closely to the ways in which it has evolved throughout history.

1. TWO VIEWS OF RACISM

Two kinds of analysis of racism dominate the philosophical literature.² On the *political view*, racism is analyzed in terms of systematic oppression of one or more racial groups by a society's basic institutions.³ For example, racism in

- 2 This is not to say that these two kinds of analysis exhaust the range of available views. Critical race theory, in particular, deserves mention as an important view. Yet it does not dominate the mainstream philosophical literature in the way these other two views do. Because this article seeks to intervene in that literature, it will take as its target the two kinds of view discussed in the text. But see Curry, "Will the Real CRT Please Stand Up?" for trenchant discussion of critical race theory's exclusion from the philosophical mainstream and the argument that this is connected to the very problem under consideration—namely, racism. And see note 27, below, for discussion of some connections between critical race theory and the third view argued for here.
- 3 Shelby, "Is Racism in the 'Heart'?" and "Racism, Moralism, and Social Criticism"; Mills, *The Racial Contract* and "'Heart' Attack."

housing is analyzed in terms of racially targeted injustices in the various institutions that make up the housing sector. These include entities such as the US Department of Housing and Urban Development, mortgage lending and servicing institutions, and local government agencies that enact and enforce zoning and tax regulations. On this view, racism with respect to housing stems from unjust social relations perpetuated and enshrined by these various institutions. On the *moral view*, by contrast, racism is primarily a matter of individuals' attitudes.⁴ Racism in housing is analyzed in terms of hatred, indifference, or disrespect toward people on the basis of their racial designation on the part of the individuals who play relevant roles in the housing market. For example, a region may be said to exhibit racist housing practices when its housing market is dominated by contract sellers who actively prey on home buyers of a particular racial designation out of malice or ill will. Because of their race, these home buyers are only offered predatory loans with deliberately unfair terms with the goal of swindling them as quickly as possible. On this second view, racism with respect to housing stems from the attitudes implicated in the conduct of particular individuals.

Each of these views has something going for it. The moral view has an easy time accounting for our condemnation of racism. If racism stems from objectionable attitudes, such as hatred and disrespect, then it cries out for moral censure. This appears to give this view a leg up on its rival. The moral significance of racism, on the political view, is not so straightforward. For one thing, it is not at all clear how to conceive of the (im)morality of institutions and their practices. For another, it is not clear how to properly account for the link between moral condemnation and moral responsibility in this context.

One response to these worries on behalf of the political view would be to deny the claim that racism is always immoral.⁵ Instead, one might claim that racism is always morally significant.⁶ When we identify racism, this is the beginning of the moral conversation, not the end. In the final analysis, we may not find anyone or anything that is deserving of moral condemnation. The thought here, expressed by some proponents of the political view, is that there may be people who innocently harbor racist beliefs or perpetuate racial injustice.⁷

The political view is not primarily concerned with individual morality, which

4 Garcia, "The Heart of Racism"; Blum, "Racism: What It Is and What It Isn't"; Glasgow, "Racism as Disrespect."

5 Cf. Garcia, "The Heart of Racism"; Haslanger, "Oppressions"; Glasgow, "Racism as Disrespect."

6 Shelby, "Is Racism in the 'Heart'?"

7 See discussion of the "benevolent" racist in Mills, "'Heart' Attack," and Shelby, "Is Racism in the 'Heart'?"

seems to be part of its attraction. This account of racism focuses squarely on the large-scale and pressing issues of entrenched, institutionalized racial bias and harm. In this way, it appears to place the scope and depth of the problem squarely in its analytical crosshairs. Racism touches our lives in many ways. It results in material harm for some and privilege for others. It also influences our attitudes and relationships, with respect to people and institutions, in ways that too often go unnoticed. By focusing on the social structures that contribute to material inequality and interpersonal bias, the political view appears to focus our attention on ways to effect widespread and impactful change, after which changes in hearts and minds may follow. Tackling the basic institutions that structure a racist society seems the best way to make a real and lasting impact in the fight against racism. Thus, there appears to be a pragmatic reason to prefer this account. Moreover, we ignore the racist structure of our institutions not only at the peril of those whose lives are made worse or cut short as a result, but also at the peril of an accurate sense of the way things work.⁸

It would be a mistake to think that the moral view does not also have something insightful to say about institutional racism. Some may claim that racism is always and only a matter of individual attitudes. But this seems obviously mistaken, and disingenuous. The moral view may analyze racism in terms of individual attitudes, but it need not stop there. It is compatible with recognizing the role that racist attitudes have played in the origination, development, and maintenance of the institutions that structure society. Institutional racism, on this view, results when the racist attitudes of individuals infect the social fabric.⁹ Educational institutions, for instance, may be said to become racist when the ill will of individual actors within them leads to the adoption and maintenance of policies with unjust racial disparities. Once these policies are in place, they infect the attitudes of those who pass through the racist educational system. Students learn a whitewashed historical narrative that reflects the beliefs of those who designed it—for example, beliefs about the inferiority of indigenous people and enslaved African people. Internalizing this narrative, many of these students come to harbor these very same beliefs. Then they pass them on to the next generation. The foundations of institutional racism, conceived of as stemming from individual attitudes, are thus self-reinforcing. And institutions can be racist even when none of the individuals who currently make them up harbor objectionable attitudes, so long as the infection has become sufficiently entrenched in the institution's policies and practices. Thus, the moral view can be seen to offer a more sophisticated analysis than initially meets the eye.

8 This is one of the insights behind Curry's "necessary knowledge thesis" ("Race").

9 Garcia, "The Heart of Racism."

The political view, too, comes in a cruder and a more sophisticated version. It has become a commonplace belief that explicit racism is on the decline; people do not spout epithets and endorse overtly racist policies like they used to. There are those who think that people do not harbor racist attitudes anymore. But if the 2016 election and its aftermath taught us anything, it is that this is simply not true. There's nothing like a political victory to make one feel comfortable screaming in public what a short time ago was fit for the country club or chat room only. This is not anathema to the analytical framework of the political view. Those who analyze racism in terms of institutional oppression can make sense of "real racists" in terms of participation in and habituation to a racist society, undergirded by racist institutions. All of us, to some degree or another, come to hold racist attitudes—explicit, implicit, or both—because of the way we are shaped by our social context. It is precisely because of the racist institutions that shape our collective modes of thought that we come, as individuals, to harbor the very attitudes some have claimed are only features of the past.

As should be clear by now, the sophisticated versions of the moral and political views largely agree when it comes to the scope of racism. They differ, mainly, in terms of explanatory priority. This has consequences for how they envision effective change. According to the moral view, institutional racism is real, but it is ultimately explained by appeal to individuals' attitudes. Systemic change requires that we change hearts and minds. The political view, by contrast, holds that the racist structure of social institutions ultimately explains the racist beliefs and intentions of individuals. And eradicating or mitigating racism, on this view, is a matter of restructuring the scaffolding on which society is built. Only by changing racist policies and practices will we change the hearts and minds of individual racists.

Each of these views has its attractions. But they cannot both be correct. Thus, it may seem as if a firm grasp of what racism is and how to combat it requires choosing between them. But that is not the case. Careful consideration of an example drawn from the 2016 election and transition, as well as examination of the impression they give about the relationship between race and racism, suggests that neither view is satisfactory.

2. THE PRAGMATIC AIM AND THE CASE OF JEFF SESSIONS

Though they offer distinct analyses of racism, the political view and the moral view are both closely connected to what we can call the *pragmatic aim*: the aim of eliminating, or at least mitigating, racism. This aim is apparently furthered by the political view's commitment to focusing on institutional and social struc-

tures, the elimination of which would effect widespread and materially beneficial changes in the lives of those who suffer most from racial oppression. As suggested above, this seems to be one major attraction of the political view. But the pragmatic aim is related also to the moral view in that it is justified by that view's commitment to the necessary immorality of racism.¹⁰ To claim that racism is immoral is to claim that there are reasons to combat it. Thus, it seems fair to ask how the two familiar views of racism fare with respect to furthering the pragmatic aim. Do these views offer proper guidance to those who aim to combat racism?¹¹

In short: no. And the battle over Senator Jeff Sessions's nomination and confirmation for the post of attorney general in the Trump administration can help us see why.

Many objected to Sessions's nomination on the grounds that he is racist and that the Department of Justice would, under his leadership, fail to combat, and likely exacerbate, the racism endemic to the American criminal justice system. But some opposed to his nomination argued that Sessions's own attitudes were beside the point; opposition to his nomination should focus on his record.¹² This would seem to comport well with the political view's focus on institutions. The call to focus on a public official's record is a call to focus on the institutional policies enacted during his tenure and, in this case, their consequences vis-à-vis racial justice. At the same time, however, it would appear to undermine the pragmatic aim of reforming racist institutions. Institutional change comes about through the actions of individuals, who are themselves moved by their attitudes.¹³ Attending to what is in the heart and mind of the individual in charge is not a distraction from, but rather a key element of, any plan to effect institutional change.¹⁴ Thus, the political view undermines efforts to combat structural injustice if it ignores the personal attitudes of certain key actors in the institutional structure. If the goal is to mitigate or eliminate racism in the criminal justice system, Sessions's attitudes are relevant to his qualifications for attorney general.

It may be objected that the political view does not really ignore individuals'

10 Garcia, "The Heart of Racism"; Glasgow, "Racism as Disrespect." This commitment is not confined to proponents of the moral view; see Haslanger, "Oppressions: Racial and Other."

