JOURNAL *of* ETHICS
& SOCIAL PHILOSOPHY

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.

SORTING OUT SOLUTIONS TO
THE NOW-WHAT PROBLEM

François Jaquet

WOULD YOU SAY that donating to charity is right and rape is wrong, that generosity is a virtue and vanity a vice, that children must respect their parents and husbands not beat their wives? Then you are mistaken according to the moral error theory, which takes all moral judgments to be false. On the main variant of the view, these judgments are false because they presuppose the existence of queer nonnatural facts—and these facts are queer because they would generate categorical reasons, i.e., reasons one would have regardless of one's desires.[1]

Some error theories naturally lead to the abandonment of their target judgments. Thus, religious error theorists rarely believe that God is omniscient or that He disapproves of homosexuality. Still, other error theories do not share this eliminativist tendency. We should continue to believe that $1 + 1 = 2$ even assuming that all arithmetical beliefs are false because they presuppose the existence of queer arithmetic facts.[2] The opposite policy would have terrible consequences, which suggests that our arithmetical beliefs are useful regardless of their truth. The moral error theory therefore raises the following question: What should we do with our moral beliefs? Should we get rid of them (as atheists get rid of their religious beliefs) or retain them (as we would retain our arithmetical beliefs if we discovered that they are false)? This question is known as the "now-what problem" for moral error theorists.[3]

---

1   Mackie, *Ethics*; Joyce, *The Myth of Morality*.

2   Field, *Realism, Mathematics, and Modality*.

3   Lutz, "The 'Now What' Problem for Error Theory." Two clarifications are in order. First, as opposed to a moral question, which would make no sense for error theorists to ask, this question presupposes only the existence of hypothetical reasons, i.e., reasons that depend on their bearer's desires (Joyce, *The Myth of Morality*, 221; Lutz, "The 'Now What' Problem for Error Theory," 353; Svoboda, "Why Moral Error Theorists Should Become Revisionary Moral Expressivists," 50). It simply does not arise for error theorists who deny the existence of such reasons (e.g., Streumer, *Unbelievable Errors*). Second, this question does not presuppose the truth of doxastic voluntarism, i.e., the view that we have direct control over our

Until recently, three solutions had been advanced in the literature. As its name indicates, abolitionism advocates the abolition of our moral beliefs.[4] No more surprisingly, conservationism advises us to keep our moral beliefs, as if nothing happened.[5] Finally, fictionalism is a form of revisionism: it recommends that we replace our moral beliefs with fictional attitudes.[6] Lately, however, two other revisionary theories have emerged in this debate. According to revisionary expressivism, we should replace our current moral beliefs with revised moral judgments constituted by conative attitudes: disapprove of rape instead of believing that rape is wrong.[7] According to revisionary naturalism, we should replace our current moral beliefs with revised moral judgments constituted by beliefs in natural facts: believe that rape causes significant psychological distress instead of believing that rape is wrong, for instance.[8]

In this paper, I argue that both revisionary expressivism (hereafter, "expressivism") and revisionary naturalism (hereafter, "naturalism") are in the end mere variants of abolitionism. In section 1, I present Toby Svoboda's case for expressivism and Stan Husi's, Matt Lutz's, and Wouter Kalf's case for naturalism. These authors list several desiderata that a suitable solution to the now-what problem should satisfy and then argue that their favorite theory satisfies these desiderata better than abolitionism, conservationism, and fictionalism. In section 2, I contend that, on closer inspection, neither expressivism nor naturalism fares better than abolitionism in these respects: abolitionists can help themselves to all the tools expressivists or naturalists use to satisfy their desiderata. Then, in section 3, I argue that the pieces of advice put forward by expressivists and naturalists actually amount to forms of abolitionism because there is nothing moral about the attitudes these philosophers say we should replace our moral beliefs with. Finally, in section 4, I show that this argument does not affect these theories merely because they are revisionist, since the main variant of fictionalism does not face

---

beliefs. The issue is not what to do with our moral beliefs at a given instant but whether to cultivate a disposition to accept moral propositions, as a "life strategy" so to speak (Joyce, *The Myth of Morality*, 219, 223–24).

4   Hinckfuss, *The Moral Society*; Garner, "Abolishing Morality"; Ingram, "After Moral Error Theory, After Moral Realism."

5   Olson, *Moral Error Theory*.

6   Joyce, *The Myth of Morality*; Nolan et al., "Moral Fictionalism versus the Rest."

7   Svoboda, "Why Moral Error Theorists Should Become Revisionary Moral Expressivists."

8   Husi, "Against Moral Fictionalism"; Lutz, "The 'Now What' Problem for Error Theory"; Kalf, *Moral Error Theory*. Revisionary expressivism and naturalism should be sharply distinguished from their hermeneutic homonyms, according to which moral judgments *already are* conative attitudes or beliefs in natural facts. These views do not aim to solve the now-what problem since they are incompatible with the error theory.

it. At the end of the day, force fictionalism is the only truly revisionary solution to the now-what problem.

1. THE CASES FOR EXPRESSIVISM AND NATURALISM

Assume with the error theorist that our current moral judgments are beliefs about nonnatural facts. Expressivism is the view that we should replace them with revised moral judgments constituted by conative attitudes. Thus, instead of believing that donating to charity is right, we should approve of donating to charity, and instead of believing that rape is wrong, we should disapprove of rape. Naturalism, by contrast, is the view that we should replace our current moral judgments with revised moral judgments constituted by beliefs about natural facts. It comes in two varieties. On the one hand, the objective naturalist's replacement judgments are beliefs in mind-independent natural facts. Instead of believing that donating to charity is right we should believe that donating to charity maximizes pleasure, and instead of believing that rape is wrong we should believe that rape causes significant distress. On the other hand, the subjective naturalist's replacement judgments are beliefs about our attitudes. Instead of believing that donating to charity is right we should believe that we approve of donating to charity, and instead of believing that rape is wrong we should believe that we disapprove of rape.

According to their respective proponents, expressivism and naturalism are the best solutions to the now-what problem because they meet four desiderata that all former theories failed to satisfy. First, a good solution would have us avoid moral error, for "There is an epistemic tension involved in making utterances that one believes to be false."[9] To be clear, the reason why we should not knowingly believe false propositions is not fundamentally epistemic. Epistemic reasons are relevant to the now-what problem only indirectly since the "should" in the question is hypothetical. Rather, it is because believing falsehoods would impede our interests, by providing us with misinformed desires, that we should avoid doing so.[10]

Second, a good solution to the now-what problem would give us reasons to act—both *motivating* and *normative* reasons; it would let our moral judgments move us to some extent and provide us with considerations that favor certain courses of action. This should come as no surprise since moral thought is es-

---

9   Svoboda, "Why Moral Error Theorists Should Become Revisionary Moral Expressivists," 57; see also Lutz, "The 'Now What' Problem for Error Theory," 355; Husi, "Against Moral Fictionalism," 88.

10 Joyce, *The Myth of Morality*, 179.

sentially practical. As Svoboda puts it, "intrapersonal motivation is a feature of morality worth preserving because it bolsters one's commitment to act for certain ends, increases one's self-control, and helps overcome weakness of will."[11] Should we replace our moral beliefs with motivationally inert attitudes, the moral practice would become worthless. Besides, morality would be even more effective if it provided us with normative reasons to act.[12]

Third, a good solution to the now-what problem would make it possible for two subjects to have a genuine moral disagreement. In other words, it would sidestep the disagreement problem often raised against speaker subjectivism. Suppose that Jim judges that homosexuality is wrong while Pam judges that it is right. According to speaker subjectivism, he thereby means that he disapproves of homosexuality while she thereby means that she approves of homosexuality. As a result, Jim and Pam's disagreement is merely apparent; they are talking past each other. But this is extremely implausible, as Jim and Pam clearly disagree in this case. The idea behind the third desideratum is that a solution to the now-what problem should entail that two subjects would genuinely disagree when one of them would judge that an act is wrong and the other that it is right. For it is on this condition that the revised morality will successfully help us to coordinate our behaviors.[13]

Finally, a good solution to the now-what problem would leave room for moral reasoning.[14] In particular, it would not bump into the Frege-Geach problem. This problem was initially addressed at hermeneutic expressivism, that is, the claim that moral sentences express conative attitudes. Although the worry concerns meaning more than it does reasoning as such, it is often framed in those terms: this brand of expressivism cannot account for the validity of moral arguments. A nice illustration is provided by moral *modus ponens* such as: stealing is wrong; if stealing is wrong, then fencing stolen goods is wrong; therefore, fencing stolen goods is wrong. Indeed, if a moral sentence's meaning is to be accounted for in terms of the conative attitude this sentence expresses, then the meaning of "stealing is wrong" cannot remain constant across this argument,

---

11 Svoboda, "Why Moral Error Theorists Should Become Revisionary Moral Expressivists," 67.

12 Lutz, "The 'Now What' Problem for Error Theory," 358.

13 Svoboda, "Why Moral Error Theorists Should Become Revisionary Moral Expressivists," 60; see also Olson, "Getting Real about Moral Fictionalism," 186. Naturalists tend to phrase the third desideratum directly in terms of coordination: a good solution to the now-what problem would help us to solve coordination issues (Husi, "Against Moral Fictionalism," 92; Lutz, "The 'Now What' Problem for Error Theory," 366; Kalf, *Moral Error Theory*, 162–70).

14 Svoboda, "Why Moral Error Theorists Should Become Revisionary Moral Expressivists," 55; Husi, "Against Moral Fictionalism," 92.

whose second premise does not express a negative attitude toward stealing. But then the argument is invalid because it commits the fallacy of equivocation. In the present context, the point is slightly different. Because solutions to the now-what problem are not in the business of describing the current moral practice, they need not account for the validity of such arguments. The idea is rather that a morality in which such arguments are invalid would be useless.

Expressivists and naturalists insist that the three older solutions to the now-what problem fail to satisfy these desiderata. While fictionalism avoids moral error, it does not allow for moral motivation, disagreement, and (in the case of force fictionalism) reasoning. Fictional attitudes would fail to motivate their bearers, disagreement cannot occur between people who accept different fictions, and making a fiction of mutually inconsistent propositions is not inconsistent. Conversely, while conservationism allows for moral motivation, disagreement, and reasoning, it cannot avoid moral error. Given my topic in this paper, I will focus on the contrast between expressivism and naturalism on the one hand and abolitionism on the other. So, how does abolitionism cope with these desiderata? For obvious reasons, it avoids moral error: if we stopped making moral judgments, we would no longer presuppose the existence of queer nonnatural facts. For reasons no less evident, however, abolitionism violates the remaining three requirements: since it involves the abolition of morality, it does not allow for moral motivation, disagreement, or reasoning. All in all, "Moral abolitionism would avoid moral error, but it also misses out on the useful features of morality."[15]

By contrast, according to Svoboda, expressivism satisfies all four desiderata. First, it avoids moral error: because the revised moral judgments would be conative states rather than beliefs, they would not be false—they would be neither true nor false. As they would not even purport to represent the world, they would not presuppose the existence of queer nonnatural facts.[16] Second, expressivism allows for moral motivation. Because the revised moral judgments would be conative states, states that are characterized by their world-to-mind direction of fit, they would prompt us to act accordingly.[17] As conative states, they would also provide us with hypothetical normative reasons, reasons whose existence would depend on them. Third, expressivism allows for the existence of moral disagreements. When Jim judges that homosexuality is wrong while Pam judges

15   Svoboda, "Why Moral Error Theorists Should Become Revisionary Moral Expressivists," 55.

16   Svoboda, "Why Moral Error Theorists Should Become Revisionary Moral Expressivists," 66–67.

17   Svoboda, "Why Moral Error Theorists Should Become Revisionary Moral Expressivists," 67.

that it is right, they do not contradict each other in the sense that either of their judgments must be false. Still, they disagree in attitude, just as someone who wants the end of the war and someone who wants the war to continue disagree in attitude even though neither of their attitudes is false.[18] Finally, expressivism allows for the possibility of moral reasoning. This might come as a surprise since the Frege-Geach problem was originally designed as an objection to hermeneutic expressivism. For his variant of expressivism to avoid the Frege-Geach problem, Svoboda relies on Simon Blackburn's higher-order attitudes account. On this account, to judge that fencing stolen goods is wrong if stealing is wrong is to disapprove of a combination of conative attitudes: disapproving of stealing without disapproving of fencing stolen goods. The *modus ponens* mentioned above is valid because if one accepts its second premise then one disapproves of accepting its first premise while rejecting its conclusion.[19] Now, some have objected that this account does not vindicate the argument's validity even though it shows that one must be somehow irrational to accept its premises while rejecting its conclusion.[20] One might expect Blackburn's solution to face this objection too when applied to revisionary expressivism, but Svoboda denies this:

> If we grant that expressivist moral judgments do not admit of logical relations among one another, they still provide good pragmatic reasons to those who hold them (e.g., to adopt or relinquish some moral attitude), and this may be enough to establish the possibility of moral reasoning.[21]

While moral arguments would not be strictly speaking valid in the revised practice, our conative attitudes would allow us to reason about morality.

Naturalism too satisfies all four desiderata, according to its advocates. First, it avoids moral error. Indeed, by contrast with moral beliefs, which presuppose the existence of nonnatural facts, nonmoral beliefs are true whenever the natural facts they purport to represent obtain. Second, naturalism provides us with both motivating and normative reasons to act morally. As Lutz puts it, "While there might not be any moral reasons, the salvaged concept can refer to something that can ground strong non-moral reasons for actions that are commonly considered to be moral."[22] Suppose that we would replace our moral beliefs with

18   Svoboda, "Why Moral Error Theorists Should Become Revisionary Moral Expressivists," 68.

19   Blackburn, *Spreading the Word*, 195–96.

20   Hale, "Can Arboreal Knotwork Help Blackburn out of Frege's Abyss?"

21   Svoboda, "Why Moral Error Theorists Should Become Revisionary Moral Expressivists," 69.

22   Lutz, "The 'Now What' Problem for Error Theory," 369.

beliefs about our states of approval and disapproval, as suggested by subjective naturalists. Then, assuming a Humean theory of reasons and provided that the replacement judgments are true, they would provide us with motivation and reasons to act. By judging that rape is wrong, you would judge that you disapprove of rape, which would provide you with both a disposition and a hypothetical reason against rape—one grounded in your conative attitude. Third, naturalism allows for the possibility of moral disagreement. Jim and Pam will disagree if he believes that homosexuality causes significant distress while she believes that it is harmless—and they will disagree in attitudes if he believes that he disapproves of homosexuality while she believes that she approves of homosexuality (provided that their beliefs are true). Finally, naturalism makes moral reasoning possible. Nothing unexpected there, since hermeneutic naturalism does not meet the Frege-Geach problem. As long as moral reasoning is made of inferences between beliefs about natural facts, it is plainly possible.

Expressivists and naturalists thus take themselves to put forward views that are not only distinct from but also superior to abolitionism. I shall question both assertions, arguing that expressivism and naturalism fare no better than abolitionism with respect to the desiderata and that they ultimately collapse into variants of abolitionism.

2. EXPRESSIVISM, NATURALISM, ABOLITIONISM, AND THE DESIDERATA

Proponents of expressivism and naturalism maintain that their respective views do better than abolitionism at meeting the requirements any solution to the now-what problem should meet. In this section, I will argue that they are wrong. It will be my contention that every desideratum that they can satisfy can be satisfied just as well by abolitionists. My argument to this effect will be twofold. Expressivists and naturalists can characterize their views as being revisionary either about moral thought only or about moral thought and discourse, but they will fail to outperform abolitionism on both characterizations, albeit for different reasons. Let us examine each option in turn, starting with the latter.

Expressivists and naturalists sometimes speak as though we should revise our moral attitudes and talk. Not only should we replace our moral beliefs with conative attitudes or beliefs in natural facts; we should also replace recognizably moral discourse with discourse of another kind. As Svoboda explains in the case of expressivism:

> One way of transitioning to revisionary expressivism is for participants in
> moral discourse to bring their moral language into line with some kind of

expressivism.... The replacement moral discourse would be recognizably non-cognitivist. Instead of saying, "It is true that donating to charity is morally obligatory," one might say, "Hooray for donating to charity!" or "Donate to charity."[23]

Likewise, a proponent of naturalism might maintain that, instead of saying that donating to charity is right, we should say that donating to charity maximizes pleasure and that, instead of saying that rape is wrong, we should say that rape generates significant distress. While the corresponding attitudes would be our revised moral thought, these sentences would constitute our revised moral language.

Whether or not it allows expressivism and naturalism to satisfy the desiderata listed in section 1, the problem with this move is that it is available to abolitionists too. Indeed, abolitionists do not merely urge us to dispense with moral thought and discourse. Because they recognize that morality fulfills a number of functions, they also advise us to adopt new tools for the same purposes. Thus, on Richard Garner's characterization of the view, "The moral abolitionist ... recommends that we abandon the practice, *or better, replace it with some motivational aids that allow us to acknowledge and deal with things as they are*."[24] Although he is an abolitionist, Garner suggests that we "replace the moral overlay with more effective and less duplicitous devices."[25] Likewise, according to Stephen Ingram, "The essence of the abolitionist position is a prohibition on uttering sentences and making judgements that ascribe moral properties to acts."[26] Abolitionists are therefore at liberty to argue that we should replace our moral beliefs with conative attitudes or nonmoral beliefs, and our moral sentences with nonmoral sentences that would express those attitudes. Indeed, this would not amount to uttering sentences and making judgments that ascribe moral properties to acts.

One rejoinder consists in saying that the replacement thought and discourse would be moral, hence out of reach for the abolitionist, who enjoins us to get rid of anything moral. It is unclear, however, why we should think of the sentences "Hooray for donating to charity!" and "Rape generates significant distress" as moral sentences since they do not contain terms we generally think of as moral. Anticipating this objection, Svoboda argues that these sentences would remain moral insofar as they would express moral judgments.[27] This answer would be

---

23   Svoboda, "Why Moral Error Theorists Should Become Revisionary Moral Expressivists," 63.

24   Garner, "Abolishing Morality," 504, emphasis added.

25   Garner, "Abolishing Morality," 505.

26   Ingram, "After Moral Error Theory, After Moral Realism," 231.

27   Svoboda, "Why Moral Error Theorists Should Become Revisionary Moral Expressivists," 63.

fair enough if the judgments in question were recognizably moral. But are they? Thus far, we have never thought of our approval of donating to charity and our belief that rape generates distress as moral judgments. We have always thought of them as paradigmatic examples of *non*-judgment (in the case of the former) and *non*-moral judgment (in the case of the latter). There does not seem to be any reason why abolitionists should suddenly take these attitudes to be moral judgments and consequently dissuade us from using them in place of our flawed moral beliefs.

Lutz too seems vaguely aware of this worry. He addresses the objection on behalf of his subjective brand of naturalism, which he calls "substitutionism." On this view, we should replace our moral beliefs with beliefs about our attitudes of approval and disapproval. Acknowledging that we already have such attitudes and that the abolitionist has nothing to object to our having them, Lutz maintains that, "unlike the abolitionist, the substitutionist will replace every discarded moral belief with a new belief about his attitudes."[28] Thus, "what is distinctive of the substitutionist approach is that it advocates for a kind of replacement procedure, where *every* moral belief is replaced by a corresponding belief about one's attitudes."[29] To generalize a bit, according to naturalism, we should not be satisfied with the nonmoral beliefs we already have; we should sometimes replace our moral beliefs with corresponding new beliefs in natural facts. And— the rejoinder goes—abolitionists cannot subscribe to the latter recommendation, and fail as a result to meet the desiderata.

Now, two cases must be distinguished: either your nonmoral beliefs are already in line with your moral beliefs or they are not. Assuming that you believe that donating to charity is right, either you already believe that you approve of donating to charity or you do not. If you do, then following the naturalist's recommendation you will simply get rid of your belief that donating to charity is right. In this case, abolitionism and naturalism will provide you with the same piece of advice. Suppose now that, although you believe that donating to charity is right, you do not approve of donating to charity—and as a result do not believe that you approve of donating to charity. Then, following the naturalist's recommendation, you will get rid of your belief that donating to charity is right and start to believe that you approve of donating to charity—and start to in fact approve of donating to charity.[30] But once again, nothing prevents the abolitionist from recommending the very same thing. Garner makes this quite explicit

---

28 Lutz, "The 'Now What' Problem for Error Theory," 365–66.

29 Lutz, "The 'Now What' Problem for Error Theory," 366, emphasis added.

30 Otherwise, you would form a false belief, and naturalism would violate a requirement very similar to the first desideratum: although it would avoid moral error, it would lead to non-

when he says, in the passage quoted above, that the abolitionist recommends that we replace our moral beliefs with motivational states that would prompt us to act in line with our previous moral beliefs. In this precise case, she would advise you to abandon your belief that donating to charity is right and form instead an attitude of approval vis-à-vis donating to charity, just as the subjectivist does. More generally, as long as expressivism and naturalism advise us to revise not only moral thought but also moral language, the tools they use to satisfy the desiderata are available to abolitionism as well.

Turn now to the second option expressivists and naturalists could pursue. Unlike the abolitionist, they might encourage us to continue using moral language but to endow moral terms and the sentences containing them with a different meaning. We should use moral sentences to express conative attitudes or beliefs in natural facts rather than beliefs in nonnatural facts.[31] In the revised moral practice, the sentence "Donating to charity is right" would no longer express the belief that donating to charity is right (i.e., instantiates the nonnatural property *rightness*) but the approval of donating to charity or the belief that donating to charity maximizes pleasure. By contrast, the sentence "Rape is wrong" would no longer express the belief that rape is wrong (i.e., instantiates the nonnatural property *wrongness*) but the disapproval of rape or the belief that rape generates significant distress.

This option seemingly has, over the previous one, the advantage of allowing expressivists and naturalists to make recommendations that are out of reach for abolitionists. Indeed, while the idea here is to revise moral thought but keep moral language unchanged, abolitionists believe that we should dispense with both moral thought and language. Moreover, on this characterization of expressivism and naturalism, we would have a reason to call the replacement judgments "moral": they would be expressed by moral sentences (assuming that a sentence is moral insofar as it contains recognizably moral words—more on that below).

A problem remains. Even if expressivism and naturalism advise us to keep using moral words and sentences, they will meet the desiderata no better than abolitionism. For it is our thoughts—whether one calls them moral (as the expressivist and the naturalist do) or nonmoral (as the abolitionist does)—that would allow morality's function to be performed, not our language. It is our attitudes, not our utterances, that motivate us and provide us with reasons to act,

---

moral error. But if there is something bad with moral error, then that thing is surely shared by error more generally.

31 Svoboda, "Why Moral Error Theorists Should Become Revisionary Moral Expressivists," 63; Husi, "Against Moral Fictionalism," 89; Lutz, "The 'Now What' Problem for Error Theory," 366.

possibly in ways that are currently considered moral. It is between our respective mental states, not between our respective utterances, that disagreements take place—in the absence of a disagreement in attitude, a disagreement in utterances is a disagreement in appearance only. And it is arguably with our judgments, not our utterances, that we reason and make inferences. Accordingly, as long as expressivists or naturalists are in agreement with abolitionists about the attitudes we should have, their views will yield the same results: when it comes to the desiderata, all will succeed or all will fail.

Of course, abolitionists cannot satisfy all the requirements *as they are phrased*: although they might in principle give us reasons to act in ways that are generally considered acceptable, they cannot make *moral* motivation, *moral* disagreement, and *moral* reasoning possible. Still, they could secure the same kind of motivation, disagreement, and reasoning expressivists and naturalists deem important. In phrasing the desiderata in explicitly moral terms, expressivists and naturalists make it impossible from the start for the abolitionist to meet them. But then they beg the question. If their recommendations amount to those of abolitionism in substance, they simply cannot argue that some important requirement is not met by abolitionism on the grounds that it makes no room for *moral* motivation, *moral* disagreement, and *moral* reasoning.[32]

3. MERE VARIANTS OF ABOLITIONISM

I have just argued that expressivism and naturalism constitute no improvement over abolitionism in terms of the desiderata listed in section 1. Abolitionists can make use of all the tools expressivists and naturalists appeal to in order to satisfy these requirements. In the present section, my claim will be stronger. I will argue that expressivism and naturalism are actually versions of abolitionism. This will be my argument:

1. Unless their replacement attitudes are genuine moral judgments, expressivism and naturalism are mere variants of abolitionism.
2. Expressivism's and naturalism's replacement attitudes are not genuine moral judgments.
3. Therefore, expressivism and naturalism are mere variants of abolitionism.[33]

32 Perhaps expressivism and naturalism can meet other desiderata that abolitionism cannot. But the burden of proof lies on the expressivist and naturalist to show this to be the case.

33   For the sake of readability, I will focus on moral thought from now on. Bear in mind that the argument is intended to apply just as forcefully to moral discourse.

Let me say a few words in support of each premise.

The first premise is *prima facie* plausible and does not require much argument. As we saw earlier, abolitionists only maintain that we should stop making moral judgments. They do not oppose our replacing these judgments with attitudes of another kind. Hence, if expressivists and naturalists advise us to replace our moral judgments with attitudes that, as a matter of fact, are not genuine moral judgments, they do not make any recommendation that the abolitionist needs to reject, and their views amount to different variants of abolitionism. To constitute authentic alternatives to abolitionism, expressivism and naturalism must recommend the making of genuine moral judgments (albeit revised ones).

Premise 2 calls for more support. According to the error theory's conceptual claim, moral judgments presuppose the existence of nonnatural facts. Importantly, this feature of them is supposed to be "nonnegotiable," in the sense that any judgment that does not presuppose the existence of nonnatural facts is not a genuine moral judgment—it is at best a "schmoral" judgment.[34] This nuance is of the utmost importance for the error theory. If this feature of moral judgments were negotiable, then moral realists could rightly propose "reforming definitions" of moral terms.[35] Moral judgments would survive the discovery that nonnatural facts do not exist just as motion judgments survived the discovery that absolute motion does not exist.[36] Now, all this provides support for premise 2. The expressivist and naturalist replacement judgments would not presuppose the existence of nonnatural facts, meaning that they would not share with current moral judgments one of their nonnegotiable features. This, in turn, means that they would not be genuine moral judgments.

Here is an objection that might be raised against premise 1: even assuming that there would no longer be genuine moral judgments and sentences after the reforms advocated by expressivists and abolitionists, the word "wrong" would still be in use. Does this not show that expressivism and naturalism are incompatible with abolitionism? Are abolitionists not as much opposed to moral words as they are to moral judgments and sentences? In a sense they are, but this sense does not offer any help to the objector. Consider a variant of abolitionism according to which we should replace our moral judgments with beliefs in natural facts. Proponents of this view could legitimately recommend that we

---

34 Joyce, *The Myth of Morality*, 3–5, and "Moral Anti-Realism," 24.

35 Brandt, "A Theory of the Right and the Good"; Railton, "Moral Realism."

36 Joyce, *The Myth of Morality*, 96. Expressivists and naturalists presumably know how crucial this is to the error theory—in fact, most say it explicitly (Husi, "Against Moral Fictionalism," 81–82; Lutz, "The 'Now What' Problem for Error Theory," 363–65; Kalf, *Moral Error Theory*, 4–5, 15–16).

use a new term, say "rong," to express these beliefs. Now, "wrong" as used after the naturalistic replacement procedure and "rong" as used after the abolitionist replacement procedure would mean the very same thing. Under these circumstances, it would be far-fetched to hold that "wrong" is a moral term while "rong" is not. If one is a moral term and the other not, that must be in virtue of their meaning. The difference just cannot be the respective sounds we make when we utter them or the letters we respectively use when we write them down. In consequence, it would make no sense for abolitionists to object to our using the word "wrong." Premise 1 still stands.

As we saw in section 2, Svoboda seems to accept it. However, he objects to something like premise 2 along the following lines:

> Whether this objection goes through depends on what makes some kind of judgment or discourse moral. If moral judgments necessarily involve beliefs, for example, then revisionary expressivist judgments are not genuinely moral ones. However, there is good reason to suspect that the best conceptions of what moral judgment and discourse *can* be are general enough to permit non-cognitivist varieties. After all, while traditional cognitivists reject the view that moral judgments and utterances are in fact non-cognitivist, this is grounded in their finding non-cognitivism problematic in some respect … other than its alleged failure to deliver genuinely moral judgments and utterances. Further, in order to avoid begging the question against non-cognitivists (e.g., by defining a moral judgment in cognitivist terms), cognitivists initially must work with some conception of moral judgment or discourse that is general enough to be susceptible to both cognitivist and non-cognitivist accounts. These considerations suggest that moral non-cognitivism would yield judgments and utterances that are recognizably moral.[37]

I remain unconvinced by this reply. Svoboda is correct when he claims that one cannot ground the denial that moral judgments are conative attitudes in hermeneutic expressivism's failure to deliver proper moral judgments. Hermeneutic cognitivists must indeed identify some other problematic feature displayed by expressivism. But the feature in question must be such as to indicate that hermeneutic expressivism fails to deliver genuine moral judgments, or we would not have an argument against it.

As for Svoboda's assertion that, in order not to beg the question, "cognitivists initially must work with some conception of moral judgment or discourse

37 Svoboda, "Why Moral Error Theorists Should Become Revisionary Moral Expressivists," 65–66.

that is general enough to be susceptible to both cognitivist and non-cognitivist accounts," it is beside the point. It may well be true but does not entail that cognitivists must *end up* with such a conception. As a matter of fact, error theorists accept the error theory because they end up with a much more specific conception of moral judgment and discourse. Indeed, the view that moral judgments ascribe nonnatural properties is typically a conclusion they reach by doing conceptual analysis rather than some sort of metaphysical investigation. When Joyce argues that this is a conceptually nonnegotiable feature of moral judgment, for instance, he relies on our conceptual intuitions, such as our intuition that a judgment presupposing only hypothetical reasons on the part of the agent or a judgment that would lack what he calls "practical clout" would not qualify as a moral judgment.[38] Now, is he thereby begging the question against hermeneutic expressivists and naturalists? He would certainly be if he assumed from the very beginning that moral judgments state nonnatural facts. But this is not what he does: this claim is the conclusion of the piece of conceptual analysis he provides, not an assumption he makes from the outset. It is a conceptual truth that moral judgments state nonnatural facts, yet one that, on pain of begging the question, must be established rather than assumed.

Absent a better expressivist or naturalist rejoinder, it is safe to conclude that the attitudes with which these views advise us to replace our current moral judgments do not deserve to be called "moral." From this, it follows that expressivism and naturalism are not distinct views from abolitionism. Does this mean that all revisionary theories are similarly doomed to collapse into a variant of abolitionism? As we shall see now, fictionalism might well be an exception.

4. WHAT ABOUT FICTIONALISM?

As a reaction to the objection that their solution to the now-what problem is in the end a form of abolitionism, expressivists and naturalists may identify a companion in guilt in fictionalism. In fact, Husi does just that:

> By the same token, however, one might equally wonder whether fictionalism represents but a version of abolitionism, surrendering morality nonetheless. After all, the fictionalist departing assumption is that morality is *fatally flawed*, in light of which moral discourse hardly could keep running just exactly as it did before this revelation.[39]

In short, if naturalism is a form of abolitionism, then so is fictionalism. Ulti-

---

38 Joyce, *The Myth of Morality*, 32, 42–43, and *The Evolution of Morality*, 57–64.

39 Husi, "Against Moral Fictionalism," 83.

mately, such a rejoinder would rest on an argument analogous to that raised in section 4 against expressivism and naturalism, but this time targeting fictionalism:

4. Unless its replacement attitudes are genuine moral judgments, fictionalism is a mere variant of abolitionism.
5. Fictionalism's replacement attitudes are not genuine moral judgments.
6. Therefore, fictionalism is a mere variant of abolitionism.

On the face of it, this argument sounds plausible. Indeed, one might suspect that expressivism and naturalism meet this objection just because they are brands of revisionism and that, *a fortiori*, any revisionary theory would meet it, including fictionalism. This would be unconvincing as a rejoinder to the objection: should fictionalism reduce to abolitionism too, this would not rescue expressivism or naturalism. But I believe that this rejoinder fails for yet another reason: one particular form of fictionalism is immune to this objection.

You might recall from the introduction that fictionalism, broadly construed, is the view that we should replace our moral beliefs with fictional attitudes of some sort. Time has come to give some more flesh to this characterization. Because we entertain two kinds of attitudes vis-à-vis fictional works, there are two versions of fictionalism. Let me illustrate this with the TV show *True Detective*. In the storyline, detectives Rust and Marty investigate the murder of Dora Lange. You might then have either of two attitudes toward this specific fictional fact. On the one hand, you could adopt an external perspective and talk about the fiction; telling your friend about the plot, you might say, "Rust and Marty investigate the murder of Dora Lange." In this case, the fiction bit would figure in the content of your attitude: your attitude would be a proper belief, and its content would be the proposition *In the fiction, Rust and Marty investigate the murder of Dora Lange*. On the other hand, you might adopt a perspective internal to the fiction and properly engage with it. In a way, you would also think that Rust and Marty investigate the murder of Dora Lange. In that case, however, the fiction bit would appear in the mode rather than the content of your attitude: your attitude would have the proposition *Rust and Marty investigate the murder of Dora Lange* as its content, but it would not be a belief; it would be a "*make*-belief."

As indicated, this gives rise to two variants of fictionalism. Content fictionalism is the view that we should adopt vis-à-vis morality an external perspective akin to that which you adopt toward *True Detective* when you tell your friend about its plot: we should replace our moral beliefs with beliefs whose content would contain a fictional operator.[40] Instead of believing that rape is wrong, say,

---

40 Nolan et al., "Moral Fictionalism versus the Rest."

we should believe that rape is wrong *in the moral fiction*. Force fictionalism, by contrast, is the view that we should adopt vis-à-vis morality an internal perspective akin to the one you adopt toward *True Detective* when you truly engage with the narrative: we should replace our moral beliefs with make-beliefs whose contents would be genuine moral propositions.[41] Instead of believing that rape is wrong, we should *make*-believe that rape is wrong.

In the present section, my claim will be that, while content fictionalism reduces to a variant of abolitionism, force fictionalism does not. If this is indeed correct, the latter is the only consistently revisionary solution available to the now-what problem. Since I accepted premise 1 earlier, I will now assume the truth of premise 4, lest I be accused of applying a double standard. Everything will therefore hinge on premise 5. To make my point, I will have to argue first that beliefs in propositions about a moral fiction would not deserve to be called "moral," and second that moral make-beliefs would.