11 It bears repeating that, as mentioned in note 1, above, those who take racism to be a permanent feature of society may consistently aim to combat it.

12 Bouie, "Jeff Sessions Fights for Racist Outcomes. Who Cares What's in His Heart?"

13 The conceptual claims about institutions in the text, both here and *passim*, have been influenced by Searle's account of institutional facts (*The Construction of Social Reality*).

14 See Madva, "A Plea for Anti-Anti-Individualism," for a convincing argument for this claim, in the context of combating discrimination, based on empirical evidence from the social sciences.

attitudes. Rather, it calls our attention to the fact that institutions can be racist even if the individuals who run them do not harbor racial biases or other problematic attitudes.¹⁵ It is true that the political view calls our attention to this fact and is correct to do so.¹⁶ But this does not adequately address the complaint that *changing* institutional policy requires more attention to individual attitudes than the political view appears to give. The existence of a racist institution may not depend on the attitudes of the individuals involved in its present-day operations, but combating institutional racism does. It is exceedingly unlikely, if not impossible, for a racist institution to adopt and implement non-racist policies in the absence of leadership that is both attuned to the problem and motivated to do something about it.

Consider now the moral view. Does it fare any better in accounting for the Sessions controversy? It seems not. To begin with, the difficulties inherent in trying to determine the attitudes that reside in a person's heart may preclude coming to any firm conclusion about whether or not Sessions is racist. This difficulty was reflected in the confirmation process and its coverage by the press. But even if we were to set that aside, the moral view, too, faces the problem of undermining the pragmatic aim of combating racism. The trouble here is different from the one facing the political view. While the moral view calls for attending to the attitudes of individuals, it focuses too narrowly on those that constitute ill will, possibly manifested by indifference.¹⁷ The worry about Sessions was not just that he might be actively opposed or indifferent to the interests of African Americans (among other groups), but also that his leadership would fail to bring about (and even forestall) changes in the criminal justice system necessary to combat the institutional racism already present.

Given the worry that Sessions's tenure in charge of the Department of Justice would perpetuate the racism already endemic to the criminal justice system, there is good reason to focus more widely than just on attitudes that constitute ill will or indifference. It takes more than lack of active antipathy toward or indifference to the plight of the oppressed to motivate effective change at the institu-

15 I would like to thank two anonymous reviewers for this journal for raising the concerns addressed here in comments on an earlier version of this article.

16 It is also true, as noted above, that the sophisticated moral view also allows that racist institutions may be run by individuals with no objectionable attitudes. It insists, however, that such institutions must have been infected by racist attitudes at some earlier time.

17 Garcia, "The Heart of Racism." For insightful discussion of the expressive significance of "emotional indifference" in the context of moral responsibility, see Smith, "Responsibility for Attitudes," esp. 242–46. Smith's Wybrow case may not be a perfect fit for the present discussion, but her comments are suggestive of the claim that one may exhibit objectionable attitudes toward others simply by exhibiting indifference toward their interests.

tional level. It takes positive concern for effecting this change. Even supposing that we could somehow determine that Sessions did not harbor racial ill will, the pragmatic aim of combating racism calls for more in this context. Even on the moral view, many of the racial disparities in the criminal justice system count as racist—the institution has a long history of infection by the racist attitudes of individuals implicated in the setting and carrying out of policy.¹⁸ Eliminating or at least mitigating this requires more than putting in charge someone without racial ill will; it requires putting in charge someone with attitudes that will motivate the necessary reforms. The moral view recommends looking into Sessions's heart, but not deep enough.¹⁹

At this point it may be objected that proper guidance in combating racism is not an appropriate criterion of adequacy for these analyses of what racism is; neither the political view nor the moral view need have this as a goal.²⁰ There is something to this objection. An analysis of racism need not be committed to combating it. Indeed, it may turn out that the correct analysis reveals that racism is ineradicable and that we cannot do anything to mitigate its effects. Nevertheless, the preceding considerations should be enough to motivate, at the very least, a hard look at the moral and political views by those who share the aim of combating racism.

3. RACE AND RACISM

The lesson from consideration of Sessions's nomination is that the analyses of racism offered by the political view and the moral view undermine the pragmatic aim because they do not take proper account of individuals' attitudes in the context of institutional racism. The political view does not pay proper attention to the attitudes of individuals; the moral view does not pay attention to the proper range of attitudes. Neither of these analyses provides an adequate account of how to combat institutional racism in this particular case. Especially in the context of the current political climate, marked by a reinvigorated commitment to the pragmatic aim, this gives us good reason to look for an alternative under-

18 For ease of exposition, we can refer to the various elements comprising the American criminal justice system as a single institution, and we can refer to the US attorney general, in his role as head of the Department of Justice, as the leader of this institution.

19 In fact, it is not enough that we consider attitudes of good and ill will. Effective institutional change requires not just motivation, but also know-how. As the point was put above, in connection with the political view, institutional leadership will be more likely to change institutional course when it is both *attuned* to the problem and *motivated* to do something about it. (I owe this point to an anonymous reviewer for this journal.)

20 I owe this objection to an anonymous reviewer for this journal.

standing of racism. But that is not the only reason to abandon the two familiar views. They also fail to adequately highlight the historical relationship between race and racism.

The moral view analyzes racism in terms of attitudes that already appeal to race—for example, beliefs *about* or hatred *of* African Americans. The political view analyzes racism in terms of institutional oppression of a particular group of people *on the basis of* their racial designation. Whether or not they require it, the suggestion, in both cases, seems to be that we do not have racism until we have appeals to race. This gets it backward.²¹ When we find racism, throughout time and across the globe, what we find is oppression that becomes racialized when the bodies of its victims are cited as justification. We see race invoked for the purpose of justifying domination of certain human beings by others.

This can be seen in historical discussions of the development of racism in the West.²² The broad-strokes historical narrative is as follows. One result of Europeans' increased mercantile travel in the late Medieval and early Renaissance eras was greater awareness of different cultures and skin tones. They came into contact with societies and body types they had not known existed or had little previous contact with. A second consequence was, of course, the colonization, enslavement, or eradication of these people and their lands. Through increased exploration, the European elite not only acquired new knowledge of what the world was like, but also new means of exploiting people for their own material gain. This changed internal class relations in European society. It became more economically advantageous to exploit foreigners rather than the serfs of one's own country. And this exploitation came to be justified in racial terms. These people did not have rights to land, labor, or bodily integrity *because they were "black" (or "brown" or "red" or "yellow")*.

This historical narrative stands in opposition to the suggestion that we do not have racism until we have appeals to race. The analyses of racism offered by the moral and political views may be taken to suggest that the initial harmful treatment of Africans or indigenous Americans at the hands of European colonizers did not count as racism. It only came to be racist once these people were oppressed on the grounds of their perceived racial designations. The moral view gives the impression that racism resulted when the individuals engaged in the harmful treatment internalized the relevant racial designations and acquired attitudes of ill will that reflected them. The political view gives the impression that racism did not result until institutions reflected these racial designations in their

21 As Ta-Nehisi Coates puts it: "race is the child of racism, not the father" (*Between the World and Me*, 7).

22 Mills, *The Racial Contract*; West, "A Genealogy of Modern Racism."

oppressive policies. This suggests that there was a moment when the oppression of these groups of people was profoundly transformed. It suggests that with the advent of race as a justifying factor, something new came on the scene. Harmful treatment of conquered peoples became exploitation and extermination of subhuman groups. But it seems more accurate to say that the availability of the notion of race allowed for a new way of justifying more of the same. The people being exterminated and enslaved were already regarded as subhuman, but there was not yet a need to recognize that this was the case, let alone offer a justification for it. That need came, among other things, from the readily apparent conflict between such treatment and the Enlightenment ideals of equality and freedom.

4. A THIRD VIEW

We have good reason to look for a view that highlights the historical interplay between the concept of race and the attitudes, structures, and conduct that we recognize as racist. Let us begin with a closer look at a historical account of a clearly racist system of practices that involved both individual racist conduct and racist institutional policies: slavery in America. As Barbara J. Fields details, the driving forces behind the enslavement of Africans in America were economic.²³ Africans were not the first people exploited for their labor in the American colonies—poor English and Irish serfs were brought over as cheap labor from the beginning of the American colonial program. But African captives quickly became the most profitable source of labor. This was made easier because they were taken from their geographic, social, and political contexts and transported to a foreign one.

In line with the historical narrative outlined in the previous section, the origins of the institution of chattel slavery in the American colonies had little if anything to do with race, and everything to do with the history of people, markets, and trade. The rationale of race may have been required for the institution of chattel slavery, where enslavement was a heritable condition, to come to fruition. But the factors that drove development of this institution were economic and political. It is more profitable to claim the offspring of one's property also as property, and it is easier to perpetually enslave a people who are not already a part of the development of the society and its defining notions. Part of the reason race was needed to justify slavery in America was that this budding country was founded on the notion of liberty and equality for all. This notion grew out of Enlightenment thinking, following long political struggles in Europe between

23 Fields, "Slavery, Race and Ideology in the United States of America."

the landed elite and the exploited serfs. Because enslaved Africans did not take part in these struggles, it was easier to leave them out of consideration when putting these ideals into practice.

When it came time to square the reality of everyday life in the American colonies with the aspirations of the revolutionary rhetoric espoused by the elites on this side of the Atlantic, something had to give. Those in power in the colonies needed a justification for treating one group of people in a manner that their professed principles deemed immoral. Their response was to exempt people of African descent from moral and political consideration—to dehumanize them. They appealed to race. The designation of the enslaved as “black” allowed the “white” people in power to distinguish between human beings on the basis of bodily appearance for the purposes of economic exploitation. Along with this came a moral and political hierarchy that purported to justify the institutionalized practice of chattel slavery. Thus was white supremacy born in the American colonies.

This is not to say that white supremacy has its exclusive origins in America. Other European colonial adventures also involved perceived and institutionalized racial hierarchies. And it is not to insist that the notion of English superiority over Africans was absent prior to the development of the institution of African slavery in the American colonies.²⁴ Nor is it to downplay the impact of white supremacy on the indigenous population in America, as well as in other colonized regions of the globe. But it is to claim that these various instances of white supremacy share an origin story. They featured the human drive to dominate coupled with a developing social hierarchy under the influence of Enlightenment notions of equal standing in the moral and political spheres. This combination was unstable. The claimed equal standing for all people threatened the developing economic, political, and moral hierarchy. The system and its inequities stood in need of justification. And that justification came to rest on readily apparent differences in what people looked like. It appealed to race.

This origin story supports a view of racism that differs from both the moral view and the political view in two key ways. First, the racialization of oppression—the birth of racism—derived neither primarily from individuals’ attitudes nor from social institutions. It came from both at once. The institution of slavery in America predated the racialized justifications of it that are now so familiar, and the same is true of antipathy toward enslaved people. Subjugated people, no matter their skin tone or continental origin, were believed to be infe-

24 See Vaughn, “The Origins Debate: Slavery and Racism in Seventeenth-Century Virginia,” for an overview of the debate among historians regarding the question of which came first, racism or African slavery, in the colony of Virginia.

rior. Early American colonists worked English indentured servants to death and traded them as property. Even free English colonists who did not command servants or own slaves looked down upon those who lacked independence. Racism bloomed in America when these attitudes and practices came to be justified by appeal to what people looked like. The institution of slavery became racialized, not just in operation, but also rationalization. Second, just as a bloom is the most visible form of a flower that grows from seed to bud, attitudes and practices justified in racialized terms announce to the world the racism that has been present for some time. The notion of race does not as much change these practices, as it does clarify what they have been all along. Racism comes to full fruition when existing oppressive practices are reinterpreted in racialized terms in the service of justifying the exploits of those in power. This is a moment not of transformation but of revelation.

Two elements remain constant, both before and after the invocation of race: (1) the drive to dominate and (2) the dominion of the dominant. These are the very features that serve as key ingredients in the recipe for white supremacy suggested above. Colonial elites and their European counterparts responded to the perceived need to justify the institution of African slavery in the Americas by invoking a distinction between people. The designation of one as “black”—that is, not “white”—was made available as a means of justifying oppressive practices and attitudes at the heart of colonial American society. The belief that certain people were inferior had its origins in the observation that they were unfree. But when it came to be the belief that this was so *because they were “black”* it served to justify their subjugation at the hands of those who were superior—now, *because they were “white.”* The subjugation came first and the racialized justification second, but, in contrast to the impression given by the familiar moral and political views, it was racist long before it was readily recognizable as such.

This historical narrative suggests that we have good reason to prefer a view of racism that is (1) essentially historical and (2) pluralistic—that is, the key elements in the analysis of racism are both irreducible. Call it the *genealogical view*.

The genealogical view of racism denies an analytical approach that seeks to identify necessary and sufficient conditions for something to be racist.²⁵ Instead, it aims to provide an account of what racism *is* by attending to what its *causes* were.

25 While I believe that this is the approach of those views I have labeled moral and political, an anonymous reviewer for this journal has helpfully drawn my attention to the fact that there are other ways of understanding what some instances of these views are trying to do. As opposed to articulating necessary and sufficient conditions on something being racist, they may, rather, be aiming to identify key or distinctive features of racism. In that case, it is my contention that the genealogical view has the advantage of highlighting the historical dimensions of racism and placing them at the center of its analysis. Doing so allows us also

Racism, on this view, is properly understood in terms of individual attitudes, social institutions, and conceptual ingenuity that were interwoven in various ways, at various times and places. The key claim of the genealogical view is that racist attitudes and racist social structures developed in tandem from attitudes and structures marked by the drive to dominate and the dominion of the dominant, and the racialized justifications merely announced what was already present. Pre-racial oppressive attitudes and institutions evolved into racialized ones by taking aim at victims conceptualized in terms of race. The roles of oppressor and oppressed were thus reinterpreted in racial terms. “White” people—which is to say, those in power—justified their oppression of “black” people on racial grounds. And these grounds were invented, or co-opted, to do just that.²⁶ But racialized justifications did not so much transform what they were applied to as much as they clarified it. The view is historical in that it calls our attention to the development of racism over the course of actual human history; it is essentially historical in that it claims that this historical development is inseparable from a proper understanding of the concept. It would be an obfuscation to claim that racism can be understood apart from grasping the way it developed over time.

The genealogical view further differs from the moral and political views in claiming that the essential analysans of those more familiar views are both required for an adequate understanding of what racism is. Whereas the moral view claims that the analytical buck stops, ultimately, with individual attitudes and the political view claims that it stops with basic social institutions, the genealogical view claims that neither of these elements is analytically primary to the other. Indeed, it claims that they are analytically inseparable.²⁷

to better appreciate the psychological mechanisms implicated in the development of racist attitudes and institutions.

26 One view has it that the concept of race was invented as a justification for racist practices. Bernard Boxill (“Introduction”) argues, against this, that the concept of race was originally developed by Europeans in order to explain the differences in appearance and culture they discovered through increased global travel. Later on, he contends, this concept corrupted Europeans’ natural sympathy and gave rise to the racist practices that came to mark the colonial era. The view laid out in the text need not take a stand on the origins of the concept of race. It may have been invented by Europeans to justify their exploitation of non-European people, or it may have originally been invented to explain the differences between people and then coopted as a rationale for oppression. Either way, the account in the text stands: the concept of race was used as a justification for oppressive attitudes and practices in the face of Enlightenment ideals of equality and freedom.

27 There are some notable affinities between the genealogical view offered here and extant accounts that fall out of the philosophical mainstream (see note 2, above). It is worth briefly noting how the view offered here differs from these other views. West offers a genealogical account of racism (“A Genealogy of Modern Racism”), and both West’s account and that

According to the genealogical view, pre-racial oppressive social institutions, in part, explain the development of racialized attitudes, such as hatred of “black” people and beliefs about their inferiority. These attitudes were cultivated in order to justify various oppressive institutional practices, such as African slavery in the American colonies. And these same institutions, now justified in racial terms, served to inculcate racist attitudes through the perpetuation and support of ideologies. At the same time, pre-racialized attitudes toward the enslaved explained the development and maintenance of the oppressive institutions in need of justification. Slavery in the North American colonies was initially justified by outright antipathy toward the enslaved and the belief that, no matter their appearance or continental origin, they were inferior. All of this predated the development of justifications in terms of race and their internalization. But once racialized attitudes came on the scene, they explained the further development and maintenance of the racist institutions they were meant to justify. Even after the abolition of slavery, African Americans were oppressed by means of Jim Crow, followed by the “colorblind” racial injustices that characterize con-

offered by McWhorter (*Racism and Sexual Oppression in Anglo-America: A Genealogy*) have Foucaultian elements shared by the genealogical view offered here—for example, all three views allow that ideologies have the power to shape individuals and institutions and stress the importance of history. The genealogical view offered here departs from these others in at least this way: it stresses the importance of individuals’ attitudes both with respect to their genetic contribution to the relevant oppressive institutional structures and with respect to their role in maintaining or evolving these structures. There is also an affinity between the genealogical view offered here and prominent accounts in the tradition of critical race theory. For example, the legal realism that inspires Bell’s racial realism stresses the importance of personal values to a proper understanding of legal decisions (“Racial Realism”). Thus, like the genealogical view offered here, racial realism is attuned to the complex interplay between personal attitudes and institutional structures. And there is shared interest in tracing the historical development and maintenance of racist oppression, seeking to account for the interplay between attitudes and structures over time. Finally, the genealogical view offered here shares some of the central tenets of critical race theory as laid out by Curry, such as a commitment to the social construction of the concept of race and its deployment to subordinate particular groups of people (“Will the Real CRT Please Stand Up?” 4–5). But the genealogical view offered in the text departs from views in the tradition of critical race theory in at least two key respects. First, it identifies a particular element of human psychology—the drive to dominate—as playing a central role in the development and maintenance of racism. Second, it is compatible with, but not committed to, claims about the permanence of racism. According to the genealogical view offered here, the question of racism’s permanence will turn, crucially, on the questions of whether the drive to dominate is a permanent feature of human psychology and whether race is a permanent object for its expression. Thus, the view offered here places a novel feature of our psychology at center stage in the discussion of racism’s past and future.

temporary America, especially in the criminal justice system.²⁸ The evolution of institutional racial oppression in America has been propelled forward, at least in part, by the racist attitudes of individuals. But the development of these attitudes is not properly explained without appeal to the institutional contexts that gave rise to them. Thus, the genealogical view weaves a historical narrative, to which both attitudes and institutions are essential. At the heart of this narrative are the psychological element of the drive to dominate and the social fact of the dominion of the dominant.