Start with content fictionalism. Why think that its replacement attitudes would not be genuinely moral? For the same reason that had us conclude that expressivism's and naturalism's replacement attitudes would not. On this view, we should replace our moral judgments with beliefs about a moral fiction—the belief that rape is wrong, for instance, with the belief that rape is wrong in the moral fiction we have adopted. But beliefs about moral fictions do not state nonnatural facts: the belief that rape is wrong in the moral fiction we have adopted ascribes to rape the property of being wrong in the moral fiction we have adopted, and this property is as natural as the property of generating significant distress. Hence, content fictionalism's replacement attitudes do not share with moral judgments one of their nonnegotiable features, which means they are not genuine moral judgments.[42]

One might expect premise 5 to be true of force fictionalism too, but it is not. In order to make that clear, I will elaborate further on the notion of make-belief that is at play in this theory. The main proponent of the view distinguishes belief from make-belief as follows:[43]

> *Belief*: S believes that P only if (i) S has assented to P in her most critical contexts, and (ii) S is disposed to assent to P in her most critical contexts.

> *Make-belief*: S make-believes that P only if (i) S is disposed to assent to P

---

41   Joyce, *The Myth of Morality*.

42   This was predictable since content fictionalism is essentially a form of naturalism. It advocates replacing our moral beliefs with a specific kind of nonmoral beliefs.

43   Joyce, *The Myth of Morality*, 192–93.

in some contexts, (ii) *S* has assented to not-*P* in some more critical context, and (iii) *S* is disposed to assent to not-*P* in her most critical context.[44]

Two clarifications are in order if one is to apprehend force fictionalism adequately. First, assent to a proposition must be understood as a *mental* act rather than a *speech* act. The distinction between belief and make-belief could thus be phrased in terms of acceptance: both attitudes dispose us to accept their content in some contexts; however, while beliefs dispose us to accept their content in our most critical contexts, make-beliefs dispose us to reject their content in those contexts. Second, in the relevant sense, for any pair of contexts <*Cn, Cm*>, "*Cn* is more critical than *Cm* if and only if *Cn* involves scrutiny and questioning of the kinds of attitude held in *Cm* but not vice versa."[45] Accordingly, deliberative contexts are less critical than contexts in which we do metaethics because we make in the former an assumption that we question in the latter, the assumption that there are moral truths. To sum up, force fictionalism advises us to accept moral propositions in everyday deliberation and yet reject them while doing metaethics.

Let us return to the matter at hand. There is a sense in which force fictionalism's replacement judgments are not genuine moral judgments: moral make-beliefs are not moral *beliefs*. Still, two replies are available to the fictionalist. First, even though moral make-beliefs would not qualify as moral *judgments* because they are not judgments in the first place, they would nonetheless share moral judgments' nonnegotiable problematic feature since they would essentially be dispositions to accept genuine moral propositions, and thus to ascribe nonnatural properties, no less than our moral beliefs currently do. Since it is to the ascription of nonnatural properties that abolitionists object, this means that force fictionalism is incompatible with abolitionism.

Second, it should be acknowledged that fictional replacement judgments cannot count as moral judgments so long as one construes judgment as a kind of mental state, namely as belief. However, although this construal is widespread enough in the metaethical literature (which is why I have so far stuck to it for the sake of presentation), it is at best a benign simplification. Philosophers of mind do not equate judgments with beliefs; they generally take judgments to be

---

44 One might object that this characterization fails as an account of the ordinary notion of make-belief. Maybe it does, but this is immaterial to my argument. What matters for my argument is that the attitude I call "make-belief," and whose adoption force fictionalists recommend, involves genuine moral judgments—and, even more to the point, that abolitionists cannot recommend its adoption. Thanks to an anonymous referee for drawing this issue to my attention.

45 Joyce, *The Myth of Morality*, 193.

those mental acts that our beliefs dispose us to perform. Even some metaethicists make this distinction, as testified by this passage from Sigrun Svavarsdóttir:

> It seems more accurate to think of judgments as mental acts rather than mental states, although they are, of course, the onsets, expressions, or activations of mental states.... I distinguish between moral judgments and the corresponding mental states ... and talk about the former as manifesting the latter.[46]

Interestingly for our purposes, in this more accurate understanding of the notion of judgment, the acceptances mentioned in the respective characterizations of belief and make-belief above *are* judgments—they are the mental acts that our beliefs dispose us to perform or by which they are manifested in our most critical moments. Besides, and most significantly, the acceptances that our *make*-beliefs dispose us to make in everyday contexts are judgments, just like the acceptances that our beliefs dispose us to make in more critical contexts. And for a good reason: *they are the same mental acts*. As a result, the acceptances that our *moral* make-beliefs would dispose us to make in everyday deliberation are moral judgments no less than those our moral beliefs currently dispose us to make.

Provided that moral judgments are construed—following the dominant use in the philosophy of mind—as mental acts rather than states, force fictionalism therefore entails that we should continue to make moral judgments, although we should replace our moral beliefs with moral make-beliefs. This means that it is not a mere variant of abolitionism.

5. CONCLUSION

I have argued that three of the current four revisionary solutions to the now-what problem—namely expressivism, naturalism, and content fictionalism—are actually mere variants of abolitionism. Because there is nothing especially moral about the attitudes and sentences with which their proponents advise us to replace our moral judgments and sentences, these recommendations are plainly compatible with the letter of abolitionism. Not only that: they are even faithful to its spirit, which recommends that we adopt new attitudes in place of our discarded moral beliefs. If all this is correct, then there are to this day only three distinct solutions to the now-what problem: abolitionism, conservationism, and force fictionalism.[47] And the latter is the only revisionary option.

---

46   Svavarsdóttir, "Moral Cognitivism and Motivation," 167.

47   Or only two, if conservationism reduces to force fictionalism, as has been argued elsewhere (Jaquet and Naar, "Moral Beliefs for the Error Theorist?").

In the present contribution, I did not defend a particular solution to the now-what problem. Mine was merely a claim *about* the now-what debate. Still, I believe the claim in question has interesting implications *in* this debate. For if expressivism, naturalism, and content fictionalism are ultimately variants of abolitionism, then they must face the same objections it faces. For example, it has been argued by both conservationists and force fictionalists that proper moral discourse and thought bring with them advantages that we would have to dispense with should we abolish morality. Thus, moral judgments help us bolster self-control and prevent short-sighted rationalizations.[48] According to some, moral thought and discourse are so central to our psychology and practices that their abolition would be virtually impossible to implement.[49] If I am correct that expressivism, naturalism, and content fictionalism are variants of abolitionism, and that they cannot salvage proper moral thought and discourse, then one can expect them to face these worries no less than self-proclaimed abolitionists do.[50]

University of Montreal
francois.jaquet@umontreal.ca

REFERENCES

Blackburn, Simon. *Spreading the Word: Grounding in the Philosophy of Language*. New York: Oxford University Press, 1984.

Brandt, Richard. "A Theory of the Right and the Good." Amherst, NY: Prometheus, 1979.

Field, Hartry. *Realism, Mathematics, and Modality*. Oxford: Blackwell, 1989.

Garner, Richard. "Abolishing Morality." *Ethical Theory and Moral Practice* 10, no. 5 (November 2007): 499–513.

Hale, Bob. "Can Arboreal Knotwork Help Blackburn out of Frege's Abyss?" *Philosophy and Phenomenological Research* 65, no. 1 (July 2002): 144–49.

---

48 Joyce, *The Myth of Morality*, 215; Olson, *Moral Error Theory*, 195.

49 Nolan et al., "Moral Fictionalism versus the Rest," 307; Olson, *Moral Error Theory*, 180–81.

Hinckfuss, Ian. *The Moral Society: Its Structure and Effects*. Canberra: Australian National University, 1987.

Husi, Stan. "Against Moral Fictionalism." *Journal of Moral Philosophy* 11, no. 1 (January 2014): 80–96.

Ingram, Stephen. "After Moral Error Theory, After Moral Realism." *Southern Journal of Philosophy* 53, no. 2 (June 2015): 227–48.

Jaquet, François, and Hichem Naar. "Moral Beliefs for the Error Theorist?" *Ethical Theory and Moral Practice* 19, no. 1 (February 2016): 193–207.

Joyce, Richard. *The Evolution of Morality*. Cambridge, MA: MIT Press, 2007.

———. "Moral Anti-Realism." *Stanford Encyclopedia of Philosophy* (Winter 2016). https://plato.stanford.edu/entries/moral-anti-realism/.

———. *The Myth of Morality*. New York: Cambridge University Press, 2001.

Kalf, Wouter. *Moral Error Theory*. London: Palgrave Macmillan, 2018.

Lutz, Matt. "The 'Now What' Problem for Error Theory." *Philosophical Studies* 171, no. 2 (November 2014): 351–71.

Mackie, J. L. *Ethics: Inventing Right and Wrong*. New York: Penguin, 1977.

Nolan, Daniel, Greg Restall, and Caroline West. "Moral Fictionalism versus the Rest." *Australasian Journal of Philosophy* 83, no. 3 (2005): 307–30.

Olson, Jonas. "Getting Real about Moral Fictionalism." *Oxford Studies in Metaethics*, vol. 6, edited by Russ Shafer-Landau, 181–204. Oxford: Oxford University Press, 2011.

———. *Moral Error Theory: History, Critique, Defence*. Oxford: Oxford University Press, 2014.

Railton, Peter. "Moral Realism." *Philosophical Review* 95, no. 2 (April 1986): 163–207.

Streumer, Bart. *Unbelievable Errors: An Error Theory about All Normative Judgements*. Oxford: Oxford University Press, 2017.

Svavarsdóttir, Sigrun. "Moral Cognitivism and Motivation." *Philosophical Review* 108, no. 2 (April 1999): 161–219.

Svoboda, Toby. "Why Moral Error Theorists Should Become Revisionary Moral Expressivists." *Journal of Moral Philosophy* 14, no. 1 (December 2017): 48–72.

THE ELIGIBILITY OF RULE UTILITARIANISM

David Mokriski

T HE ELIGIBILITY THEORY OF MEANING (ETM), also known as the doc-
trine of reference magnetism, has played a significant role in recent dis-
cussions of metaphysics and philosophy more generally, including meta-
ethics. According to ETM, the referent of a predicate is the property that best
balances fit-with-usage, or "charity," and eligibility, where eligibility is a function
of how metaphysically natural the property is. ETM is motivated by its ability to
avoid intolerable levels of semantic indeterminacy and secure shared reference
between disputing parties. This sort of metasemantics has the potential to be
friendly toward somewhat revisionary theories—theories that do not fit so well
with some of our considered judgments—since the superior naturalness of a
candidate referent can outweigh some mismatch with usage. In this way, highly
natural properties act as "reference magnets," securing our reference despite ap-
parent counterexamples and otherwise less than optimal fit.

Using considerations of naturalness and eligibility, several philosophers have
recently argued for somewhat revisionary theories in epistemology, ontology,
and the metaphysics of truth.[1] In this paper, I add a similar argument to the
stock, applying these considerations to normative ethics. In particular, I argue
that the theory of rule utilitarianism (RU) achieves a high balance of charity and
eligibility. I will not argue that it achieves the *best* balance, relative to all possi-
ble (or popular) ethical theories, for that would be too ambitious. However, by
comparing RU to two of its common rivals, act utilitarianism (AU) and Rossian
pluralism (RP), I show how RU strikes a good balance between two extremes.
On the one hand, AU achieves a high degree of eligibility but only at a significant
cost of charity, while RP does the opposite, fitting very nicely with our consid-

---

1 Weatherson ("What Good Are Counterexamples?") defends the "justified true belief" the-
ory of knowledge against Gettier's apparent counterexamples in Gettier, "Is Justified True
Belief Knowledge?" Sider defends ontologists' ability to reject "common sense" views about
what objects exist ("Ontological Realism"), and Edwards defends "representational" theo-
ries of truth against apparent counterexamples that threaten their scope ("Naturalness, Rep-
resentation, and the Metaphysics of Truth").

ered judgments but at the price of low eligibility. A compromise between these factors would be preferable, and RU fits the bill, making it a promising theory.

My plan for the paper is as follows. In section 1, I introduce and motivate ETM and give a brief overview of metaphysical naturalness. In section 2, I give a rough account of the three rival theories in normative ethics that I will be comparing on grounds of charity and eligibility. In section 3, I take as my starting point the idea that we have some sort of a moral reason to "promote the good" and introduce five questions that must be addressed in order to clarify and precisify this thought. Each of these questions represents a dilemma for the theorist who endorses ETM, for one answer leads to a far more eligible theory like AU, while the other leads to a much more charitable one like RP. I then show how RU side-steps each dilemma, achieving a high degree of charity without sacrificing much eligibility. In section 4, I address some objections regarding whether RU is really as charitable as I claim. Finally, I conclude in section 5 with a brief discussion of the metaphilosophical costs of denying ETM or of downplaying the role of eligibility in metaethics.

1. ELIGIBILITY AND NATURALNESS

Meaning is not just a function of use. At least, this is the lesson that Lewis draws from Putnam's "model-theoretic argument" against metaphysical realism, Kripke's Wittgenstein-inspired semantic skepticism, and similar puzzles.[2] For any given term (e.g., "green"), there are far too many candidate referents that fit equally well with our usage (e.g., *being green, being grue*).[3] Likewise, for some terms (e.g., "gold"), there are bizarre candidates that may fit *better* with our usage

---

2   Lewis, "New Work for a Theory of Universals" and "Putnam's Paradox"; Putnam, "Realism and Reason"; Kripke, *Wittgenstein on Rules and Private Language*.

3   The predicate "grue" originated in Goodman, *Fact, Fiction, and Forecast*. An object is grue iff it is either green and discovered before some arbitrary future date (say, AD 3000) or blue and not so discovered. One might complain that *being grue* does *not* in fact fit with our usage of "green," since presumably fit-with-usage includes future and counterfactual usage as well as past and actual usage. Since, if prompted, we would probably say things like, "Green is the color of *all* emeralds, not only ones discovered before AD 3000," this may seem to disqualify *being grue* from being the referent of "green" on usage grounds alone. However, the problem is that there is a possible referent for "emerald," namely *gremerald*—where an object is a gremerald iff it is either an emerald and discovered before AD 3000 or a sapphire and not so discovered—and an interpretation that assigns *grue* to "green" and *gremerald* to "emerald" will fit with our usage as well as the intuitively intended one, at least when it comes to the linguistic dispositions mentioned so far. In general, it is always possible to take the intuitively intended interpretation and perform systematic permutations that yield bizarre (and intuitively unintended) interpretations that nevertheless fit with our usage equally well.

(e.g., *being gold-or-fool's-gold*) than their intuitively "intended" rivals (e.g., *being gold*).[4] As a proposed solution to these puzzles, ETM holds that reference is determined by *two* factors: how well a candidate referent fits with our usage of the term in question and the nature of the candidate referent itself.[5] In short, the referent of a predicate is the property that best balances the twin constraints of charity and eligibility.[6]

In addition to resolving indeterminacy, ETM has been put to work in securing shared reference between disputing parties, thus explaining the possibility of genuine disagreement despite diverging usage of a common term.[7] In metaethics, ETM has been proposed as a solution to the Moral Twin Earth challenge, offering an explanation of how our core moral term (e.g., "morally permissible") and the orthographically identical moral term of our "twins" on Twin Earth could refer to the same property, even if our patterns of usage are somewhat different (e.g., we are committed deontologists and our twins committed consequentialists).[8] If there is one highly natural property in the vicinity, then it would serve as a "reference magnet" and secure shared reference despite our somewhat different usages.[9] The merit of this response lies in the fact that it follows from an independently plausible, *general* metasemantics, solving a problem in metaethics that has plagued some versions of moral realism for decades.[10]

It is worth here clarifying the metaphysics that is presupposed by ETM, namely the distinction between natural and unnatural properties. Metaphysical natu-

---

4   Sider, "Criteria of Personal Identity and the Limits of Conceptual Analysis," 191.

5   For the most comprehensive discussions of ETM, see Lewis, "New Work for a Theory of Universals," 370, and "Putnam's Paradox"; and Sider, *Writing the Book of the World*, sec. 3.2. For arguments that the eligibility constraint is independently motivated and hence no mere *ad hoc* solution to the semantic puzzles, see Williams, "Eligibility and Inscrutability," 371; and Sider, *Writing the Book of the World*, 28.

6   Weatherson argues that this statement of ETM is an oversimplification but still a useful heuristic ("The Role of Naturalness in Lewis's Theory of Meaning"). For some problems with the simple account and suggestions on how to supplement it, see Williams, "Eligibility and Inscrutability"; and Hawthorne, "Craziness and Metasemantics."

7   Weatherson, "What Good Are Counterexamples?" 7–8; Sider, "Ontological Realism," sec. 11.

8   For the problem, see Horgan and Timmons, "New Wave Moral Realism Meets Moral Twin Earth." For the solution that appeals to ETM, see van Roojen, "Knowing Enough to Disagree"; Edwards, "The Eligibility of Ethical Naturalism"; and Dunaway and McPherson, "Reference Magnetism as a Solution to the Moral Twin Earth Problem."

9   In other words, even though distinct properties fit best with each community's respective usage, the same highly natural property would achieve the best *balance* of fit-with-usage and naturalness for both.

10   Dunaway and McPherson, "Reference Magnetism as a Solution to the Moral Twin Earth Problem," 641.

ralness is the gradation of properties exemplified by the pair *being green* and *being grue*; the property *being green* is more natural—less "gerrymandered"—than the property *being grue*.[11] The naturalness dimension ranges from the *perfectly natural* to the hopelessly gruesome. The perfectly natural properties are the fundamental ones, and this binary distinction of perfect naturalness is used to analyze the scalar notion of comparative naturalness on the traditional Lewisian view.[12] Every (less than perfectly natural) property has a canonical definition—a definition in terms of the perfectly natural properties (and logical operators)—and one property is more natural than another to the extent that the canonical definition of the former is less gerrymandered than that of the latter. Factors that contribute to gerrymanderedness include length, complexity, and the miscellaneousness of its constituents.[13] For example, the canonical definition of *being green* is intuitively much less gerrymandered than that of *being grue* (e.g., *being green-and-discovered-before-AD-3000-or-blue-and-not-so-discovered*); the latter is longer and more complex, and its constituents are more miscellaneous. Note that this difference in gerrymanderedness is highly plausible even given our inability to produce the full canonical definitions of such properties—we need not know the basis of *being green* in fundamental reality in order to know that it is a more natural property than *being grue*. We will return to this point soon.