5. OBJECTIONS

Let us now consider two objections to the genealogical view, both of which amount to the charge that it does not offer a proper analysis of racism.²⁹ The first objection is that the genealogical view does not provide the proper tools to identify instances of racism because it does not offer a necessary condition on something's being racist. Racist attitudes and institutions might arise in some other ways than they actually have; even if history had been different, racism might still exist. The second objection is that the genealogical view does not appear to offer clear answers to questions we want answered by an analysis of a concept like racism. For instance, it does not tell us exactly when an attitude or institution comes to be racist.

These objections do not appear to apply to the moral or political views. The moral view suggests that a necessary condition on racism, whether personal or institutional, is the presence of attitudes of ill will toward people on account of their racial designation. And it tells us that personal or institutional conduct becomes racist at the point when these attitudes infect it. The political view suggests that a necessary condition on racism is the presence of institutional structures that enshrine and perpetuate unjust social relations with disparate racial impact. And it tells us that conduct or policies become racist when appropriately influenced by institutions that target groups or individuals for harmful treatment on the basis of their racial designations. The above objections suggest criteria of adequacy for a satisfactory account of racism, and the two familiar views appear to pass with flying colors.

How should the proponent of the genealogical view reply? One strategy would be to articulate ways in which the view, despite appearances, really does satisfy the suggested criteria for adequacy. But the best response is to insist that

28 Woodward, *The Strange Career of Jim Crow*; Alexander, *The New Jim Crow*.

29 I owe these to correspondence with Jorge L. A. Garcia on a different version of some of these ideas.

these objections address the wrong questions. The request for a necessary condition on the existence of racism rests on counterfactual aspirations. It asks: if things had gone differently, would such and such still count as racist? The essential historicity of the genealogical view is antithetical to this approach. Rather than ask what would be the case if things had been thus and so, the genealogical view focuses on the question: how did things come to be as they are? There is value in unearthing the actual unfolding of history and analyzing racism as we find it.

One might think that this is to miss the point. The counterfactual aspirations behind the request for a necessary condition need not supplant historical inquiry. They may rather supplement it in service of the pragmatic aim of mitigating and eradicating racism. If we have a clear means of identifying racism, however it might arise, then we will be in a better position to nip it in the bud as novel forms creep into existence.

This brings us to the second objection. It would seem that the pragmatic aim requires being able to identify when a given attitude, conduct, policy, or institution comes to be racist. But that is not so. The genealogical view has help to give in identifying, and perhaps even preventing, new and novel instances of racism, even though it eschews a precise answer to the question when a given individual or institution comes to be racist. For one thing, the view highlights the central role of the drive to dominate in the development of racist people and institutions. Where we find this drive operating in a context in which it either targets or is likely to target individuals or groups on the basis of a racial designation, we have reason to suspect racism is in the offing. Moreover, the view highlights the dominion of the dominant as also playing an important role. Thus, increasingly entrenched dominance of one racial group over another, whether interpersonal or institutional, is a red flag. And the historical focus of the genealogical view provides us with blueprints for the rise of new forms of racism based on past patterns. It prompts us to learn history's lessons.

6. MAKING SENSE OF 2016

Let us now to return to the case of Sessions's nomination to the post of attorney general. Some of the genealogical view's virtues will become clearer through investigation of how it fares better than the two familiar views in making sense of this case.

Recall the lessons learned from considering the moral and political views in relation to the Sessions case. In light of the pragmatic aim, the moral view appeared to focus too narrowly on attitudes constitutive of ill will. When it

comes to examining the attitudes of a nominee to head up an institution with racially disparate harmful effects, it will not do to make sure he does not harbor ill will toward members of disadvantaged racial groups. Change in the policies and procedures of the institution—here, the American criminal justice system—requires attitudes that spur active pursuit of positive change in the treatment of disadvantaged groups. The political view appeared to focus primarily, if not exclusively, on matters at the level of institutional policy and public record. Effective change in these areas, however, requires action by individuals, especially those in charge of relevant policy decisions. An institution already on course to make racially disparate negative impacts will not steer a different course if its captain is not motivated and prepared to change tack.

Now consider how the genealogical view fares in its treatment of the Sessions case. To begin with, it is not subject to either of the above difficulties. Unlike the political view, the genealogical view does not privilege institutional structures in its analysis of racism. Thus, it does not call our attention to these structures at the expense of investigating also the attitudes of individuals, especially those in positions of power and capable of steering the institution's course. Indeed, the genealogical view would justify special concern about the attitudes of these individuals, as they would be the ones in positions of dominance.

Unlike the moral view, the genealogical view does not focus narrowly on attitudes constitutive of ill will. The key psychological ingredient in the genealogical view is the drive to dominate, and myriad attitudes may serve to refocus this drive in order to promote what amounts to racist conduct or policy.³⁰ Yet this drive and (at least some of) these attitudes are not themselves constitutive of ill will. The desire for profit is an instructive example. This desire is not in itself constitutive of ill will, but it can be a significant part of the explanation of why a given individual or institution comes to oppress people on the basis of their racial designation. We saw this in the above historical narrative of African slavery in the colonies. And it has been implicated in the development of the “prison-industrial-complex” in the context of the US criminal justice system in its present form.³¹ The basic idea, in both cases, is that the drive to dominate comes to focus on a particular racial category because this is profitable. The resulting institutional structures are then reinterpreted in order to justify these practices in racialized terms. Just as slavery before it, criminality has come to be justified by appeal to a person's skin color—“black” men are “superpredators.”

The genealogical view not only avoids the apparent pitfalls of the moral and

30 The characterization in the text of the drive to dominate has been influenced by Katsafanas, “Nietzsche's Philosophical Psychology.”

31 Davis, *Abolition Democracy*.

political views, it also makes better sense of the widespread alarm over Sessions's nomination to the post of attorney general. This largely had to do with history, both his personal history and the history of the institution he was being nominated to lead. Sessions's record as an elected and appointed official—US attorney for Alabama, Alabama attorney general, and US senator from Alabama—provided what many found to be ample grounds for concern about racism.³² They were concerned about his personal attitudes regarding African Americans and about how he would steer federal policy with respect to their treatment by elements of the criminal justice system (e.g., oversight of local and state police). These concerns were especially pointed in the context of his nomination as US attorney general because this would put him in charge of an institution that has been well documented as a tool for racial oppression.³³ It would be difficult to sympathetically understand the extent of concern over his nomination without an adequate grasp of the history of this institution in America. Here the essential historicity of the genealogical view shows itself to be a real advantage. It can make sense of the level of concern about Sessions's nomination by highlighting not only the history of the American criminal justice system, but also the role Sessions has played in this history during his time in public office.

7. CONCLUDING REMARKS

Trump's electoral victory has reinvigorated anti-racism movements on the left. But the familiar moral and political views do not provide adequate analyses of racism, and they fall short of providing proper guidance for the anti-racist project. We should look to a view that focuses on racism's lineage and avoids the monistic focus on a single, ultimate analytical factor—attitudes or institutions—in favor of a pluralistic focus that recognizes the historical interplay between them. This is what the genealogical view seeks to do. It is thus able to capture the messy reality of racism and put us in a position to more effectively combat it.³⁴

Sam Houston State University
bmy@shsu.edu

32 Serwer, "What Jeff Sessions's Role in Prosecuting the Klan Reveals about His Civil-Rights Record."

33 Alexander, *The New Jim Crow*.

34 Several anonymous reviewers deserve thanks for very helpful comments on earlier versions of this article. I thank Jorge L. A. Garcia for comments on an earlier version of some of these ideas and, especially, for encouraging me to pursue them further. My debt to his own work on racism is, I hope, evident. Special thanks are due to David Wright for many helpful conversations on these ideas throughout all stages of their development.

REFERENCES

- Alexander, Michelle. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. New York: The New Press, 2012.
- Bell, Derrick. "Racial Realism." *Connecticut Law Review* 24, no. 2 (Winter 1992): 363–79.
- Blum, Lawrence. "Racism: What It Is and What It Isn't." *Studies in Philosophy and Education* 21, no. 3 (May 2002): 203–18.
- Bouie, Jamelle. "Jeff Sessions Fights for Racist Outcomes. Who Cares What's in His Heart?" *Slate*, January 11, 2017. http://www.slate.com/articles/news_and_politics/politics/2017/01/jeff_sessions_fights_for_racist_outcomes_who_care_what_s_in_his_heart.html.
- Boxill, Bernard. "Introduction." In *Race and Racism*, edited by Bernard Boxill, 1–42. New York: Oxford University Press, 2001.
- Coates, Ta-Nehisi. *Between the World and Me*. New York: Spiegel and Grau, 2015.
- Curry, Tommy J. "Race." In *Political and Civic Leadership: A Reference Handbook*, vol. 2, edited by Richard A. Couto, 550–59. Thousand Oaks: SAGE Publications, 2010.
- . "Will the Real CRT Please Stand Up? The Dangers of Philosophical Contributions to CRT." *The Crit* 2, no. 1 (2009): 1–47.
- Davis, Angela Y. *Abolition Democracy: Beyond Empire, Prisons, and Torture*. New York: Seven Stories Press, 2005.
- Fields, Barbara J. "Slavery, Race and Ideology in the United States of America." *New Left Review* I, no. 181 (May–June 1990): 95–118.
- Garcia, Jorge L.A. "The Heart of Racism." *Journal of Social Philosophy* 27, no. 1 (March 1996): 5–46.
- Glasgow, Joshua. "Racism as Disrespect." *Ethics* 120, no. 1 (October 2009): 64–93.
- Haslanger, Sally. "Oppressions: Racial and Other." In *Racism in Mind*, edited by Michael P. Levine and Tamas Pataki, 97–123. Ithaca: Cornell University Press, 2004.
- Katsafanas, Paul. "Nietzsche's Philosophical Psychology." In *Oxford Handbook of Nietzsche*, edited by John Richardson and Ken Gemes, 727–55. New York: Oxford University Press, 2013.
- Madva, Alex. "A Plea for Anti-Anti-Individualism: How Oversimplified Psychology Misleads Social Policy." *Ergo* 3, no. 27 (2016): 701–28.
- McWhorter, Ladelle. *Racism and Sexual Oppression in Anglo-America: A Genealogy*. Bloomington: Indiana University Press, 2009.
- Mills, Charles W. "'Heart' Attack: A Critique of Jorge Garcia's Volitional Conception of Racism." *Journal of Ethics* 7, no. 1 (March 2003): 29–62.

- . *The Racial Contract*. Ithaca: Cornell University Press, 1997.
- Searle, John R. *The Construction of Social Reality*. New York: Free Press, 1995.
- Serwer, Adam. "What Jeff Sessions's Role in Prosecuting the Klan Reveals about His Civil-Rights Record." *The Atlantic Monthly*, January 9, 2017. <https://www.theatlantic.com/politics/archive/2017/01/sessions-klk-case/512600>.
- Shelby, Tommie. "Is Racism in the 'Heart'?" *Journal of Social Philosophy* 33, no. 3 (Fall 2002): 411–20.
- . "Racism, Moralism, and Social Criticism." *Du Bois Review* 11, no. 1 (Spring 2014): 57–74.
- Smith, Angela. "Responsibility for Attitudes: Activity and Passivity in Mental Life." *Ethics* 115, no. 2 (January 2005): 236–71.
- Vaughn, Alden T. "The Origins Debate: Slavery and Racism in Seventeenth-Century Virginia." *The Virginia Magazine of History and Biography* 97, no. 3 (July 1989): 311–54.
- West, Cornel. "A Genealogy of Modern Racism." In *Race Critical Theories*, edited by Philomena Essed and David Theo Goldberg, 90–112. Malden: Blackwell, 2002.
- Woodward, C. Vann. *The Strange Career of Jim Crow*. New York: Oxford University Press, 1966.

DOES CONTRARY-FORMING PREDICATE NEGATION SOLVE THE NEGATION PROBLEM?

Robert Mabrito

SOLVING EXPRESSIVISM'S Frege-Geach problem requires specifying the attitudes expressed by arbitrarily complex moral sentences. Nicholas Unwin emphasizes the problems that arise in doing so for even the relatively simple case of negated atomic sentences.¹ Terry Horgan and Mark Timmons believe that contrary-forming predicate negation offers a solution to this negation problem.² I argue that their solution is incomplete.

1. THE NEGATION PROBLEM

Consider Marie, who is contemplating the morality of stealing. One possibility is that

w: Marie thinks that stealing is wrong.

Following Horgan and Timmons, assume that expressivists take it that for Marie to think that stealing is wrong is for her to oppose stealing.³ As Unwin points out, expressivists have a problem accounting for all the possible views Marie might have. To illustrate Unwin's point, consider the following:

	<i>Marie's View</i>	<i>Expressivist Interpretation</i>
N	Marie does not think that stealing is wrong.	Marie does not oppose stealing.
w~	Marie thinks that not stealing is wrong.	Marie opposes not stealing.
~w	Marie thinks that stealing is not wrong.	?

1 Unwin, "Quasi-Realism, Negation, and the Frege-Geach Problem."

2 Horgan and Timmons, "Expressivism and Contrary-Forming Negation."

3 Horgan and Timmons, "Expressivism and Contrary-Forming Negation," 98.

The negation problem is the problem of specifying—according to expressivism—what it is for Marie to think that stealing is not wrong. Given the reasonable assumption that Marie’s state of mind in $\sim w$ is distinct from her state in either $w\sim$ or N , expressivists cannot say that in $\sim w$ Marie opposes not stealing or simply does not oppose stealing. Indeed, as an argument due to Mark Schroeder shows, expressivists cannot take it that in $\sim w$ Marie’s attitude is one of opposition to anything.⁴

Horgan and Timmons attempt to solve this negation problem by distinguishing between two types of negation.⁵ The “not” in

Marie does not oppose stealing.

might express contradictory-forming sentential negation, in which case the above sentence is equivalent to

SN: It is not the case that Marie opposes stealing.

Or, it might express contrary-forming predicate negation, in which case the sentence is equivalent to

PN: Marie is unopposed to stealing.

For Horgan and Timmons, SN describes Marie’s state in N while PN describes her state in $\sim w$.⁶ Of course, Marie’s states in N and $\sim w$ are distinct if and only if SN and PN are not equivalent.

Horgan and Timmons argue that SN and PN are not equivalent because the concept of opposition is trivalent.⁷ Associated with every trivalent concept is a feature and an anti-feature, which are such that everything falls into one of three non-overlapping and non-empty groups: (i) those things with the feature, (ii) those with the anti-feature, and (iii) those with neither.⁸ Thus, while it is impossible to simultaneously possess both an anti-feature and its corresponding feature, it is possible to possess neither. The referent of a term that expresses a trivalent concept is the feature while the anti-feature is the referent of the term produced by applying a prefix such as “un-” or “in-.” Horgan and Timmons offer “pleasant” as an example; it picks out a feature while “unpleasant” picks out the relevant anti-feature.⁹ Thus, the contradictory-forming sentential negation of

4 Schroeder, *Being For*, 45–46.

5 Horgan and Timmons, “Expressivism and Contrary-Forming Negation,” 96–97.

6 Horgan and Timmons, “Expressivism and Contrary-Forming Negation,” 98.

7 Horgan and Timmons, “Expressivism and Contrary-Forming Negation,” 96–98.

8 Horgan and Timmons, “Expressivism and Contrary-Forming Negation,” 96.

9 Horgan and Timmons, “Expressivism and Contrary-Forming Negation,” 96.

P_1 : A is pleasant

is

P_2 : It is not the case that A is pleasant

while the contrary-forming predicate negation is

P_3 : A is unpleasant.

The two negations are not equivalent. P_2 is just the *negative* claim that A lacks the relevant feature, while P_3 is the *positive* claim that A possesses the relevant anti-feature.

If the concept of opposition is trivalent, then the term “unopposed” refers to the relevant anti-feature. If so, to be unopposed to stealing is not simply a matter of being not opposed, just like being unpleasant is not simply a matter of being not pleasant. To be unopposed is to possess an attitude that stands to the attitude of opposition as an anti-feature stands to a feature. Thus, SN and PN are not equivalent. SN is just the negative claim that Marie lacks the attitude of opposition while PN is the positive claim that Marie has its anti-feature, i.e., the attitude of unopposition. Thanks to trivalence, the proposal of Horgan and Timmons assigns distinct states to Marie in $\sim w$, $w\sim$, and N.

2. A COMPLETE SOLUTION?

In addition to assigning distinct states to Marie in $\sim w$, $w\sim$, and N, any acceptable expressivist solution to the negation problem must entail that Marie’s state in $\sim w$ is inconsistent with her state in w . Horgan and Timmons believe that their solution does so; they say it “provides the resources to explain why it is logically inconsistent to be simultaneously both opposed and unopposed to the same thing.”¹⁰ But it is important to distinguish two claims:

- A: An agent simultaneously being both opposed to x and being unopposed to x is a logically inconsistent state of affairs.
- B: Being opposed to x is logically inconsistent with being unopposed to x in the way in which a belief that p is logically inconsistent with a belief that $\sim p$.

Say two states are *incompatible* just in case a claim analogous to A is true of them; an agent cannot simultaneously instantiate two incompatible states. Say two states are *inconsistent* just in case a claim analogous to B is true of them; inconsis-

¹⁰ Horgan and Timmons, “Expressivism and Contrary-Forming Negation,” 99.

tent states clash in the way beliefs in inconsistent propositions do. Incompatibility does not entail inconsistency. (I leave it open whether inconsistency entails incompatibility.) Having a headache and not having a headache are incompatible states, but they are not inconsistent. Marie's headache and Angela's lack of a headache are not inconsistent in the way, say, Marie's belief that snow is white and Angela's belief that it is not white are. To borrow Allan Gibbard's way of putting the point, in the headache case there is a difference between Marie and Angela without there being disagreement, while in the belief case there is both.¹¹

That the concept of opposition is trivalent entails both that being unopposed is *distinct* from simply not being opposed and that it is *incompatible* with being opposed. These two claims follow, by definition, from the fact that unopposition stands to opposition as an anti-feature stands to a feature. But the fact that the concept of opposition is trivalent does not by itself entail that opposition and unopposition are inconsistent. Consider the concept I will call *hensitivity*. (Compare to James Dreier's example of *hiyo* or Gibbard's example of *yowee*.¹²) Say one is hensitive toward x just in case one possesses:

H: the disposition toward developing a headache when exposed to x .