Many theorists, even those otherwise sympathetic toward the appeal to naturalness in philosophy, have rejected the traditional Lewisian account of comparative naturalness, typically in favor of primitive degrees of naturalness.[14] However, I think many of the standard objections are overstated and the traditional account is worth maintaining. First, the traditional account is reductive, analyzing comparative naturalness in terms of fundamentality, a notion that many already have use for in philosophy. Second, the traditional account captures the paradigm examples of differences in comparative naturalness very well (e.g., *green* versus *grue*)—unnatural properties always seem like "merely arbitrary constructions" compared to their more natural counterparts.[15] Third, the traditional account does not leave our comparative naturalness judgments

---

11 For the most systematic discussions of metaphysical naturalness, see Lewis, "New Work for a Theory of Universals"; Sider, *Writing the Book of the World*; and Dorr and Hawthorne, "Naturalness."

12 Lewis, "Putnam's Paradox," 228.

13 Guigon, "Overall Similarity, Natural Properties, and Paraphrases," 8.

14 For objections, see Williams, "Eligibility and Inscrutability"; and Hawthorne, "Craziness and Metasemantics." Dunaway and McPherson are among those who opt for primitive degrees of naturalness ("Reference Magnetism as a Solution to the Moral Twin Earth Problem").

15 Hirsch, *Dividing Reality*, 55.

unconstrained the way primitive degrees of naturalness seem to.[16] On the primitivist view, different theorists are bound to find their preferred properties to be more natural than those of their rivals, whereas on the traditional view, there is pressure to converge on naturalness judgments insofar as our judgments of gerrymanderedness converge.

In spite of these benefits, many have objected that the traditional view is inadequate. The main objection is the worry that many properties of interest only have canonical definitions that are infinitely long, and there is no way to distinguish the comparative naturalness of such properties on the traditional view.[17] However, as Guigon notes, this objection ignores the fact that length is not the only factor that contributes to gerrymanderedness.[18] Even if two properties have canonical definitions that are both infinitely long, one may be more complex (e.g., be a conjunction of disjunctions rather than just an extended disjunction) or one may have more miscellaneous constituents. Furthermore, this objection is predicated on the idea that infinite canonical definitions of interesting properties are common. However, this seems to assume a sort of hyper-microphysicalist view according to which the only perfectly natural properties are certain microphysical ones, and all other properties are only definable as infinite disjunctions of realizations in microphysical terms. But why should we think the only canonical definition of, say, *being a person* is of the form *being P_1 or P_2 or ...,* where each P_i is a complete description of a possible person in microphysical terms? As Sider discusses, many of these properties are more plausibly defined in finite functional terms.[19] Likewise, we may have good reasons to countenance some properties at levels other than just the microphysical as perfectly natural.[20]

Even if one grants the traditional Lewisian view is correct, one might complain that it leaves us completely in the dark about the comparative naturalness facts, since we rarely know any canonical definitions in full detail. However, as

16 Dunaway and McPherson, "Reference Magnetism as a Solution to the Moral Twin Earth Problem," 653.

17 Another common objection is the worry that the traditional view yields the wrong verdict about "reasonably" natural properties like *being green* or *being a person*, since such properties are plausibly wildly complex when spelled out in terms of the fundamental properties. However, we must keep in mind that naturalness is a *comparative* matter, so when we think of properties like *being green* or *being a person* as "reasonably" natural, this is because they are far simpler than surrounding properties like *being grue* or *being a person-not-born-on-a-Tuesday*. It does not matter that such "reasonably" natural properties are much more complex, and hence much less natural, than fundamental properties like *spin* or *mass*.

18 Guigon, "Overall Similarity, Natural Properties, and Paraphrases," 7–9.

19 Sider, *Writing the Book of the World*, 130.

20 For discussion, see Schaffer, "Two Conceptions of Sparse Properties."

I mentioned above, we can typically gauge the gerrymanderedness of such canonical definitions even when we only have access to the ordinary-language definitions. For instance, we can tell that the canonical definition of *being red* is less gerrymandered than that of *being green-and-discovered-before-AD-3000-or-blue-and-not-so-discovered*, even without having access to either of these canonical definitions. This is because the canonical definition of *being red*, however gerrymandered it might be, is intuitively on par with those of *being green* and *being blue*, and the former is therefore much less gerrymandered than a complex and arbitrary construction involving both of the latter as well as the further complications involving *being discovered*, AD 3000, and the Boolean operators.

This suggests the following epistemology of comparative naturalness: to the extent that the ordinary-language definition of a property is more complex and arbitrary than another, this is evidence that the former property is less natural than the latter.[21] After all, if most of the terms that occur in two such ordinary-language definitions denote properties that are roughly on par with respect to naturalness, then any difference in the gerrymanderedness of these two definitions will roughly track a difference in the gerrymanderedness of their corresponding canonical definitions. In general, the more gerrymandered the ordinary-language definition of a property is, the more gerrymandered its canonical definition will be, since the canonical definition is typically obtained by taking the ordinary-language definition and replacing its terms with the canonical definitions of their referents. For example, the canonical definition of *being gricular* (i.e., *being green or circular*) is plausibly "*being G or C*," where *G* and *C* are the canonical definitions of *being green* and *being circular*, respectively.

In my argument for RU, I will appeal to this methodology quite a bit, gauging the naturalness of a property by how gerrymandered its ordinary-language definition is. The definition of *permissible_{AU}* is highly simple and nonarbitrary, but AU does severe damage to our moral intuitions. RP, on the other hand, fits very nicely with our moral intuitions, but the definition of *permissible_{RP}* ends up being highly complex and arbitrary. Compared to AU and RP, RU achieves a nice balance; the definition of *permissible_{RU}* is moderately simple and nonarbitrary, while achieving a moderately high degree of fit with our considered judgments. I will soon make the case for this in greater detail.

It is worth briefly pausing here to outline the package of assumptions about naturalness that I made in this section and on which the main argument of this

---

21 This methodology fails when our language is of the bizarre sort discussed in Hirsch (*Dividing Reality*), in which there are primitive terms like "gricular" for unnatural properties like *being green or circular* and complex expressions for more natural properties like *being green*. However, ordinary languages are typically not like this.

paper depends.[22] While I think there is good motivation for each, I also grant that they are highly controversial, and opponents of my conclusion would not be in bad company if they ended up denying them. These assumptions are:

N1. The comparative naturalness of a property is a function of how gerry-mandered its canonical definition is.

N2. The degree of gerrymanderedness of such a definition is a function of length but also other factors such as complexity and the miscella-neousness of its constituents.

N3. The perfectly natural properties that figure in such canonical defini-tions are not necessarily limited to microphysical properties.

N4. The degree of gerrymanderedness of a property's ordinary-language definition is (typically) a good guide to the degree of gerryman-deredness of that property's canonical definition.

Theorists who think we should ultimately reject one or more of these assump-tions may still find it valuable to see their implications for normative theory on a metaethics that includes ETM. Before I defend these implications, I will give a brief, and somewhat rough, overview of each theory I will be comparing.

2. THREE RIVALS IN NORMATIVE ETHICS

AU, RP, and RU obviously do not exhaust the options in normative ethics. Yet they make for useful comparisons, especially as different ways of developing a theory of normative ethics from the plausible starting point that, other things being equal, it is in some sense morally preferable to make the world better for all.[23] I will interpret each view as primarily a theory of moral *permissibility*, though much of what I say could instead be put in terms of moral *rightness*, what we morally *should* or *ought* to do, or what we have most moral *reason* to do.

Concerning metaethics, I will assume a naturalist or reductive view, accord-ing to which moral properties are identical to naturalistic properties of the sort that are in principle investigable by natural science (e.g., *being an action that maximizes overall happiness*).[24] I make this assumption because on the non-re-

---

22 Thanks to an anonymous reviewer for this suggestion.

23 Hooker, for instance, often defends his theory of rule consequentialism by making compari-sons with RP, as well as with act consequentialism ("Ross-Style Pluralism versus Rule-Con-sequentialism"; *Ideal Code, Real World*). Similarly, Sidgwick's discussion of the "methods" of ethics involves a systematic comparison of utilitarianism (and egoism) with RP, using the label "intuitionism" for the latter (*Methods of Ethics*).

24 Naturalists include Railton, "Moral Realism"; Boyd, "How to Be a Moral Realist"; and Brink, *Moral Realism and the Foundation of Ethics*.

ductive view, on which moral properties are held to be fundamental, there is
no difference in eligibility between AU, RU, RP, or any other first-order ethical
theory; on all such theories, the moral property is perfectly natural, no matter
how unnatural the naturalistic property is on which the purported fundamen-
tal moral property supervenes. Furthermore, I assume that the naturalness of a
moral property that is identical to naturalistic property N is determined by the
canonical definition of N. This rules out the possibility of a naturalistic property
with a highly gerrymandered canonical definition turning out to be highly natu-
ral simply because it is identical to a moral property.[25]

Finally, I also assume that the reductive view in question will identify the
moral property with the naturalistic property that is described by the substan-
tive first-order theory (e.g., *being an action that maximizes utility, respects auton-
omy*) rather than some sort of a response-dependent construction (e.g., *being
permitted by the moral framework that an idealized subject would endorse*).[26] On
such a response-dependent view, the naturalness of the moral property would
be settled by the canonical definition of this response-dependent property itself
and would be independent of the content of the first-order theory it happens
to track, and so all first-order theories would once again be tied for eligibility.
As a matter of fact, given my argument in this paper, I think the prospects for
maintaining the first-order theory of RP are most promising if we accept a re-
sponse-dependent reduction of the sort described above. This could give us the
plausible first-order consequences of RP without its otherwise poor degree of
eligibility.

2.1. Act Utilitarianism

According to one characterization of AU, an action is morally permissible if and
only if it leads to at least as much overall well-being as any other available ac-
tion.[27] Well-being is typically characterized as the balance of pleasure over pain,
preference satisfaction, or else some combination of features that includes these
as well as things like autonomy, friendship, accomplishments, etc.[28] Although

---

25 Thanks to an anonymous reviewer on this point.

26 Thanks to an anonymous reviewer for encouraging me to make this assumption explicit.

27 Proponents of AU include Bentham, *Introduction to the Principles of Morals and Legislation*;
Mill, "Utilitarianism"; Sidgwick, *Methods of Ethics*; Singer, "Is Act-Utilitarianism Self-De-
feating?"; Smart, "An Outline of a System of Utilitarian Ethics"; and, more controversially,
Hare, *Moral Thinking*. It may be unfair to characterize some of these earlier authors as pro-
ponents of AU over RU, given that this distinction was not made explicitly during their time,
and some of the things they say are open to interpretation.

28 Proponents of the hedonistic view of pleasure over pain include Bentham, *Introduction to
the Principles of Morals and Legislation*; Mill, "Utilitarianism"; and Sidgwick, *Methods of*

I characterized AU in terms of *actual* well-being produced, one could also hold that it is *expected* well-being that matters (i.e., for each possible outcome, take the well-being that would be produced and multiply it by the probability of that outcome occurring, and then sum these). There are a number of other possible variations on the view, but the characterization above should do for my purposes.

To a moral theorist who takes a fundamental "first principle" approach to ethics, in which one searches for a highly simple, general principle, AU may be somewhat plausible. However, there are several obvious criticisms of the view. Arguably the most common one is that AU is easily "counterexampled" by finding intuitively immoral actions that AU prescribes. For instance, if an instance of murder (theft, torture, etc.) maximizes overall well-being—and it is not at all difficult to come up with cases like this—then AU says it is permissible. Other criticisms include that AU is too demanding in what it requires of you, too simplistic in its account of what is valuable, leaves no room for personal projects and relationships, and is not sensitive to how well-being is distributed. Proponents of AU have a number of responses to these worries, but this at least demonstrates a *prima facie* conflict between AU and our considered moral judgments.[29]

2.2. Rossian Pluralism (aka "Commonsense Morality")

RP, unlike AU, eschews the idea of a single fundamental principle underlying moral permissibility.[30] For this reason, it is difficult to give a concise statement of the view. However, the rough version is this: an action is morally permissible if and only if we do *not* have an *all-things-considered* duty not to do it, where this is determined by the balance of *pro tanto* duties, of which there is an irreducible plurality.[31] These duties can conflict and, for the resolution of such conflicts,

Ethics. For the preference-satisfaction account, see Smart, "An Outline of a System of Utilitarian Ethics"; and Harsanyi, "Rule Utilitarianism and Decision Theory." If Hare counts as a utilitarian, then he is also of the preference-satisfaction variety (*Moral Thinking*). This pluralistic account of well-being is often called the "objective list view." Proponents include Parfit, *Reasons and Persons*; and Hooker, *Ideal Code, Real World*.

29 For a critique of AU—in particular, for how badly it fits with our strongest moral convictions—see Williams, "A Critique of Utilitarianism"; Harsanyi, "Rule Utilitarianism and Decision Theory," 31; and Scheffler, *Consequentialism and Its Critics*, 1–13.

30 For the *locus classicus* of such a view, see Ross, *The Right and the Good*. Other proponents of RP include Nagel, "Fragmentation of Value"; and Audi, *The Good in the Right*.

31 Ross (*The Right and the Good*) unfortunately chose to label these as "*prima facie* duties," which suggests that they merely *seem* to be duties but may end up being morally irrelevant upon reflection. It is widely acknowledged that they are more accurately called "*pro tanto* duties," since they are considerations that count in favor of there being a duty proper but that might be outweighed by countervailing considerations. Kagan, *The Limits of Morality*, 17n13; Hooker, "Ross-Style Pluralism versus Rule-Consequentialism," 534n6.

are assigned specific weights (rather than being arranged in a lexical hierarchy). Furthermore, there is no fundamental principle governing the relative weights of the several principles to which we can appeal; rather, we must exercise opinion or judgment as to which *pro tanto* duties are defeated in which cases.[32] The *pro tanto* duties include, among other things, duties to keep one's promises, not to harm (steal, lie, etc.), and to promote well-being (giving priority to those to whom we bear special connections).

When compared to AU, the merits of RP should be obvious. Such a view, often described as the best regimentation of "commonsense morality," is very difficult, if not impossible, to "counterexample."[33] Anytime you construct a case that intuitively has some moral status that is not entailed by the current version of the theory, you can simply add a new *pro tanto* duty to your list that does cover it. Likewise, anytime several *pro tanto* duties conflict, this can be resolved by assigning the greatest weight to the one that intuitively should take precedence. However, there are some common criticisms of RP that are unrelated to how well it handles specific cases. The main charge, unsurprisingly, is that it is not systematic enough.[34] It paints morality as a "heap of unconnected duties," and its lack of an underlying principle governing which sort of actions are duties and their relative weights renders the view rather unexplanatory and unsatisfying from a theoretical perspective.[35] Likewise, the appeal to "judgment" for the purpose of resolving conflicts, rather than a principled explanation of the relative weights, is un-illuminating and perhaps even *ad hoc*. It is more like an evasion of the problem than a method.[36] If our intuitions are silent about the relative strengths of two competing *pro tanto* duties, there is simply no way to find out the truth of the matter.

2.3. Rule Utilitarianism

According to the version of RU that I favor, an action is morally permissible if and only if it is permitted by the rules whose general internalization has the great-

---

32 Ross, *The Right and the Good*, 19; Hooker, *Ideal Code, Real World*, 534.

33 Skelton, "William David Ross."

34 Joseph, *Some Problems in Ethics*; Rawls, *A Theory of Justice*; Skelton, "William David Ross."

35 McNaughton, "An Unconnected Heap of Duties?" 434. Ross, anticipating some of these "theoretical" worries, writes, "Loyalty to the facts is worth more than a symmetrical architectonic or a hastily reached simplicity" (*The Right and the Good*, 23). Of course, in order to avoid begging the question, he should replace the term "facts" with the term "evidence" or "considered moral judgments." However, if ETM is correct, then considerations of simplicity (and nonarbitrariness) must be included in our total evidence alongside our considered moral judgments.

36 Hare, *Moral Thinking*, 34.

est expected value in terms of overall well-being.[37] The notion of *internalization* here is meant to be distinct from that of full compliance; rather it is a matter of *accepting* the rules as the shared moral code in one's community, allowing for the possibility of failing to always live up to it. The expected value of the internalization of rules takes into account the costs of both implementation and maintenance.[38] There are more details that I am unable to currently fill in, so in this paper I will leave the account of RU as roughly stated above, while acknowledging that it is greatly in need of refinement.[39] Finally, RU also owes us some account of well-being, and the options are the same as they were for AU.

The attraction of RU—indeed its main motivation for those otherwise sympathetic toward utilitarianism or consequentialism in general—is that it can (allegedly) avoid common objections to AU while staying relatively systematic. I will make the case for this in more detail in section 3.2, albeit in terms of charity and eligibility. For now, I will mention the most common criticisms.