There is a disposition that stands to H as an anti-feature stands to a feature, namely:

NH: the disposition toward *not* developing a headache when exposed to x .

One cannot possess both dispositions, but one might lack both because one's tendency to develop headaches is unrelated to one's exposure to x . Thus, the concept of hensitivity is trivalent. One is hensitive toward x if one possesses the feature H, unhensitive if one possesses its anti-feature NH, and neither if one lacks both. As in the case of opposition, that the concept of hensitivity is trivalent entails that being unhensitive is *distinct* from not being hensitive and *incompatible* with being hensitive. But it is not plausible to take it that hensitivity and unhensitivity are inconsistent. If Marie is hensitive toward paint and Angela is unhensitive toward it, we would not want to say that Marie's hensitivity is inconsistent with Angela's unhensitivity in the way in which Marie's belief that snow is white and Angela's belief that it is not white are inconsistent. In Gibbard's terminology, in the hensitivity case there is a difference between Marie and Angela, but they do not appear in virtue of this difference to be disagreeing with each other. The example of hensitivity shows that trivalence does not guarantee inconsistency.

Thus, the fact that the concept of opposition is trivalent only goes so far in

11 Gibbard, *Thinking How to Live*, 60–68.

12 Dreier, "Expressivist Embeddings and Minimalist Truth"; Gibbard, *Thinking How to Live*, 65.

solving the negation problem. Establishing trivalence establishes that being unopposed is distinct from simply being not opposed, but it does not establish that it is inconsistent with being opposed. An additional argument is needed to establish this second claim. Of course, the nature of this additional argument depends on how the attitude of unopposition is specified. Horgan and Timmons offer one proposal for doing so, though they allow for the possibility of others.¹³ Their proposal appeals to *motivated dispositions*, which are dispositions “to behave-in-a-specific-way-for-a-specific-reason.”¹⁴ For Horgan and Timmons, Marie’s opposition to stealing constitutively involves her possessing certain motivated dispositions concerning particular acts of stealing, such as

D: the disposition toward [refraining from taking candy from children *because* doing so is an act of stealing].¹⁵

For Marie to be unopposed to stealing involves her constitutively possessing the corresponding *negative* dispositions, such as

ND: the disposition toward *not* [refraining from taking candy from children *because* doing so is an act of stealing].

Note that the fact that Marie possesses ND does not entail that Marie is disposed to take candy from children. According to the account of Horgan and Timmons, Marie can still be disposed to refrain from stealing candy from children so long as she is not disposed to refrain *because* it is an act of stealing.¹⁶ She might be disposed to refrain from taking candy from children because it makes them cry. Also note that, while Marie cannot instantiate both D and ND, she might instantiate neither because she has no relevant motivated dispositions.¹⁷ Thus, Marie cannot be both opposed to and unopposed to stealing, but she might be neither. So understood, the attitude of unopposition stands to the attitude of opposition as an anti-feature stands to a feature.

Given this account, explaining why opposition is inconsistent with unopposition requires explaining, for example, why Marie’s disposition D is inconsistent with, say, Angela’s disposition ND in the way in which Marie’s belief that *p* and Angela’s belief that $\sim p$ are. If Marie opposes stealing while Angela is unopposed, they have *different* and *incompatible* motivated dispositions—such as D and ND—but, in virtue of this, do they count as disagreeing with each other? As the ex-

13 Horgan and Timmons, “Expressivism and Contrary-Forming Negation,” 100.

14 Horgan and Timmons, “Expressivism and Contrary-Forming Negation,” 100.

15 Horgan and Timmons, “Expressivism and Contrary-Forming Negation,” 101–2.

16 Horgan and Timmons, “Expressivism and Contrary-Forming Negation,” 100.

17 Horgan and Timmons, “Expressivism and Contrary-Forming Negation,” 101.

ample of hensitivity shows, that the concept of opposition is trivalent does not settle this question. Further argument is needed. For example, expressivists pursuing this line might borrow from Schroeder's proposed solution to the negation problem. It appeals to the notion of an *inconsistency-transmitting attitude*, where an attitude is inconsistency-transmitting just in case an instance of it directed toward a content x is inconsistent with an instance directed toward an inconsistent content.¹⁸ Belief is an uncontroversial example of an inconsistency-transmitting attitude. Schroeder thinks expressivists may assume there are other inconsistency-transmitting attitudes.¹⁹ If D and ND —which have inconsistent contents—are two instances of an inconsistency-transmitting attitude, then this would explain why opposition and unopposition are inconsistent. Unfortunately, it is not clear that motivated disposition is an inconsistency-transmitting attitude, since, in general, disposition is not, as the discussion of hensitivity demonstrates.

Note that the claim here is not that it is impossible to argue that opposition and unopposition are inconsistent in the relevant sense. (Perhaps motivated disposition is an inconsistency-transmitting attitude even though disposition in general is not.) Rather, the claim is that Horgan and Timmons have failed to provide such an argument. This failure may be due to a failure to distinguish between an agent instantiating an attitude and the attitude itself. For example, the failure to distinguish between the members of the following pairs:

- (1) Marie opposes x .
- (1a) The attitude of opposing x
- (2) Marie is unopposed toward x .
- (2a) The attitude of being unopposed toward x .

Horgan and Timmons have an explanation of why (1) and (2) are inconsistent—namely, that on their account opposition and unopposition are incompatible. But that explanation is not an explanation of the inconsistency of (1a) and (2a). And Horgan and Timmons need an explanation of that to completely solve the negation problem.

North Carolina State University
 robert_mabrigo@ncsu.edu

18 Schroeder, *Being For*, 43.

19 Schroeder, *Being For*, 42–44.

REFERENCES

- Dreier, James. "Expressivist Embeddings and Minimalist Truth." *Philosophical Studies* 83, no. 1 (July 1996): 29–51.
- Gibbard, Allan. *Thinking How to Live*. Cambridge, MA: Harvard University Press, 2003.
- Horgan, Terry, and Mark Timmons. "Expressivism and Contrary-Forming Negation." *Philosophical Issues* 19 (2009): 92–112.
- Schroeder, Mark. *Being For: Evaluating the Semantic Program of Expressivism*. Oxford: Oxford University Press, 2008.
- Unwin, Nicholas. "Quasi-Realism, Negation, and the Frege-Geach Problem." *The Philosophical Quarterly* 49, no. 196 (July 1999): 337–52.

ARE MORAL ERROR THEORISTS INTELLECTUALLY VICIOUS?

Stephen Ingram

CHARGING OTHER PEOPLE with intellectual vice is an important part of human life. One journalist might accuse another of being a *narrow-minded* conspiracy theorist, for example, or a lecturer might accuse her student of being *intellectually lazy* when he once again fails to do the required reading. We make “epistemic vice-charges,” as Kidd calls them, for various reasons.¹ Ideally, they can improve our dialectical situation by identifying, explaining, evaluating, and correcting bad epistemic activity. Less nobly, they can be used to stain a rival’s reputation, or to make laypersons doubt an expert’s testimony. Kidd distinguishes *robust* and *rhetorical* vice-charges.² In rhetorical cases, one agent negatively evaluates another but cannot “elaborate or ‘unpack’ the charge ... by explaining the reasoning that supports the negative judgment.”³ A rhetorical charge lacks epistemic force. Even if it is widely endorsed, without evidence to back it up it is indistinguishable from arbitrary name calling, and thus cannot advance a debate in an epistemically admissible way. But if a charge receives adequate evidential support it becomes *robust*, has real dialectical force, and can play a role in epistemic life.

Kyriacou suggests that we might need to issue an epistemic vice-charge against moral error theorists.⁴ He says this in response to an objection I make against the “moral fixed points view” defended by Cuneo and Shafer-Landau.⁵ This paper replies to Kyriacou. I show that there is little hope of making his vice-charge *robust* enough to vindicate the moral fixed points view. I begin with a brief overview of the debate. I then develop Kyriacou’s charge, before arguing that it fails to vindicate the moral fixed points view.

1 Kidd, “Charging Others with Epistemic Vice.”

2 Kidd, “Charging Others with Epistemic Vice,” 183–84.

3 Kidd, “Charging Others with Epistemic Vice,” 183.

4 Kyriacou, “Moral Fixed Points and Conceptual Deficiency: Reply to Ingram.”

5 Ingram, “The Moral Fixed Points”; Cuneo and Shafer-Landau, “The Moral Fixed Points.”

1. BACKGROUND

Cuneo and Shafer-Landau recommend that moral non-naturalists accept the following:

There are non-natural moral truths. These truths include the moral fixed points, which are a species of conceptual truth, as they are propositions that are true in virtue of the essences of their constituent concepts.⁶

A representative example of a moral fixed point (MFP) is the following:

A: It is wrong to engage in the recreational slaughter of a fellow person.

Cuneo and Shafer-Landau suggest that the essences of the concepts WRONG and RECREATIONAL SLAUGHTER OF A FELLOW PERSON are such that, when something satisfies the latter concept, it also satisfies the former.

I discuss putative attractions of this view below, but note first that it has the following entailment: *anyone who denies an MFP is not just morally mistaken, but conceptually deficient*. This includes error theorists, who hold that *all* of moral discourse is in error. Of course, to say that error theorists are conceptually deficient is not to say that they are metaethically flat-footed. Cuneo and Shafer-Landau make the subtler and more charitable claim that error theorists are misled as to what is implied by the content of their moral concepts. What misleads them, according to Cuneo and Shafer-Landau, is a suspect *methodology*. Specifically, one that involves “rejecting highly evident first-order moral propositions . . . on the basis of either highly controversial metaethical claims, or speculative empirical claims.”⁷ This supposedly underwrites the claim that error theorists are conceptually deficient. Not everyone is convinced, however. In earlier work I have argued that, when we examine the error theorist’s method carefully, it is not clear what exactly is wrong with it.⁸

More specifically, I argued that the error theorist uses a standard philosophical methodology—roughly, a four-stage process of theory development:

- (1) Get your intuitions in order. Scrutinise them, make them cohere, etc.
- (2) Develop a theory that, if true, vindicates your intuitions. Test it against standard criteria. If it passes, accept the theory. If it fails, move to (3).

6 Cuneo and Shafer-Landau, “The Moral Fixed Points,” 411–12.

7 Cuneo and Shafer-Landau, “The Moral Fixed Points,” 438.

8 Ingram, “The Moral Fixed Points,” 4.

- (3) Revise your theory. If it still fails after you have considered all available revisions, move to (4).
- (4) Accept that your intuitions were in error, and explain this error.

My suggestion was that error theorists operate in this familiar and apparently legitimate way. So, even if error theory is false, Cuneo and Shafer-Landau cannot ascribe conceptual deficiency to error theorists *on the basis that they are misled by their method*. Kyriacou, however, thinks that Cuneo and Shafer-Landau's view can be improved.

2. KYRIACOU'S VICE-CHARGE

Kyriacou proposes diagnosing the conceptual deficiency within a *virtue-theoretic* framework. The central idea is that virtues can dispose us to mastery of moral concepts, while vices can dispose us to conceptual deficiency. One way to develop this is to say that those who deny the MFPPs do so because they are *morally* vicious. Their callous and sadistic character "induces some conceptual deficiency and moral blindness."⁹ But that line of thought will not wash—the error theorists I know are lovely, not callous or sadistic at all. (Some of my best friends are error theorists!) For this sort of reason, Kyriacou suggests that error theorists do have an intuitive grasp of the MFPPs when engaging in first-order moral thought.¹⁰ However, at the second-order metaethical level, they resist recognizing them as conceptual truths.

So, strictly speaking, error theorists should be accused of *meta*-conceptual deficiency. Kyriacou holds that, to support this, we should argue that error theorists are "not sufficiently virtuous *in the relevant intellectual respects*."¹¹ He submits that error theorists are blinded by "a *stubborn and narrow-minded* insistence on the suspect philosophical methodology of strong reductionist naturalism that values ontological parsimony over and above the value of saving the phenomena."¹² If this is right, it can underwrite the accusation of meta-conceptual deficiency. This would undercut my objection, vindicating the MFPP view.

Kyriacou acknowledges that more work needs to be done to develop his suggested vice-charge. I do some of that work here, for it is necessary for the task of evaluating the prospects of making the charge robust. This matters, for if the charge remains rhetorical—that is, if it continues to lack adequate eviden-

9 Kyriacou, "Moral Fixed Points and Conceptual Deficiency," 4.

10 Kyriacou, "Moral Fixed Points and Conceptual Deficiency," 5.

11 Kyriacou, "Moral Fixed Points and Conceptual Deficiency," 5.

12 Kyriacou, "Moral Fixed Points and Conceptual Deficiency," 6, emphasis added.

tial support—then it will possess no more epistemic force than arbitrary name calling. I focus on the charge of narrow-mindedness, for if error theorists are narrow-minded *and* stubborn then the latter is likely due to the former. I also read the charge charitably as targeting a “local” rather than a “global” vice.¹³ That is, instead of asking whether error theorists have a narrow-minded psychological profile in most or all (trait-relevant) contexts—something that would be very hard to establish—I instead evaluate the claim that they are narrow-minded in some relevantly localized range of contexts.

Indeed, what Kyriacou seems to have in mind are contexts relating to the methodology of “strong reductionist naturalism.” So, let us say that the local vice-charge is that error theorists are *narrow-minded-about-naturalism*. This is what explains their meta-conceptual deficiency. To see what this amounts to, consider what is involved in narrow-mindedness for any given domain.

On one influential view, *open-mindedness* involves being motivated to be “receptive to new ideas and arguments even when they conflict with one’s own in order to ultimately get knowledge.”¹⁴ Being “receptive” to such ideas involves being willing and able to examine them without prejudice, without letting partisan commitments undercut conscientious inquiry. Correspondingly, then, *narrow-mindedness* involves failure or refusal to explore ideas or arguments that conflict with one’s own, or else a prejudicial exploration of them.

It is worth introducing a nuance here. Philosophers often use “narrow-minded” and “dogmatic” interchangeably, and this reflects part of ordinary discourse. But we often reserve “dogmatic” for an extreme form of narrow-mindedness seen in those who are *fanatically devoted* to a doctrine—those who fail to receptively entertain rival ideas because they are in the grip of a worldview that seems unarguable to them. This fanaticism does not *always* apply to those we call narrow-minded. Those who find a view they like and fail to receptively entertain its rivals (where this is due to the fact that they already have a view they like) differ from those who are *devoted to a worldview*, and who are thus *hostile* to rival ideas, and attempt to *repudiate* them.

For example, consider (1) a fascist who feels a need to burn books that espouse rival ideas, and (2) someone who has no desire to view sources of news other than *The Daily Mail* because they like what they read there, but who also feels no deep loyalty to that paper and is indifferent to where others get their news. Both are narrow-minded, but in comparing them I am inclined to reserve the charge of “dogmatism” for (1). On this use of terms, which reflects part of ordinary language, dogmatism entails narrow-mindedness but not vice versa. In

13 Cf. Doris, *Lack of Character*, 62–66.

14 Zagzebski, *Virtues of the Mind*, 269.

short, in some contexts we reserve “dogmatic” for those who are unusually dedicated to a worldview.

I note this because it forces us to ask whether error theorists should be charged with *dogmatic* narrow-mindedness. I have never met a book-burning error theorist—even if they turn out to be *dogmatically-narrow-minded-about-naturalism*, they will no doubt be at the milder end of the spectrum. But, whether or not they are dogmatic in this way, this is what the charge must be.

To see this, note that their alleged epistemic misconduct is radical. They are accused of meta-conceptual deficiency—reflectively resisting conceptual truths that they intuitively grasp. If this is underwritten by some form of narrow-mindedness, it is the dogmatic form to which we must appeal, for this is what it takes to explain why error theorists endorse a theory that conflicts with *their own intuitions*. In theorizing they fail to receptively entertain *their own ideas*. It plausibly takes a deep dedication to naturalism to produce such cognitive dissonance. So, the charge should be that error theorists are *dogmatically-narrow-minded-about-naturalism*. This is what explains the meta-conceptual deficiency involved in their denying the MFPS, and gives a way to defend Cuneo and Shafer-Landau against my earlier objection. With this charge on the table, we can now assess the prospects of making it robust rather than rhetorical. I suggest that the prospects are not good.

3. A ROBUST CHARGE?

3.1. *Intellectual Misconduct*

The first thing we have to do to get a robust charge is show that the intellectual misconduct that it is invoked to explain and evaluate *really is* intellectual misconduct. After all, if the error theorist does not misbehave, there is no reason to make the charge—no dialectical role for it to play.

Note that the misconduct here cannot just be that error theorists accept a theory that conflicts with their first-order intuitions. For if the theory is correct, and the intuitions in error, then this is the right thing to do. So, a robust charge will have to show that the putative MFPS *really are* MFPS, for if this can be reasonably rejected then it will be unclear that error theorists err in denying the MFPS, and thus that there is intellectual misconduct for the vice-charge to explain.

We must therefore assess the MFP view’s plausibility. Cuneo and Shafer-Landau argue that it explains four facts about the claims that they have in mind as MFPS: (i) that such claims hold of necessity, (ii) that they fix the boundaries for what counts as a moral framework, (iii) that their denial incites bewilderment,

and (iv) that they are knowable *a priori*.¹⁵ If this is right, then we have reason to call the MFPs conceptual truths, and there is intellectual misconduct to explain. But the MFP claim is not the only available take on (i)–(iv). Evers and Streumer argue that the essences of WRONG and RECREATIONAL SLAUGHTER OF A FELLOW PERSON do not make it a conceptual truth that

A: It is wrong to engage in the recreational slaughter of a fellow person.

They suggest that it is only a conceptual truth that

B: If anything is wrong, it is wrong to engage in the recreational slaughter of a fellow person.¹⁶

Evers and Streumer note that this suggests a different way of interpreting (i)–(iv).¹⁷ For instance, if their view is correct, it is only *on the condition* that A is true that it is necessarily true, or knowable *a priori*. And maybe Olson is right that error theory is emotionally difficult to accept—this may be what makes it so bewildering and disturbing when the MFPs are denied.¹⁸ We might also propose that B fixes the boundaries for what can count as a moral framework.

The idea that error theorists err in denying the putative MFPs (and that there is thus misconduct to explain) thus relies on the MFP explanation of (i)–(iv) being superior to this rival. We must settle the debate in favor of the MFP explanation before the charge can be seen as robust. However, to establish the superiority of the MFP explanation would effectively be to establish the MFP view, and whether this view is true is just what is at issue in this debate. So, one issue for defenders of the vice-charge is that they must find a dialectically appropriate way to establish the MFP view in order to show that there is some intellectual misconduct to explain.