The greatest objection to RU, which is largely responsible for its unfavorable reputation among moral theorists in general, takes the form of a dilemma: either RU "collapses" into AU, in which case it loses its distinctiveness and gains all the problems associated with the latter, or else it is guilty of "incoherence," which could be understood as anything from logical inconsistency to being severely unmotivated as a version of consequentialism.[40] Roughly, the thought is that either the optimal set of rules prescribed by RU includes only the single rule of AU (i.e., "Choose the optimal action"), or else the rules tell you to take sub-

---

37 Proponents of RU include Harrod, "Utilitarianism Revised"; Harsanyi, "Rule Utilitarianism and Decision Theory"; and Brandt, *A Theory of the Good and the Right*. Harsanyi ("Rule Utilitarianism and Decision Theory," 32) credits the idea behind RU to Harrod ("Utilitarianism Revised") and the terms "act utilitarianism" and "rule utilitarianism" to Brandt (*Ethical Theory*, 380, 396). My characterization here is close to Hooker's formulation of his view, with some key differences (*Ideal Code, Real World*, 32). First and foremost, he is a rule *consequentialist* rather than a proponent of the narrower RU view. This is because he includes a principle of distribution, namely some priority for the worse off, rather than being merely aggregative, although he makes this choice rather tentatively (*Ideal Code, Real World*, 65). Also, he includes a clause favoring rules closer to conventional morality as a tiebreaker between otherwise optimal rules.

38 Hooker, *Ideal Code, Real World*.

39 I acknowledge the possibility that upon attempting to address such questions, RU may lose some of the eligibility that it appears to have. However, even once these details are filled in, it will still be far more eligible than RP.

40 On the idea that RU collapses into AU, see Lyons, who defends this thesis of the extension equivalence of AU and RU (*Forms and Limits of Utilitarianism*). On the idea that proponents of RU are guilty of "rule worship," a cardinal sin among consequentialists, see Smart, "An Outline of a System of Utilitarian Ethics," 10; and Kagan, *Normative Ethics*, 230.

optimal actions (even when you know they are suboptimal), which seems to abandon the spirit of consequentialism. Proponents of RU, or similar theories, have convincingly addressed how the first horn of the dilemma can be avoided, as suggested by the formulation of RU in terms of internalization rather than compliance—the general internalization of the single rule of AU simply would not have very good consequences.[41] The second horn, however, may still be worrisome, but it is beyond the scope of this paper to address it.[42]

3. THE STARTING POINT AND THE FIVE QUESTIONS

I will begin with the somewhat vague idea that we are morally required to "promote the good" and then will address several questions that arise when attempting to develop this idea with more clarity and precision. Readers who do not find such a starting point plausible are welcome to read the rest of this paper conditionally—that is, *if* you start out with this general idea about ethics, *then* considerations of charity and eligibility should move you toward RU. An alternative starting point for those whose moral sympathies are less consequentialist could be the vague idea that we are morally required to "respect persons," and I think a similar line of argument to the one in this paper could be made that leads from there to a view like contractualism.[43] Since other moral theories may be able to achieve a similarly good balance of charity and eligibility, proponents of such views still have much to gain from this discussion.

3.1. *AU versus RP on the Five Questions*

I will now introduce five questions that must be addressed in order to develop a moral theory from our starting point, the idea that we have a moral reason to promote the good. The first questions is:

> Q1. Are there some things we may *not* do while promoting the good?

This is the question of whether there are, in the terminology of Kagan, moral *constraints*, or prohibitions on certain types of actions even when they have good consequences.[44] The second question is:

> Q2. Are there some limits as to how much good we must promote?

---

41   Hooker, *Ideal Code, Real World*, 94.

42   For an answer to the charge of incoherence, see Hooker, *Ideal Code, Real World*, sec. 4.3.

43   For an overview of contractualism, see Ashford and Mulgan, "Contractualism." For a prominent contractualist view, see Scanlon, *What We Owe to Each Other*.

44   Kagan, *The Limits of Morality*, 4.

This is the question of whether there are, again in the terminology of Kagan, moral *options*, or limits on how demanding morality is, which give us room to pursue our own aims in life.[45] The third question is:

> Q3. Is there anything worth promoting for its own sake other than well-being?

This is the question of the appropriate aims of morality. The fourth question is:

> Q4. Can we favor those closest to us while promoting the good?

This is the question of whether morality permits some partiality toward those to whom we bear special relationships, or whether we must be completely impartial. Finally, the fifth question is:

> Q5. While promoting the good, does it matter how well-being is distributed?

This question asks whether it is just the total amount of good promoted that matters, or whether it matters who gets it.

AU gives a negative answer to each of these five questions, which makes it highly counterintuitive. First, there are no moral constraints; anything that best promotes the good is permissible (and indeed required), no matter what intuitively immoral actions it involves (e.g., stealing, breaking promises, killing). Second, there are no moral options; we are required to *maximize* the good, which leaves us little to no room for any personal projects in life. No matter how generous we are with our time or money, we are still required to do more since there is always more good we can do. Third, there are no moral aims other than well-being. Everything else should only be pursued as a mere means to well-being, including intuitively valuable things such as knowledge, virtue, meaningful relationships, and personal accomplishments. Fourth, we are not allowed any partiality toward those closest to us. We must be completely impartial in our good-promoting actions, giving absolutely equal consideration to the well-being of all, strangers and loved ones alike. Finally, the distribution of well-being does not matter; all that matters is that we produce as much well-being as possible, no matter how unequally it is distributed or whether it makes its way to those deserving or undeserving. Given all these implications, AU conflicts greatly with some of our strongest moral convictions, making it an extremely uncharitable theory.

RP, on the other hand, gives an affirmative answer to each of these questions, which makes it highly intuitive. First, there are moral constraints, since there

45 Kagan, *The Limits of Morality*, 3.

are *pro tanto* duties other than our duty to promote the good. We must, in ordinary circumstances, keep our promises, not steal, not harm, etc., and only when there are strong enough countervailing considerations are we permitted to violate these constraints. Second, there are moral options, since the *pro tanto* duty to promote the good, according to RP, is plausibly interpreted in terms of *satisficing*, or doing "enough" good, rather than maximizing.[46] As long as you have done enough to satisfy the duty of beneficence, you have freedom to pursue your own projects. Third, there are aims to promote other than well-being, since in addition to countenancing a plurality of duties, RP is also pluralistic about the good. We should aim to promote well-being, but we should also aim to promote knowledge, virtue, and various other intuitively valuable goods. Fourth, we are allowed some partiality toward those to whom we bear special relationships. We can favor the well-being of ourselves and our loved ones, provided that we not disregard others' interests entirely. Finally, the distribution of well-being matters. Two plausible candidates for distribution principles include one according to desert—for instance, the "allocation of pleasure to the virtuous"—and priority toward the worse off.[47] Given these implications, RP fits very well with our moral convictions and is hence an extremely charitable theory.

Unfortunately, the high degree of charity that RP achieves comes at a very steep price in terms of eligibility. To demonstrate this, I will discuss how each affirmative answer RP gives affects the definition of *permissible*$_{RP}$. First, we begin with the following simple definition:

[RP1] *being an action that promotes the good*

Now, adding in moral constraints, we get:

[RP2] *being an action that promotes the good without C_1 or C_2 or … or C_n*

where each C_i corresponds to a constraint-violating act-type (e.g., perhaps C_2 is *breaking a promise*). However, recall that RP's constraints are non-absolute; most constraints have a set of exception clauses for when they are outweighed by stronger moral considerations. Thus, we must add a layer of complication to the definition:

[RP3] *being an action that promotes the good without [C_1 and not ($E_{1,1}$ or $E_{1,2}$ or … or $E_{1,h}$)] or [C_2 and not ($E_{2,1}$ or $E_{2,2}$ or … or $E_{2,k}$)] or … or [C_n and not ($E_{n,1}$ or $E_{n,2}$ or … or $E_{n,m}$)]*

where each $E_{i,j}$ is the *j*th exception to the *i*th constraint. For instance, if C_2 is

46 On satisficing, see Slote and Pettit, "Satisficing Consequentialism."

47 Ross, *The Right and the Good*, 140.

breaking a promise, $E_{2,3}$ might be *saving someone from severe distress*. Next, we must account for moral options, by interpreting "promotes" as *satisfices*:

> [RP4] *being an action that leads to at least quantity Q of good without $[C_1$ and not $(E_{1,1}$ or $E_{1,2}$ or … or $E_{1,h})]$ or $[C_2$ and not $(E_{2,1}$ or $E_{2,2}$ or … or $E_{2,k})]$ or … or $[C_n$ and not $(E_{n,1}$ or $E_{n,2}$ or … or $E_{n,m})]$*

where Q is some complete specification of what counts as "enough" good. Next, since RP is pluralistic about the good, we must fully specify the several components of the good and assign relative weights to each to account for trade-offs (e.g., to determine how much knowledge is worth promoting over a certain amount of well-being):

> [RP5] *being an action that leads to at least quantity Q of the sum: $(W_1G_1 + W_2G_2 + … + W_vG_v)$ without $[C_1$ and not $(E_{1,1}$ or $E_{1,2}$ or … or $E_{1,h})]$ or $[C_2$ and not $(E_{2,1}$ or $E_{2,2}$ or … or $E_{2,k})]$or … or $[C_n$ and not $(E_{n,1}$ or $E_{n,2}$ or … or $E_{n,m})]$*

where each G_i is a quantity of one of the goods worth promoting for its own sake and each W_i is its appropriate weight. Next, we must specify the details of its partiality:

> [RP6] *being an action that leads to at least quantity Q of the sum: $(W_1G_1 + W_2G_2 + … + W_vG_v)$, giving those in group S_1 priority P_1, those in group S_2 priority P_2, …, and those in group S_w priority P_w, without $[C_1$ and not $(E_{1,1}$ or $E_{1,2}$ or … or $E_{1,h})]$ or $[C_2$ and not $(E_{2,1}$ or $E_{2,2}$ or … or $E_{2,k})]$ or … or $[C_n$ and not $(E_{n,1}$ or $E_{n,2}$ or … or $E_{n,m})]$*

where each S_i is a set of individuals related to the agent in a morally relevant way and each P_i is the appropriate weight of prioritizing the well-being of those in S_i. For instance, perhaps S_1 is the agent's singleton and P_1 is the greatest prioritizing weight, S_2 includes the agent's closest friends and family members and P_2 is the second greatest weight, etc. Finally, we must account for RP's distribution principles, leading to our completed definition:

> [RP] *being an action that [leads to at least quantity Q of the sum: $(W_1G_1 + W_2G_2 + … + W_vG_v)$, where the well-being of those in group S_1 is given priority P_1, …, and the well-being of those in group S_w is given priority P_w, giving weight D_1 to the deserving, D_2 to those worse off, …, and D_r to those with feature $F_r]$ without $[C_1$ and not $(E_{1,1}$ or $E_{1,2}$ or … or $E_{1,h})]$ or $[C_2$ and not $(E_{2,1}$ or $E_{2,2}$ or … or $E_{2,k})]$or … or $[C_n$ and not $(E_{n,1}$ or $E_{n,2}$ or … or $E_{n,m})]$*

where each D_i is the appropriate weight of distributing well-being in favor of those with feature F_i.

Now that we have gone through how RP's answers to the five questions affect the definition of *permissible*$_{RP}$, we can compare it to the definition of *permissible*$_{AU}$, which is something like the following:

[AU] *being an action that leads to as much overall well-being for all, equally considered, as possible*

Clearly [RP] is far longer and vastly more complex than [AU]. Even though [AU] references complex phenomena such as well-being, whose canonical definition may itself be relatively complex, [RP] does so as well, while also referencing a lot more. Furthermore, [RP] contains a significant amount of arbitrariness, since it assigns many arbitrary weights and includes many arbitrary exception clauses. For instance, why think that we are only required to promote quantity Q of good, rather than $Q + k$ or $Q − k$? Also, why think we may violate constraint C_2 only to prevent at least quantity Q' of bad, rather than $Q' + k$ or $Q' − k$? Countless properties are extremely similar to the one expressed by [RP] that differ only in assigning a slightly different value to one of these variables, or in swapping in or out some one or two exception clauses. Any choice between them will be completely arbitrary. The fact that [RP] is extremely complex and arbitrary suggests that *permissible*$_{RP}$ is extremely unnatural and much less natural than *permissible*$_{AU}$.

Let me briefly summarize. The theories AU and RP give opposite answers to each of the five questions, where one answer leads to a much simpler and less arbitrary but highly counterintuitive theory, while the other answer leads to a highly intuitive but extremely complex and arbitrary theory. Whether I got the details of *permissible*$_{RP}$ exactly correct is debatable, but that such a property will inevitably be extremely complex and arbitrary is not. Thus, if we begin at our starting point with a vague requirement to promote the good and address the five questions, considerations of charity and eligibility seem to pull strongly in opposite directions.

3.2. *How RU Answers the Five Questions*

I will now show how RU achieves a moderately high degree of both charity and eligibility. Before we see how RU fares in addressing the five questions, let us remind ourselves how it scores on eligibility. The definition of *permissible*$_{RU}$ looks something like:

[RU] *being an action that is permitted by the rules whose general internal-*

ization has the greatest expected value in terms of overall well-being for all, equally considered

As I admitted in section 2.3, a lot of details still need to be filled in, and doing so may lead to greater complexity and arbitrariness than is now apparent. However, I think it is clear that while [RU] is not as simple or nonarbitrary as [AU], it is far simpler and less arbitrary than [RP] and will continue to be even once these details are filled in. Thus, *permissible_{RU}* is much more natural than *permissible_{RP}*.

However, unlike AU, RU can maintain this relatively high degree of eligibility without sacrificing too much charity. This is because RU gives the intuitive answers to each of the five questions, the same affirmative answers that RP gives. First, RU will include moral constraints, since the rules whose internalization has the greatest expected utility (henceforth, the "optimal rules") will not simply consist of one rule that says "Maximize well-being," but will instead consist of a plurality of (plausibly non-absolute) rules such as "Keep your promises in ordinary circumstances" and "Don't harm an innocent person (except to prevent a disastrous outcome)." The general internalization of constraints against dishonesty, promise-breaking, and violating property rights in ordinary circumstances is necessary to secure trust and other beneficial expectation effects.[48] Likewise, constraints against harm help to avoid miscalculation and abuse, and constraints against free-riding produce beneficial coordination effects.[49] In general, there is great social utility in the general acceptance of constraints and their correlative rights.[50]

Second, RU will plausibly be much less demanding than AU, with its optimal rules leaving people with options to pursue their own aims and projects.[51] Whatever rule in the optimal rule set is associated with promoting well-being will plausibly be stated in terms of satisficing rather than maximizing. This is because the costs of getting a hyper-demanding rule (e.g., "Be altruistic to the point of diminishing marginal utility") internalized among the general population and maintaining it would be extremely high.[52] Even if people could be

---

48 Harsanyi, "Rule Utilitarianism and Decision Theory," 32–33; Hooker, *Ideal Code, Real World*, 77.

49 Kagan, *The Limits of Morality*, 33–34.

50 For an extensive discussion of how RU and similar indirect consequentialist views justify constraints, see Hooker, *Ideal Code, Real World*, ch. 6. For a thorough discussion of whether such constraints are plausibly absolute or non-absolute, also see Hooker, *Ideal Code, Real World*, sec. 6.4.

51 See Hooker, *Ideal Code, Real World*, ch. 8. For dissent regarding RU's ability to avoid excessive demandingness, see Kagan, *The Limits of Morality*, 35.

52 Hooker, *Ideal Code, Real World*, 78–79.

convinced of the moral authority of such a rule, which is dubious, they would constantly fail to live up to it and thereby alienate themselves from morality in general and perhaps other important moral rules in particular.[53] Ironically, it is plausible that people would end up being motivated to do *more* good if there's a less demanding rule concerning how much good they are required to do that leaves room for supererogatory action beyond that.[54] After all, sometimes you get more with honey than with vinegar.

Third, RU's rules governing good-promotion will not tell you to just aim at promoting well-being for its own sake but other things as well, including knowledge, virtue, and justice. This is an extension of the lesson drawn from the so-called paradox of hedonism, which is the observation that "adopting as one's exclusive ultimate end in life the pursuit of maximum happiness may well prevent one from having certain experiences or engaging in certain sorts of relationships or commitments that are among the greatest sources of happiness."[55] This sort of "paradox" can be generalized into what we might call the "paradox of welfarism"—in other words, adopting overall well-being as the only direct aim in our everyday lives will likely result in less overall well-being. This is because such an attitude would preclude us from aiming *directly* at things like accomplishments, scientific or philosophical discovery, meaningful relationships, and self-improvement; these other things would be treated as purely instrumental, worth pursuing only if our direct utility calculations yield the right verdict. Such a single-minded way of deliberating, apart from being wildly impractical, seems much less conducive to overall well-being than the alternative, namely pursuing a reasonable plurality of aims.

Fourth, RU would plausibly permit some degree of partiality, since internalizing practical rules that allow or even mandate some degree of partiality would have better consequences in terms of overall well-being. Given human psychology, there would be significant costs in attempting to get and keep fully impartial practical rules internalized.[56] Furthermore, there are certain benefits that can best be secured through partiality, including personal accomplishments, which require favoring your own interests, and meaningful relationships, which require favoring the interests of those close to you.[57] In general, overall well-being is better promoted when we follow rules that prescribe some degree of partiali-

53   Kagan, *The Limits of Morality*, 35.

54   Kagan, *Normative Ethics*, 225.

55   Railton, "Alienation, Consequentialism, and the Demands of Morality," 140.

56   Hooker, *Ideal Code, Real World*, 140.

57   Hooker, *Ideal Code, Real World*, 139.

ty toward ourselves and those closest to us, since those are the interests we are most familiar with and in the best position to affect.[58]

Finally, RU's rules governing good-promotion will plausibly take distribution into account. As Hooker explains, "utilitarians have to trade off the diminishing marginal utility of material goods against the need for economic incentives."[59] The former consideration calls for a distribution principle that gives priority to the worse off, while the latter calls for a distribution principle based on desert.[60] Thus the optimal rule set will plausibly require us to give priority to the worse off and the virtuous when promoting well-being, as such a practice is much more conducive to overall well-being than the alternative.