3.2. Cross-Situational Consistency

Another, more troubling problem arises from the fact that we can ask whether error theorists have the psychological profile involved in being *dogmatically-narrow-minded-about-naturalism*. The best evidence we are likely to get here would be their acting similarly viciously in a suitable range of contexts, for this would indicate a characterological root from which the behavior springs. If error theorists are *dogmatically-narrow-minded-about-naturalism*, then it will be any context in which naturalism is at stake that is relevant. We can allow that error theorists

15 Cuneo and Shafer-Landau, “The Moral Fixed Points,” 407–8.

16 Evers and Streumer, “Are the Moral Fixed Points Conceptual Truths?”

17 Evers and Streumer, “Are the Moral Fixed Points Conceptual Truths?” 4n9.

18 Olson, *Moral Error Theory*, 143.

might inquire virtuously when participating in a debate in which naturalism is not at stake, but the charge commits us to predicting that they are out to defend naturalism generally (or that, when they are forced to pick a side in a debate in which naturalism is at stake, they seek to defend it and repudiate its rivals). Thus, if the error theorist is *dogmatically-narrow-minded-about-naturalism*, we can expect them to have naturalistic views on the metaphysics of mind, mathematics, and logic, for they will fail to receptively entertain the idea that qualia, numbers, and identity (etc.) are non-natural. This is a localized prediction of “cross-situational consistency.” How plausible is it?

It is true that some moral error theorists have been out to defend naturalism generally. Mackie considers whether numbers and identity (etc.) can fit into his naturalistic picture, and says that we must “show how, on empiricist foundations, we can construct an account of the ideas and beliefs and knowledge that we have of all these matters.”¹⁹ Of cases for which he had yet to give his own empiricist theory, he says: “I can only state my belief that satisfactory accounts of most of these can be given in empirical terms,” and that if any should “resist such treatment, then they too should be included, along with objective values, among the targets of the argument from queerness.”²⁰ Such remarks show little, however. They need not reflect failure to *receptively entertain* rival ideas, or some broader psychological profile.

And, besides, Mackie is just one man. The vice-charge targets error theorists *as a group*. Indeed, it has to do this in order to play the role that Kyriacou has in mind for it—namely, to vindicate the MFP view against my objection. I suggested that it is implausible that error theorists are conceptually deficient. Kyriacou replied that they plausibly are *meta*-conceptually deficient, at least once we explain how this came to be using a vice-charge. If the charge is to play this dialectical role alone, it must apply to all who participate in the relevant intellectual misconduct—to error theorists as a group. This is where issues emerge. There is no necessary connection between (a) the idea that non-natural *moral* truths are queer and (b) an *overarching* naturalistic worldview.²¹ That is, even if most actual error theorists endorse naturalism, nothing about error theory *forces* them to do so. We can imagine someone who finds irreducible normativity queer and thus believes that error theory is true, but who also thinks that qualia are irreducibly non-natural. Maybe nobody has defended this combination of views, but it is easy to *imagine* someone doing so. Nothing about error theory excludes this possibility.

What might one say to explain the alleged meta-conceptual deficiency of this

19 Mackie, *Ethics*, 39.

20 Mackie, *Ethics*, 39.

21 Cf. Olson, *Moral Error Theory*, 86.

person? A new explanation is needed, for they are engaged in the same activity as other error theorists, but not because they are *dogmatically-narrow-minded-about-naturalism*. We can be certain of this, for they forsake naturalistic views of qualia. An obvious reply is to further localize the charge. Instead of saying that error theorists are *dogmatically-narrow-minded-about-naturalism*, perhaps we could say that they are *dogmatically-narrow-minded-about-metaethics*. Someone with this vice can inquire viciously on metaethical issues, like the MFPPs, without this having any bearing on their inquiry about qualia and such. But this does not help, for we can imagine error theorists who inquire virtuously in metaethical debates that do not bear on their chosen form of error theory.

Consider someone who says that moral truths would have to be irreducibly normative, that irreducible normativity is queer, and that moral discourse is thus in error. This form of error theory is compatible with any view of moral motivation. Error theorists of this sort might inquire virtuously in debate over judgment internalism and externalism, for instance. So, the charge must localize further. And we can keep running this line until we reach the charge that error theorists are *dogmatically-narrow-minded-about-moral-error-theory*. But there are problems with such a narrow charge. First, it is uninformative. When the claim was that error theorists are *dogmatically-narrow-minded-about-naturalism*, we made interesting predictions about what they will do when their naturalistic worldview is at stake. We cannot say the same if the charge is that they are *dogmatically-narrow-minded-about-moral-error-theory*. The prediction in this case will be that the moral error theorist will defend moral error theory if it is at stake, and that, if she is shown an argument against moral error theory, she will seek to reject one of its premises. This prediction is uninteresting; it is what we expect of error theorists whether or not they are vicious.

Second, such a charge does not identify a *vice*. Local traits are one thing, but this charge is *so* narrow that the vice is not really an aspect of character. Doris does discuss hyper-local traits, but I agree with Cassam that “the sweet spot is somewhere between the mythical global traits which demand absolute consistency and ultra-fine-grained, situation-specific local traits which carry no implications for a person’s conduct other than in a single case.”²² The original charge plausibly does lie in the sweet spot but has limited dialectical force. The revised charge (that they are *dogmatically-narrow-minded-about-moral-error-theory*) is too localized to pick out a real character trait—something that could play a useful explanatory role.

A related issue with such a narrow charge is that it must be bolstered by a story of how the error theorist actually became dogmatically committed to error

22 Doris, *Lack of Character*, 66; Cassam, “Vice Epistemology,” 174.

theory. This is also true of the broader original charge, but getting a plausible story here is easy—for instance, we might say that error theorists are impressed (as many are) by the success of the natural sciences. But, having abandoned the original charge for the hyper-localized one, we must now ask how they came to this very specific vice-trait. And this is a harder task. Kyriacou suggests that the error theorist suffers from “theory-induced blindness.”²³ This might explain why they have a hyper-localized vice of being *dogmatically-narrow-minded-about-moral-error-theory*. However, not everyone with a theory also has theory-induced blindness, so we need to explain why this afflicts error theorists in particular. Otherwise the debate will descend into “an interminable exchange of charge and countercharge,” with error theorists accusing non-naturalists of having a theory-induced blindness that has led them to become *dogmatically-narrow-minded-about-non-naturalism*.²⁴ Maybe it is doable, but it is a high hurdle to jump.

It is worth emphasizing that, even if the original charge were to succeed for every *actual* error theorist, the issue raised here still emerges. The point of Kyriacou’s vice-charge was to explain the error theorist’s alleged intellectual misconduct. If such misconduct can in principle occur without this vice, as it evidently can, then some other explanation—whether a narrower vice-charge or something else entirely—will be needed to supplement or refine the original charge.

4. CONCLUSION

I do not wish to rule out the possibility that such explanations could be given, but they would have to survive the sort of reasoning outlined above, at least if we wish to make the charge in question robust enough for it to be able to vindicate the MFP view. There are major difficulties with identifying a relevant vice-charge that can (a) be made robust and (b) play the dialectical role needed to support the MFP view. So, without a robust charge the MFP view remains in trouble, for its *prima facie* implausible commitment—that error theorists are conceptually (or meta-conceptually) deficient—has yet to be made plausible. Until plausible diagnoses of this intellectual misconduct can be given, we must reject this commitment, and the MFP view with it.²⁵

University of Manchester
stephen.ingram@manchester.ac.uk

23 Cf. Kahneman, *Thinking, Fast and Slow*.

24 Kidd, “Charging Others with Epistemic Vice,” 184.

25 For comments on earlier drafts, I am grateful to Jimmy Lenman and Ian Kidd. I am also grateful to Lizzy Kirkham, Denise Fox, and Shirley Carter.

REFERENCES

- Cassam, Quassim. "Vice Epistemology." *The Monist* 99, no. 2 (April 2016): 159–80.
- Cuneo, Terence, and Russ Shafer-Landau. "The Moral Fixed Points: New Directions for Moral Nonnaturalism." *Philosophical Studies* 171, no. 3 (December 2014): 399–443.
- Doris, John M. *Lack of Character: Personality and Moral Behaviour*. Cambridge: Cambridge University Press, 2002.
- Evers, Daan, and Bart Streumer. "Are the Moral Fixed Points Conceptual Truths?" *Journal of Ethics and Social Philosophy* (March 2016).
- Ingram, Stephen. "The Moral Fixed Points: Reply to Cuneo and Shafer-Landau." *Journal of Ethics and Social Philosophy* (February 2015).
- Kahneman, Daniel. *Thinking, Fast and Slow*. New York: Farrar, Straus, and Giroux, 2011.
- Kidd, Ian James. "Charging Others with Epistemic Vice." *The Monist* 99, no. 2 (April 2016) 181–97.
- Kyriacou, Christos. "Moral Fixed Points and Conceptual Deficiency: Reply to Ingram." *Journal of Ethics and Social Philosophy* (March 2017).
- Mackie, J. L. *Ethics: Inventing Right and Wrong*. London: Penguin, 1977.
- Olson, Jonas. *Moral Error Theory: History, Critique, Defence*. Oxford: Oxford University Press, 2014.
- Zagzebski, Linda T. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. New York: Cambridge University Press, 1996.

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.