Let us take stock of how RU fares when addressing the five questions. Given that the affirmative answers RU gives are not *assumed* as part of the theory, like they are on RP, but are instead *derived* from the theory together with empirical considerations, RU is able to secure a degree of eligibility that far surpasses that of RP. It avoids numerous complications and countless arbitrary choices about exactly how to assign specific values and where to draw certain lines (between, for instance, cases that are exceptions to a certain constraint and those that are not). However, the (complex and arbitrary) contingent, empirical facts being what they are, the theory yields a highly complex set of practical rules that map onto our considered moral judgments in a fairly comfortable manner. Again, this fit will be far from perfect—for instance, RU may still be a bit more demanding than we expected morality to be—but there is a world of difference between the charity of RU and that of AU. Thus, if we begin our moral theorizing from the starting point of a vague requirement to promote the good and address the five questions in order to clarify and precisify this intuition, RU looks like a very promising moral theory, securing a nice balance of charity and eligibility.

---

58 Jackson makes this point in a particularly compelling way using his "crowd control" thought experiment ("Decision-Theoretic Consequentialism and the Nearest and Dearest Objection," 474).

59 Hooker, *Ideal Code, Real World*, 64.

60 An additional merit of this sort of approach is that it is consistent with a thoroughgoing skepticism about basic, non-consequentialist desert. See Caruso ("Skepticism about Moral Responsibility") for discussion. The sort of desert invoked at the practical level is merely consequentialist. The virtuous do not deserve a benefit in any deep sense that requires a controversial sort of moral responsibility or free will. Rather, they "deserve" it because treating them so is part of a useful practice, providing incentive effects that are conducive to overall well-being.

4. OBJECTIONS

In this section, I will consider and address three objections. They all concern whether RU is really as charitable as I suggest. Some of my responses will be rather concessive, granting in many cases that RU's fit with our considered judgments is far from perfect. However, when we keep ETM and considerations of eligibility in mind, I think RU's mismatch with usage is far from decisive.

4.1. The "Wrong" Rules and Lines in All the "Wrong" Places

First, one might grant that I have successfully shown that RU can make room for *certain* constraints, options, pluralistic aims, degrees of partiality, and principles of distribution, but not that I have shown that RU can get the *intuitively correct* ones. Given the contingent, empirical facts, the optimal constraints, weights, etc. might be somewhat different from what we intuitively think. For instance, RU's line that marks where satisficing ends, where the constraint against promise-breaking gives way to the exception for preventing distress, or where some quantity of virtue outweighs some quantity of well-being, might not be exactly where our moral intuitions want it to be. Thus RU may end up being *much* less charitable than RP, since the latter can take the constraints, weights, lines, etc. to be *exactly* as they intuitively seem (except when our intuitions are inconsistent), while the practical rules of RU are hostage to contingent, empirical facts.

In response to this objection, I will first point out that even getting *some* constraints, options, etc. is still a considerable achievement and puts RU light-years ahead of AU with respect to charity. The counterintuitiveness of AU is altogether a difference in kind, given that it has *no* constraints, options, etc., whereas the counterintuitiveness of RU's imperfect (with respect to our considered judgments) constraints and line placements is just a matter of degree. That there are *some* constraints, options, etc. may be close to a "Moorean fact" about moral permissibility, whereas we seem to be more open to revising exactly where we draw certain lines.

Second—and this will be somewhat of a recurring theme in my responses to objections—I expect the superior eligibility of RU (over RP) to pick up the slack wherever its charity falls short. After all, ETM, in contrast to a charity-only metasemantics, gives revisionary theories a chance of being true despite some mismatch with usage. I see RU as an instance of this general idea; the superior naturalness of *permissible*$_{RU}$ makes up for its less than perfect fit with our considered judgments (e.g., some lines in the intuitively wrong places).

4.2. *The Contingency of the Rules and Otherworldly "Counterexamples"*

Next, one might grant that RU can yield the intuitively correct (or close enough) practical rules in the *actual* world, given the *actual* facts about human psychology and our environment. However, when we consider other possible worlds with other such possible facts about agents and environments, the practical rules derived from RU may be drastically different and highly counterintuitive.[61] For instance, perhaps we can imagine alternative agent psychologies or laws of nature such that, in those worlds, internalizing rules that permit or mandate torture, theft, etc. has a very high expected value in terms of overall well-being. Thus RU, though it can avoid (most of) the easy, *this*-worldly counterexamples to AU, may still be subject to damning, *other*worldly counterexamples. After all, the constraints, weights, lines, etc. on RU, unlike on RP, are an entirely *contingent* matter.

In response, I think that not all counterexamples are created equal in terms of their theoretical import. When it comes to the metasemantic constraint of charity, fitting with our usage of the term over typical and familiar cases counts more in reference determination than fitting with our usage of the term over far-fetched and unfamiliar cases. Thus if there is a candidate referent that fits somewhat poorly with our dispositions to apply the term to extremely atypical cases, but otherwise fits very well, then considerations of charity should not disqualify it from being the referent of that term. Hence, the fact that RU can handle (most of) our considered judgments about typical, actual, and *nearby* possible cases gives it the degree of charity it needs to be a strong competitor in the battle of theory choice. Considerations of eligibility can take it the rest of the way.

4.3. *The Right Rules but for All the "Wrong" Reasons*

Finally, one may grant that RU does a good enough job at fitting our considered judgments about the practical rules of morality but then complain that it does a bad job at fitting our judgments about *why* those are the correct rules. Perhaps RU can correctly account for the fact that we are required to keep our promises, promote virtue, pursue meaningful relationships, etc., but its explanation for these facts may be highly counterintuitive. We typically think that promise-keeping, virtue, meaningful relationships, etc. are valuable *in themselves*, whereas RU holds that their value is derivative, wholly explained by their (indirect) relationship to the value of well-being. In general, if we have a constraint against performing

---

61 For a recipe for constructing some such (modally distant) counterexamples to RU or rule consequentialism in general, see Podgorski, "Wouldn't It Be Nice?" Thanks to an anonymous referee for this reference.

actions of a certain type, that's because it seems like there is something wrong with those actions *in themselves* rather than merely because they are prohibited by the optimal rules. Thus even if RU can yield the intuitively correct practical rules, it does so for the intuitively wrong reasons. The revisionary nature of RU's moral explanations means that the theory must take a hit with respect to charity.

I will give two responses to this objection, the first combative and the second concessive. First, even when we think about our judgments of moral explanations, it is not obvious that a view like RP is that much more charitable than RU. While RP, unlike RU, can agree with our intuitions that constraints, pluralistic aims, etc. have fundamental moral value, we also seem to have the intuition that there is some deeper, unifying explanation behind ordinary moral rules. This is evident from the fact that, long before the notions of naturalness or eligibility were anywhere on the scene, moral philosophers were complaining that views like RP are too unsystematic—mere "shopping lists" of disconnected principles and unexplained arbitrary weights. These two intuitions, that multiple sorts of things are fundamentally valuable and that there is a unifying explanation behind all of morality, seem to be in conflict. RP does justice to the former and RU the latter. Thus, even taking into account our convictions about moral explanations, there may not be as big a difference in charity between RP and RU as the objection suggests.

Second, even if the objection is correct that RU's moral explanations are somewhat revisionary, this does not mean we should not accept them. If we are realists about ethics, just like realists about anything, we should be open to somewhat surprising explanations behind ordinary phenomena. We have learned from modern science that there are all sorts of extremely surprising and counterintuitive explanations (e.g., atomic theory, quantum mechanics) behind the behavior of ordinary things at the macroscopic level. RU can be seen as just another instance of this general theme, albeit in the moral domain. Once again, charity is not the be-all and end-all; eligibility must be given its due weight.

5. CONCLUSION

Before I close, it is worth briefly discussing the costs of denying ETM in meta-ethics, or of downplaying the strength of the eligibility constraint to the point where a view like RP could end up achieving the best balance of charity and eligibility in spite of its low eligibility. If charity were given near full authority in the metasemantics of "morally permissible," then if there were to be two or more equally charitable interpretations, the term would be semantically inde-terminate between them. If our moral intuitions, together with those of the rest

of our linguistic community, were split or undecided on some matter—for instance, on the presence or absence of some particular constraint or the value of some particular weight—then there would be no fact of the matter as to what is morally permissible in cases that turn on this difference. Furthermore, if we were to encounter a moral theorist from an alternative linguistic community—or whose position could best be understood by reference to a corresponding hypothetical linguistic community—with different moral convictions about certain constraints or weights, then many of our disputes with her about what is morally permissible would be verbal.[62] Thus there are significant metaphilosophical costs of downplaying the role of eligibility.

However, it should be noted that even if we do adopt ETM there is no guarantee of securing shared reference for every linguistic community with a term that plays the permissibility role. For instance, a community of committed act utilitarians, whose usage of "morally permissible" aligns very closely with what AU entails, may refer to *permissible*$_{AU}$ after all, since it is this property that will best balance charity and eligibility in that community. This is, however, the exception that proves the rule. It is only because the usage of these act utilitarians is so vastly different from our own that we end up expressing distinct properties by our respective moral terms.[63] For any linguistic community whose usage is in the vicinity of what we consider "commonsense morality," shared reference will be secured to *permissible*$_{RU}$ due to its decent fit and high degree of eligibility. Thus most moral disputes will still come out as nonverbal.

My concession that ETM, under the assumptions about naturalness N1–N4 outlined in section 1, does not provide the strong *guarantee* of shared reference for all possible moral communities may seem to undermine the main motivation for ETM in metaethics, namely its use as a general solution to the Moral Twin Earth challenge. If this concession is too much for some theorists, then this

62 This is assuming Hirsch's ("Physical-Object Ontology, Verbal Disputes, and Common Sense") account of verbal disputes, which is motivated by its ability to remain faithful to Burge's ("Individualism and the Mental") social externalist insight, namely that what we mean is partly determined by the patterns of usage in our wider linguistic community (i.e., meaning is not a completely private matter). A dispute's being verbal on this account does not require that the disputants mean different things (since members of the same linguistic community typically speak a shared language) but only that the hypothetical linguistic communities with the parties' differing usages would mean different things.

63 Perhaps it could be argued that *permissible*$_{AU}$ is not in fact a candidate referent for "permissible," since it cannot play all of the permissibility role in our thought and discourse, which includes action guidance. If so, then perhaps *permissible*$_{RU}$ is the most natural candidate after all, in which case there may be more of a guarantee of shared reference due to reference magnetism, at least to the extent that RU achieves a high enough degree of charity for every possible community with a term that plays the permissibility role.

may give them more reason to reject one or more of N1–N4. However, I think securing a reasonable amount of shared reference—in particular, for all moral communities in the vicinity of "commonsense morality"—is motivation enough.

In this paper, I have argued that RU is a very promising theory if we adopt a metasemantics that includes ETM. On RU, the moral property comes out as fairly simple and nonarbitrary, especially when compared to views like RP. Since her moral property is relatively natural, the proponent of RU can reap the benefits of reference magnetism, which includes limiting semantic indeterminacy and securing shared reference between alternative linguistic communities with somewhat diverging usages, thus avoiding verbal disputes. Unlike its rival AU, RU secures this high degree of eligibility without sacrificing too much by way of charity. Hence, RU should be taken very seriously by any moral philosopher who aims to "carve nature at its joints."[64]

University of California, Santa Barbara
dmokriski@gmail.com

REFERENCES

Ashford, Elizabeth, and Tim Mulgan. "Contractualism." *Stanford Encyclopedia of Philosophy* (Summer 2018). https://plato.stanford.edu/archives/sum2018/entries/contractualism/.

Audi, Robert. *The Good in the Right: A Theory of Intuition and Intrinsic Value*. Princeton: Princeton University Press, 2004.

Bentham, Jeremy. *Introduction to the Principles of Morals and Legislation*. London: T. Payne and Son, 1789.

Boyd, Richard. "How to Be a Moral Realist." In *Essays on Moral Realism*, edited by Geoffrey Sayre-McCord, 181–228. Ithaca, NY: Cornell University Press, 1988.

Brandt, Richard B. *Ethical Theory*. Upper Saddle River, NJ: Prentice-Hall, 1959.

———. *A Theory of the Good and the Right*. Buffalo, NY: Prometheus Books, 1979.

Brink, David O. *Moral Realism and the Foundation of Ethics*. Cambridge: Cambridge University Press, 1989.

Burge, Tyler. "Individualism and the Mental." *Midwest Studies in Philosophy* 4, no. 1 (September 1979): 73–122.

Caruso, Gregg. "Skepticism about Moral Responsibility." *Stanford Encyclopedia of Philosophy* (Spring 2018). https://plato.stanford.edu/archives/spr2018/entries/skepticism-moral-responsibility/.

Dorr, Cian, and John Hawthorne. "Naturalness." In *Oxford Studies in Metaphysics*, vol. 8, edited by Karen Bennett and Dean W. Zimmerman, 1–70. Oxford: Oxford University Press, 2013.

Dunaway, Billy, and Tristram McPherson. "Reference Magnetism as a Solution to the Moral Twin Earth Problem." *Ergo* 3, no. 25 (2016): 639–79.

Edwards, Douglas. "The Eligibility of Ethical Naturalism." *Pacific Philosophical Quarterly* 94, no. 1 (March 2013): 1–18.

———. "Naturalness, Representation, and the Metaphysics of Truth." *European Journal of Philosophy* 21, no. 3 (September 2013): 384–401.

Gettier, Edmund. "Is Justified True Belief Knowledge?" *Analysis* 23, no. 6 (June 1963): 121–23.

Goodman, Nelson. *Fact, Fiction, and Forecast.* Cambridge, MA: Harvard University Press, 1955.

Guigon, Ghislain. "Overall Similarity, Natural Properties, and Paraphrases." *Philosophical Studies* 167, no. 2 (January 2014): 387–99.

Hare, R. M. *Moral Thinking: Its Levels, Method, and Point.* Oxford: Oxford University Press, 1981.

Harrod, R. F. "Utilitarianism Revised." *Mind* 45, no. 178 (April 1936): 137–56.

Harsanyi, John C. "Rule Utilitarianism and Decision Theory." *Erkenntnis* 11, no. 1 (January 1977): 25–53.

Hawthorne, John. "Craziness and Metasemantics." *Philosophical Review* 116, no. 3 (July 2007): 427–40.

Hirsch, Eli. *Dividing Reality.* Oxford: Oxford University Press, 1993.

———. "Physical-Object Ontology, Verbal Disputes, and Common Sense." *Philosophy and Phenomenological Research* 70, no. 1 (January 2005): 67–97.

Hooker, Brad. *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality.* Oxford: Oxford University Press, 2000.

———. "Ross-Style Pluralism versus Rule-Consequentialism." *Mind* 105 no. 420 (October 1996): 531–52.

Horgan, Terence, and Mark Timmons. "New Wave Moral Realism Meets Moral Twin Earth." *Journal of Philosophical Research* 16 (1991): 447–65.

Jackson, Frank. "Decision-Theoretic Consequentialism and the Nearest and Dearest Objection." *Ethics* 101, no. 3 (April 1991): 461–82.

Joseph, H. W. B. *Some Problems in Ethics.* Oxford: Clarendon Press, 1931.

Kagan, Shelly. *The Limits of Morality*. Oxford: Oxford University Press, 1989.

———. *Normative Ethics*. Abingdon, UK: Routledge, 1998.

Kripke, Saul. *Wittgenstein on Rules and Private Language*. Cambridge, MA: Harvard University Press, 1982.

Lewis, David. "New Work for a Theory of Universals." *Australasian Journal of Philosophy* 61, no. 4 (December 1983): 343–77.

———. "Putnam's Paradox." *Australasian Journal of Philosophy* 62, no. 3 (September 1984): 221–36.

Lyons, David. *Forms and Limits of Utilitarianism*. Oxford: Clarendon Press, 1965.

McNaughton, David. "An Unconnected Heap of Duties?" *Philosophical Quarterly* 46, no. 185 (October 1996): 433–47.

Mill, John Stuart. "Utilitarianism." In *Oxford Philosophical Texts*, edited by Roger Crisp. Oxford: Oxford University Press, 1998.

Nagel, Thomas. "Fragmentation of Value." In *Mortal Questions*, 128–41. Cambridge: Cambridge University Press, 1979.

Parfit, Derek. *Reasons and Persons*. Oxford: Oxford University Press, 1984.

Podgorski, Abelard. "Wouldn't It Be Nice? Moral Rules and Distant Worlds." *Noûs* 52, no. 2 (June 2018): 279–94.

Putnam, Hilary. "Realism and Reason." *Proceedings and Addresses of the American Philosophical Association* 50, no. 6 (August 1977): 483–98.

Railton, Peter. "Alienation, Consequentialism, and the Demands of Morality." *Philosophy and Public Affairs* 13, no. 2 (Spring 1984): 134–71.

———. "Moral Realism." *Philosophical Review* 95, no. 2 (April 1986): 163–207.

Rawls, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971.

Ross, W. D. *The Right and the Good*. Oxford: Clarendon Press, 1930.

Scanlon, T. M. *What We Owe to Each Other*. Cambridge, MA: Belknap Press of Harvard University Press, 1998.

Schaffer, Jonathan. "Two Conceptions of Sparse Properties." *Pacific Philosophical Quarterly* 85, no. 1 (March 2004): 92–102.

Scheffler, Samuel, ed. *Consequentialism and Its Critics*. Oxford: Oxford University Press, 1988.

Sider, Theodore. "Criteria of Personal Identity and the Limits of Conceptual Analysis." *Philosophical Perspectives* 15 (2001): 189–209.

———. "Ontological Realism." In *Metametaphysics: New Essays on the Foundations of Ontology*, edited by David Chalmers, David Manley, and Ryan Wasserman, 384–423. Oxford: Clarendon Press, 2009.

———. *Writing the Book of the World*. Oxford: Clarendon Press, 2011.

Sidgwick, Henry. *Methods of Ethics*, 7th ed. London: Macmillan, 1907.

Singer, Peter. "Is Act-Utilitarianism Self-Defeating?" *Philosophical Review* 81, no. 1 (January 1972): 94–104.

Skelton, Anthony. "William David Ross." *Stanford Encyclopedia of Philosophy* (Summer 2012). https://plato.stanford.edu/archives/sum2012/entries/william-david-ross/.

Slote, Michael, and Philip Pettit. "Satisficing Consequentialism." *Proceedings of the Aristotelian Society* 58 (1984): 139–76.

Smart, J. J. C. "An Outline of a System of Utilitarian Ethics." In *Utilitarianism: For and Against*, by J. J. C. Smart and Bernard Williams, 3–76. Cambridge: Cambridge University Press, 1973.

Van Roojen, Mark. "Knowing Enough to Disagree." In *Oxford Studies in Metaethics*, vol. 1, edited by Russ Shafer-Landau, 161–94. Oxford: Oxford University Press, 2006.

Weatherson, Brian. "The Role of Naturalness in Lewis's Theory of Meaning." *Journal of the History of Analytic Philosophy* 1, no. 10 (2013): 1–19.

———. "What Good Are Counterexamples?" *Philosophical Studies* 115, no. 1 (July 2003): 1–31.

Williams, Bernard. "A Critique of Utilitarianism." In *Utilitarianism: For and Against*, by J. J. C. Smart and Bernard Williams, 77–150. Cambridge: Cambridge University Press, 1973.

Williams, J. Robert G. "Eligibility and Inscrutability." *Philosophical Review* 116, no. 3 (July 2007): 361–99.

COUNTER THE COUNTERSTORY

NARRATIVE APPROACHES TO NARRATIVES

Hilde Lindemann

A COUNTERSTORY, as I have developed the concept, is a story that is told for the purpose of resisting a socially shared narrative used to justify the oppression of a social group.[1] The socially shared story enters the tissue of stories that constitute the group's identity, damaging that identity and thereby constricting group members' access to the goods on offer in their society. The counterstory sets out to uproot some part of the oppressive story and replace it with a more accurate one. In this way it can sometimes repair the damage to the identity. Whether it succeeds is a question of uptake: enough people in the dominant group must accept the new story and treat the members of the group accordingly.

In this paper, I explore some of the difficulties that arise in getting a counterstory to succeed. I will focus in particular on six narrative strategies that people in dominant social positions use to counter a counterstory and thereby keep an oppressive social order in place, and then I will offer a tiny hope. But first I will have to say a little more about how personal identities work.

1. THE SOCIAL PRACTICE OF PERSONAL IDENTITIES

What I have called the social practice of personal identities is a moral practice we engage in constantly, but it has not received much recognition *as* a moral practice: it is the practice of initiating human beings into personhood and then holding them there. I argue that personhood consists of four necessary elements: (1) a human being has sufficient mental activity to constitute a personality; (2) aspects of this personality are expressed bodily; (3) other persons recognize it as the expression of a personality; and (4) they respond to what they see. Recognition and response are often a matter of understanding who someone is and interacting with them on that basis.

These understandings—both the *self*-conception we express and the recog-

---

1   Lindemann, *Damaged Identities, Narrative Repair*.

nition that is *others'* sense of who we are—consist of a web of stories depicting our most important acts, experiences, characteristics, roles, relationships, and commitments. Some of the stories are personal (I grew up in a German-speaking household), while others are widely circulated stories we use to make sense of the social groups to which we belong—white, middle-class women, university professors, mothers, and so on (I am all of the above). This narrative tissue constitutes our personal identities, which play a crucial role in the practice of personhood.

Our personal identities fuel the practice because they indicate who we are with respect to other persons and in that way guide not only how we are supposed to treat those others, but also how we are supposed to conduct ourselves. As the practice of personhood is governed by rules that are both socially and morally normative, my description of how we engage in the practice aims to capture something important about how morality—an interpersonal practice—functions.

We are initiated into personhood though interactions with other persons, and we both develop and maintain personal identities through interactions with others who hold us in our identities, as we hold them in theirs.[2] This holding can be done well or badly. Done well, it supports an individual in the creation and maintenance of a personal identity that allows her to flourish personally and in her interactions with others. Done badly, we hold people in invidious, destructive narratives. In some cases, the damaging narratives identify the *social group* to which someone belongs as socially and morally inferior and in that way uphold abusive power relations between "us" and "them." In other cases, the stories that purport to represent the *individual's* acts or experiences or other important personal characteristics are flat-out false or get the proportions wrong. In this paper, though, I am going to set the flawed depictions of individuals to one side and focus on the stories that damage the identities of the *social groups* to which the individuals belong.

2. OPPRESSIVE MASTER NARRATIVES

For convenience, let us call the socially shared stories that everybody knows the *master narratives* of that society. Many fairy tales qualify as master narratives—think of "Snow White" or "Goldilocks and the Three Bears"—and so do nursery rhymes such as "Humpty Dumpty." Children's books such as the Harry Potter series and the Shakespearean plays *Hamlet* and *Macbeth* count too. So do the biblical tales of Adam and Eve, Samson and Delilah, and the Prodigal Son.

Master narratives are crucial for any social interaction, because they depict

how we are supposed to behave in specific settings. The school narrative tells you that students should sit in their desks and raise their hand for permission to speak but that the teacher may speak and move about freely. The restaurant narrative tells you that when you finish your meal, you must not leave without paying.

Master narratives like these that guide us through life are often reasonably benign, but there are some whose dictates serve to reinforce unjust distributions of social power by pretending to justify them. They do it by depicting certain groups of people—"them"—as inferior to "us." Jews are money-loving self-promoters. Brown-skinned people are dirty and dangerous. Women are stupid sluts. And so on. The trouble is, when a group of people is stereotyped in that way, the people doing the stereotyping find it normal and natural to treat members of the group accordingly. That puts restrictions on how the group members can act. In the United States, for instance, if you are a man with brown skin, the likelihood that you will be killed by the police is three times greater than if you are white.[3] If you are a young black man, the odds go up to twenty-one times greater.[4] If you are a woman, the chances are one in four that you will be raped during your lifetime. In both cases, the stories that masquerade as an explanation for why you deserve that sort of treatment are so powerful that even when we know they are false, they still retain some of their hold over us. One way this works is that master narratives do not essentialize, but tolerate exceptions, and that adds to their durability. One can continue to believe the narrative because the cases where one can see plainly that it is false can be explained away as an aberration: "Well, yeah, she didn't dress like a slut, but most girls do." "Oh, he's different—he's a credit to his race."

Where does their power come from? What is it about damaging master narratives that makes them so much like dragons devouring everything in their path? First, they are organic ensembles of stories that grow and change. Second, they constitute a world view. Third, they are epistemically rigged. And fourth, because of these three features, they are very good at assimilating opposition.

2.1. Organic Ensembles

Let us take these features in order. As is perhaps apparent by now, talk of *a* master narrative is really just a manner of speaking, because the term does not designate a single story with a specific plot and a fixed cast of characters. Instead, they are ensembles of repeated themes that take on a life of their own. Fragments of

---

3   Sikora and Mulvihill, "Trends in Mortality Due to Legal Intervention in the United States, 1979 through 1997."

4   National Public Radio, "ProPublica Analyzes 3 Decades of Deadly Police Shootings."

history, biography, film, fables, jokes, and similar narrative forms ring changes on the theme, as do proverbs, music, advertising slogans, and other cultural artifacts. Master narratives are capacious, as cluttered and untidy as a Victorian attic, but unlike the attic, they grow or shrink over time. An altered political economy or other social change can push a master narrative in a new direction. What moves it along are additional images, stories, songs, and slogans.

Take the story of Cinderella, whose only hope of a good life is for a powerful man to recognize her sexual appeal, marry her, and carry her off to his castle. This got an update in 1990 in the movie *Pretty Woman,* which highlights women's identity as a slut—Julia Roberts plays a sex worker—but her sexual appeal is strong enough to persuade the rich and handsome character played by Richard Gere to marry her anyway. Why? Because she demonstrates that she is just as good at serving others without complaint, in her own way, as her original was at looking after her stepmother and stepsisters, which connects her to all the other stories you know about women whose sluttishness can be forgiven as long as they know their place: subservient to men.

Master narratives take on a fresh vigor with each new accretion, but they are also strengthened by their ties to other master narratives. The master narratives depicting brown-skinned people as dirty purport to justify giving them the dirty work: somebody has to collect the trash and mop the floors, and it is fitting that brown men are the ones to do it. As for brown women, well, since they are sluts and servants, but *also* fit only for dirty work, their job is to serve white people by cleaning their toilets or changing their children's diapers. By interlocking in these ways, master narratives reinforce one another. Because they can incorporate an enormous diversity of even the humblest items in the cultural store, and because they can link themselves to any number of other master narratives that do the same thing, master narratives infiltrate every corner of society.

2.2. Worldview

Precisely because they are so widely known, the master narratives of a society allow its members to understand who they are with respect to that society, as well as how the world works. They create a picture of the world that is both comprehensive and reasonably unified because the narratives that constitute them are not only far-reaching but also mutually reinforcing. A theme—say, that women are to serve men—plays itself out not only in the Cinderella and domestic worker narratives but in the master narrative of modern science, which characterizes nature as the feminine Gaia and reason as the masculine figure in a white lab coat who dominates her. You also see it in the narratives that feminize "lower" races, the ones that prescribe what categories of men may have sexual access to what

categories of women, and the countless advertisements that equate feminine beauty with heterosexual sexiness. One master narrative thus confirms another, and the interlocking, intertwining web of these narratives creates a plausible worldview.

In that narrative web of our beliefs, some of the stories depend on other stories. Those that lie farthest out on the periphery can be given up without much damage to the overall structure, but the closer the story is to the center, the more it supports. Tear out one strand there and a great deal of the web will have to be rebuilt. But since the stories at the center of the web are taken deeply for granted, they are hard to dislodge even when there is evidence that calls them into question, since the evidence can be dismissed as an exception. Even the oppressed, who have every reason to repudiate them, can find it hard to shift them. A master narrative that depicts their group as unworthy or morally subpar can infiltrate the consciousness of some members of the group, causing them to see themselves in the hateful light the story sheds on them. For them, dislodging the story would require calling their own identities into question. Then too, to the extent that group members share the dominant culture's worldview, they have an interest in keeping that narrative web intact. What is at stake is nothing less than their society's—and their own—understanding of life and their place in the universe.

2.3. Epistemic Rigging

The third thing that makes damaging master narratives so hard to contest is that they are rigged to make it seem as if nothing bad is going on. Life is much more comfortable for the dominant class if its members can think of their society as just and good. It is hard to do that if subjugated or excluded classes are continually rebelling, and the use of whips and chains to keep them down is just not very efficient. Much better to deal with resistance by stopping it before it starts. And that calls for stories that hide the coercive power circulating through society so that it is not visible unless you look closely. If you can make the justification for keeping those people in their place seem really plausible, the resulting master narrative will be so smooth and slippery that the opposition cannot grab hold of it.

One way to hide the coercion is to *naturalize* it—to make it seem inevitable that certain classes of people must occupy subordinate positions, or obvious why they have no place in decent society at all. An excellent way to do this is to link the identity to some feature the person was born with, such as a vagina or skin color, because then it appears as if the person's identity came into the world with her, due to some straightforward biological fact about her. The identity does not look as if it was socially constructed at all. No indeed. It was God-given.

If you cannot rig the master narrative by linking the identity to a trait the group member was born with, you might try *privatizing* it. Take disability, for example. Most people with disabilities were not born with them—they acquire them later in life. The easiest thing then is to hide them away in nursing homes, or refuse to make the accommodations that allow them to participate in society. You can build your cities with plenty of curbs and steps that cannot be navigated in wheelchairs, force people who are hearing impaired to hire their own interpreters who know sign language, and require blind people to pass the same written exams as everyone else. As for gays and lesbians, just closet them so you do not have to think about what they are doing in there.

If these tactics fail, you can always *normalize* the identity. This is done when the norms that are supposed to regulate certain kinds of behavior in fact create the identity. Women's identities are normalized by imposing norms of conduct on them that deflect attention away from how they are treated. It is all too common to criticize rape victims for what they were wearing when the rape occurred, or being in a part of town that left them open to attack. Easy, too, to blame the rape on the fact that the victim was drunk at the time, or high on recreational drugs. And if dark skin is supposed to be a badge of subservience, you can be indignant at the person's insubordination without having to stop and wonder *why* dark-skinned people have to be subordinate.

3. COUNTERING THE COUNTERSTORY

If I am right about all this (and I bet I am), it is pretty obvious that counterstories have their work cut out for them. The master narratives they go after are so interconnected, so deeply rooted in our psyches, and so smooth and slippery that the counterstory mostly just bounces off and rolls away harmlessly. And if it cannot get inside the narrative to repair the tissue that is damaging the identity, the people bearing the identity remain unable to move about freely within their society. But let us suppose that the counterstory manages to engage with the master narrative rather than being deflected. At that point the counterstory still faces a huge obstacle, namely, master narratives' ability to *assimilate opposition*. In what is left of this paper, I am going to describe six ways these narratives absorb opposition, thereby retaining their power and continuing to inflict damage.

3.1. Make the Language Pretty

Counterstories protesting violence against members of a subgroup can be tweaked so that they represent what happened in language that is more comfortable for the oppressor. Domestic assaults, committed far more often and more

violently by men on women than the other way around, become written up in the *New York Times* as "domestic disputes," which simultaneously erases the violence and makes it appear that both parties contributed equally to a disagreement. "Racial unrest" is a euphemism that hides both the violent behavior of the subgroup protesting its oppression and the violence used by police or other servants of the oppressive social order to beat the subgroup back into submission. "Some of my best friends are Jews" or "I never think of you as black; to me, you're just human" are pretty sentiments that deny the power differential between the speaker and the subgroup, thereby making the subgroup's counterstories seem to be beside the point, or inaptly targeted. Recently I heard someone dismiss sexual harassment as "romantic entanglements," which erases not only the power hierarchy between predator and prey, but also diverts attention from the sense of entitlement that accompanies unwanted sexual groping. Attractive commercials that depict elderly people as retaining their (white) beauty and vigor well into the "golden years," where they play golf and tennis at exclusive country clubs, or walk hand in hand on sunset beaches, do not just paper over the very real indignities of old age's diminishing physical and social power, but make it impossible for counterstories that destigmatize vulnerability and dependence to get heard at all.

3.2. Play Devil's Advocate

Here is a second way to counter a counterstory. Let us take the master narratives about lazy, shiftless poor people that arguably fuel the Trump administration's cutbacks in federal spending. Suppose a group of friends is discussing these cutbacks and, by way of a counterstory, someone points out that many people living below the poverty level actually work two or three jobs and still cannot make ends meet. Now suppose, "just to play devil's advocate," someone else says it is a well-known fact that poor people have babies on purpose, so they can receive welfare benefits. That assimilates the resistance by sending the conversation off into a discussion of whether people should procreate if they cannot support their children, rather than engaging with the counterstory about the working poor. The devil's advocate technique can also be used to counter counterstories that display the sexism behind the wage gap. When this comes up there is usually someone—not always a man—who, just for the sake of argument, says that women are paid less because they choose to take time off from work to care for their children or elderly relatives. It's a theory. For the sake of argument. Here, as in the case of cutbacks for the poor, the master narratives purporting to justify the oppression absorb the counterstory by wrapping it in a related master narrative for which the teller takes no responsibility. Often, people who play devil's

advocate actually believe the arguments they say they are making just for fun, but because they suspect that the arguments would make them sound arrogant or privileged, they put them in the devil's mouth.

These discussions may feel like fun to the devil's advocate, but that is because the issues under discussion do not directly affect them. If they are not one of the working poor their motives for having children are not questioned, and if they are a man they are not expected, in any case, to cut back their hours at work to take care of the children. But there may well be people present in these conversations who do not think devil's advocacy is fun at all. As a blogger on Feministing puts it:

> It is physically and emotionally draining to be called upon to prove that these systems of power exist. For many of us, just struggling against them is enough—now you want us to break them down for you? Imagine having weights tied to your feet and a gag around your mouth, and then being asked to explain why you think you are at an unfair disadvantage. Imagine watching a video where a young man promises to kill women who chose not to sleep with him and then being forced to engage with the idea that maybe you are just a hysterical feminist seeing misogyny where there is none. It is incredibly painful to feel that in order for you to care about my safety, I have to win this verbal contest you have constructed "for fun."[5]

3.3. *Play What about Me?*

A third way to counter a counterstory is to deflect the conversation back to the dominant group. Since the 1990s a loose coalition calling themselves Men's Rights Activists have been countering feminist counterstories by insisting that it is men who are oppressed and women the oppressors. Men's rights groups have identified women as the source of emasculation and subjugation, claiming that they file false paternity suits and maliciously accuse innocent men of raping them. Also, some women make more money than some men do. (The eyebrow shoots skyward.) Paul Elam, founder of A Voice for Men, has declared publicly that if he ever sits on the jury of a rape trial, he will vote to acquit even in the face of overwhelming evidence that the defendant is guilty. Why? Because the legal system "is patently untrustworthy when it comes to the offense of rape. . . . In this, the age of misandry, not one aspect of a rape case can be trusted. . . . The accuser cannot be trusted."[6] Now, Elam's way of countering the counterstory by turning the focus back on the poor, mistreated men is not particularly good because it is

5 Britto Schwartz, "An Open Letter to Privileged People Who Play Devil's Advocate."
6 Ford, "A Lesson for Men's Rights Activists on Oppression," 14.

too crude to be plausible, but it does display how the mechanism works: let us not talk about you when we could be talking about the much more important and put-upon Me.

That example is pretty blatant, so let us look at one that is more subtle. I refer to the mothers of young men accused of rape, who defend their sons vigorously with the claim that they could not possibly have done it—it was not in their character. This looks altruistic, but often it is really a variant on "What about Me?" because what they are actually saying is, "Look at all I've done for him. I'm not the kind of mother who would have a rapist for a son."

Another example of the "What about Me?" technique has been described by the philosopher Uma Narayan. Since at least the publication in 1978 of Edward Said's *Orientalism,* many counterstories have been launched against colonialist master narratives of the barbarism and backwardness of the brown-skinned people living in the Near East and on the Indian subcontinent. But Narayan reports that in their eagerness to distance themselves from Western imperialism, Western feminists are unwilling to say anything negative about Eastern cultures. And they are equally unwilling to hear Narayan's criticisms of the Indian social order. As a result, in their conversations with her, they keep pulling the conversation back to how Western colonialism is responsible for all of India's ills.

> What starts as a proper stress upon negative Western attitudes and interferences in the Third World grows into a focus on the Big Bad West, one that operates so as to virtually eclipse the Third World and its agents, institutions, and responses from view. The Third World virtually vanishes, except as a flat backdrop or frame for the Bad Deeds of the West. I have often felt like Alice, watching "the Third World" slowly disappearing from view until it has all the substantiality of the Cheshire Cat's grin.[7]

This sort of breast-beating is not confined to groups with a colonialist history, of course. It also manifests itself as white liberal guilt, which is far more interested in how bad I feel about racism than in what it is like for you to be oppressed by it.

3.4. Require Victims to Be Blameless

The fourth way to counter a counterstory is a specific kind of victim blaming, which says that a victim of oppression is a victim, all right, but only if she is of sterling moral character. Did he get angry when your friend told that anti-Semitic joke? Well, everybody knows these Jews do not have a sense of humor. Was her skirt too short or her history of sexual activity too long? Then it was not really rape or, at least, she was asking for it. Had he ever been accused of petty

---

7   Narayan, *Dislocating Cultures*, 140.

theft? Then even though he was not armed, you cannot really blame the police for gunning him down and leaving his body on the road for four hours.

In the United States it is not uncommon to hear Zionists defend Israel's oppression of Palestinians by pointing out that Hamas is a terrorist organization, as if that justifies the oppression. While Hamas's activities certainly make it harder to end the violence in that part of the Middle East, the assumption that oppression is somehow legitimate because not everybody in the oppressed group is of upright moral character can just as certainly be questioned.

3.5. Change the Subject

When all else fails, a master narrative can absorb a counterstory by changing the subject. When women complain about sexual harassment in the workplace, for instance, the conversation frequently veers off in the direction of the woman's work performance or her attitude toward her bosses. The same thing can happen when someone tries to start a discussion protesting some racist incident or pattern of incidents. Often the conversation morphs into a discussion about all the ways in which society is now "post-racial," or why it is racist to bring up the topic of race in the first place. What gets lost sight of in these conversations is the oppressive practice or incident that was the initial reason for the discussion.

4. HOW TO SLAY A DRAGON

By now you are probably wondering why, if master narratives are so good at repelling or assimilating them, anyone would bother to launch a counterstory at all. If I am right about any of this—and mind you, what I have provided is not even close to an exhaustive list of the ways to counter a counterstory—then it seems as if Bilbo the hobbit's hope of defeating the gargantuan Smaug who absorbs all the dwarfs' treasure is doomed from the outset. So I am going to end this paper by explaining why that simply is not so.

To see that it is not, we have only to look at a counterstory that has been quite effective in getting uptake by the oppressors, to the point where many members of the oppressed group are now able to exercise their agency more freely in the wider society. In 2006, Tarana Burke founded the #MeToo movement to help survivors of sexual violence, particularly young women of color from poor communities, speak out in public about what had happened to them and find resources for healing. The movement was aimed at destigmatizing survivors by highlighting the breadth and impact sexual violence has had on women and disrupting the systems that allow this violence to flourish. #MeToo had already been in existence for eleven years when, on October 15, 2017, Alyssa Milano reig-

nited it on Twitter: "If you've been sexually harassed or assaulted, write 'me too' as a reply to this tweet." Burke says it is more than a hashtag—it is the start of a public conversation and a space for community healing. Many women have now come forward under the hashtag to report their experiences with sexual harassment, molestation, and rape.

The #MeToo movement seems to be creating a cultural shift that has made it easier to prosecute men like Bill Cosby and Harvey Weinstein, and perhaps also propelled the downfall of Judge Roy Moore. But to change the culture, the point has to be not that women are revealing abuse in the workplace, but that it occurs. Koa Beck, former editor-in-chief of the website Jezebel, urged that the takeaway should not be that men and women cannot work together, but rather that men and women step up as bystanders when they see harassment at work. The leaders of the movements are now trying to shift public focus away from the high-profile cases and onto women like Suzette Wright, who suffered in silence for years at Ford Motor Company. The shifted focus might begin to erode the culture in which gendered messages about sex, power, and courtship start young and are deeply ingrained.

The current social climate makes it easy for sexual predators not to repent. President Donald Trump and others have repeatedly expressed sympathy for Supreme Court Justice Brett Kavanaugh, who, before he was confirmed, had been accused by Christine Blasey Ford of violent attempted rape when they were teenagers. "I feel so badly for him. This is not a man who deserves this," Trump said. Kate Manne calls this kind of concern himpathy—"the inappropriate and disproportionate sympathy powerful men often enjoy in cases of sexual assault, intimate partner violence, homicide and other misogynistic behavior."[8] Himpathy is on display everywhere: in the actions of a former editor of the *New York Review of Books* who published a self-indulgent essay by a disgraced Canadian talk-show host accused of sexual harassment by many women; in Senator Lindsay Graham's outraged indignation at how Kavanaugh was treated; in the "boys will be boys" attitude visible all over Facebook and Twitter. All this focuses exclusively on the perpetrator's pain, the perpetrator's future, the tarnishing of the perpetrator's good name. And the higher the perpetrator stands in the social hierarchy, the more himpathy he attracts, so the bulk of social concern, care, respect, and attention goes to the most privileged men in our society. Himpathy goes a long way to explain why women who have been harassed or assaulted are not believed. The warm feelings all go to the male perpetrator: he was just a boy, the rules were different then, she is trying to get him into trouble. Because

8   Manne, "Brett Kavanaugh and America's 'Himpathy' Reckoning."

himpathy is such a powerful moral emotion, it has to be consciously recognized and strenuously resisted.

If that is so, why do I see hope in the #MeToo counterstory? Kavanaugh was appointed to the Supreme Court in spite of it, but the movement has accomplished one thing: in the heated discussions surrounding his confirmation, almost no one attacked Blasey's character. No one said she was a scheming slut, or deserved what she got because she had been drinking. Instead, those who disbelieved her speculated that she had misidentified her assailant, or made too much out of an incident that happened long ago. That may not seem like much, but it is evidence that, despite the master narrative's attempts to absorb it, the counterstory is getting uptake.

The reason why any counterstory has a chance of succeeding, despite whatever defenses a master narrative might set up against it, is that oppressive master narratives are always *false.* That means there are always cracks in the story, soft spots in Smaug's underbelly that let the arrow find its mark. It is stories like Mann's that show us the master narrative's mechanism, that ultimately bring the narrative down. If you shoot enough counterstories at any master narrative, you will find that it is not so invincible after all.

One story alone will not do it, of course. Nor will even a raft of stories. But if you loose enough counterstories at that soft spot on Smaug's underbelly where one of his scales fell off, you will discover that he is not so invincible after all.

Michigan State University
hlinde@msu.edu

REFERENCES

Britto Schwartz, Juliana. "An Open Letter to Privileged People Who Play Devil's Advocate." *Feministing*, May 30, 2014. http://feministing.com/2014/05/30/an-open-letter-to-privileged-people-who-play-devils-advocate.

Ford, Clementine. "A Lesson for Men's Rights Activists on Oppression." *The Drum*, June 19, 2014. http://www.abc.net.au/news/2014-06-19/ford-a-lesson-for-mens-rights-activists-on-real-oppression/5533412.

Lindemann, Hilde. *Damaged Identities, Narrative Repair.* Ithaca, NY: Cornell University Press, 2001.

———. *Holding and Letting Go: The Social Practice of Personal Identities.* New York: Oxford University Press, 2014.

Manne, Kate. "Brett Kavanaugh and America's 'Himpathy' Reckoning." *New York Times,* September 26, 2018.

Narayan, Uma. *Dislocating Cultures: Identities, Traditions, and Third World Feminism.* New York: Routledge, 1998.

National Public Radio. "ProPublica Analyzes 3 Decades of Deadly Police Shootings," October 14, 2014. https://www.npr.org/2014/10/14/356045089/propublica-analyzes-data-on-male-victims-of-deadly-police-shootings.

Said, Edward. *Orientalism*. London: Routledge and Kegan Paul, 1978.

Sikora, Andrew G., and Michael Mulvihill. "Trends in Mortality Due to Legal Intervention in the United States, 1979 through 1997." *American Journal of Public Health* 92, no. 5 (May 2002): 841–43.

COUNTERSTORIES, STOCK CHARACTERS, AND VARIETIES OF NARRATIVE RESISTANCE

RESPONSE TO LINDEMANN

Mark Lance

I HAVE BEEN THINKING quite a bit over the last two years about Hilde Lindemann's work on our narratively structured holding of one another in personhood. I am broadly sympathetic to her approach: to the idea that a necessary dimension of personhood is being given communicative uptake, to the implication that this both allows us to constitute dimensions of positive freedom and at times constrains us in ways that can constitute oppression, and finally to the suggestion that the available patterns of uptake and understanding are themselves constituted, partly, by socially available narratives.

> We are initiated into personhood through interactions with other persons, and we simultaneously develop and maintain personal identities through interactions with others who hold us in our identities. This holding can be done well or badly. Done well, it supports an individual in the creation and maintenance of a personal identity that allows her to flourish personally and in her interactions with others. Done badly, we hold people in invidious, destructive narratives.[1]

In "Counter the Counterstory," Lindemann engages directly with the cases in which the narrative constitution of lives is directly oppressive, where the kind of character we are socialized as, on the basis of widely available narratives involving instances of our many identities, cuts off possibilities, damages our potential for flourishing, or leaves people in a socially dominated position. Her goal is to begin articulating ways that "counter-narratives" can challenge oppressive narrative formations, and the reactionary tactics they mobilize to defend against such challenges. I find myself in broad agreement with the claims and, maybe more important, welcome reflection on this social dynamic of narrative and counter-narrative. Indeed, I find this all to be one important vector in the philosoph-

---

1 Lindemann, *Holding and Letting Go*, x.

ical study of activism—a dimension of human experience that was largely ig-
nored in twentieth-century Anglo-American philosophy, and which is only just
becoming an explicit topic. For this reason, especially, I welcome the current
contribution.

In section 1 of this paper, I discuss one mechanism by which widely available
narratives provide models for social lives via the institution of "stock characters"
that function as themes upon which living improvises. My aim is not to offer
anything like a theory of the constitution of the self here, nor even at the level
of metaphor to suggest that this is the only mechanism by which narrative influ-
ences the selves we hold one another to. Rather, my goal is to highlight some vul-
nerable joints and pressure points at which social pushback against oppressive
narratives can gain purchase. In section 2, I emphasize some additional ways that
these narratives can be oppressive. Finally, in section 3, I extend Lindemann's
discussion by illustrating some of those forms of pushback.

1. WHAT'S NARRATIVE GOT TO DO WITH US?

For convenience, let us call the socially shared stories that everybody
knows the *master narratives* of that society. Many fairy tales qualify as
master narratives—think of "Snow White" or "Goldilocks and the Three
Bears"—and so do nursery rhymes such as "Humpty Dumpty." Chil-
dren's books such as the Harry Potter series and the Shakespearean plays
Hamlet and *Macbeth* count too. So do the biblical tales of Adam and Eve,
Samson and Delilah, and the Prodigal Son.

Master narratives are crucial for any social interaction, because they
depict how we are supposed to behave in specific settings.[2]

On an overly literal reading of such passages, one could think that the story goes
like this: certain stories have near-universal uptake. People in various groups are
supposed to identify with particular characters in the story and society demands
that they live out their lives as the character does. But the obvious objection here
is that none of us live with talking bears or Shakespearean witches. Society does
not demand that women cut the hair of superhero lovers. So we certainly do not
literally follow the story.

Of course one can view the stories more abstractly, as a story about being a
submissive or vengeful woman. But the worry is that if we understand narratives
such as these at a level of abstraction sufficient to allow them to apply to actual
life, then the narrative features are not doing much work. Why not simply say

---

2   Lindemann, "Counter the Counterstory," 287–88.

that society has rules—women should be submissive—rather than that stories define us?

Lindemann, in fact, makes it clear that this is not the right way to think about narrative constitution when she emphasizes that it is "the tissue of stories that constitute the group's identity."[3] It is not that we are taught by society to read our behavior off of a single story, but that we have a cluster of stories for each dimension of identity, clusters that determine both how we are meant to understand ourselves and how society gives us uptake. Many stories have a character of a mother, for example. While these mothers do very different things in very different situations, commonalities of behavior, character, and motivation emerge from our exposure to them. The idea, I suggest, is not that we memorize a list of imperatives—if in this situation a person of type t must do x—but that we come to grasp a "stock character."

While offering nothing close to a theory, I think this literary trope is helpful in understanding the way that social narratives shape self- and other-understanding. A stock character, recall, is a stereotypical or archetypal character that can be seen as re-appearing in multiple stories. From the ancient Greek characters of Theophrastus and the epic heroes of national founding myths, to the hard-boiled detectives and femme fatales of film noir, literature is replete with stock characters. However, exactly, it is that we come to such a conceptualization—my goal is political strategy not cognitive science—we do recognize, say, the cynical hard-boiled detective when we see him. And when we do recognize him as such, we extrapolate typical behavior from past literary appearances of that character, even across wildly different contexts. (Rick Deckard in *Blade Runner* is immediately familiar to us, as an instance of a type that includes Philip Marlowe.)

Of course stock characters do things that surprise us. (Literature would be really boring otherwise.) So what exactly is it for something to function as a stock character? I suggest first that for there to be a stock character in a society just is for that society to recognize instances as of that stock character. So what is it to recognize someone (real or fictional) as of a stock type? Without offering anything like a complete analysis or theory, I suggest that such an understanding involves the following elements.

First, the recognition of Philip Marlowe as of the type [noir detective] provides us a sort of explanatory template. To understand him as of this type means that certain actions will appear "in character"—a cynical sneer, an unhealthy whisky and cigarette habit, a certain cavalier relationship to legal niceties and sexual consent. Someone understanding Marlowe as of the type will automatically see all this as natural. He is that character; that is what such characters do.

---

3   Lindemann, "Counter the Counterstory," 286.

None of this tells us specifically what he will say, or do, in a particular situation, but it gives us a general orientation toward him as a character, guides what surprises us and what does not.

Second, giving uptake to a character as a particular stock character does not mean that they cannot ever act out of character. One of the many reasons *Casablanca* is a great work is precisely the way that Rick breaks out of his mold as the familiar world-weary cynic, first in rejecting the great romantic love of his life, and second in embracing his beautiful friendship and returning to the cause. But he does these things *as a world-weary cynic would*. To be of a type does not mean one cannot break that type in particular actions, even actions that appear, as these do, destined to move one out of type. But first, these actions are unexpected. We are startled that Rick gives up the girl to his romantic rival. We are inspired and thrilled that he walks into the sunset with Louis. And second, unlike actions in type, they *call for explanation*. We need to understand that it is because the problems of three little people do not amount to a hill of beans in this crazy world that he is rejoining the fight. These things play their narrative role precisely as acts of originality, of defiance of who he is, in exactly the way that his drinking away his sorrows with Sam or cynically bantering do not.

Finally, none of this is explicit. To understand someone as of a certain type is not to consult a theory, or even to be consciously aware of what aspects are typical and which atypical. It is, rather, what Heidegger calls an engaged mode of "Being-with," a smooth skillful habit of engagement with someone.

And what holds for fictional characters holds as well for real ones. We understand people as instances of various stock characters and that understanding guides our interactions with them in the same way that our understanding of fictional characters guides our interpretations. That, at least, is my suggestion for fleshing out the relation between widely available social narratives and our interpretations of one another: these narratives construct for us a wide range of stock characters, and we then read one another as instances.[4] I am an academic, a radical activist, a nerd, a father, and a musician—each with its own small range of stock characters that give me a sense of what that means. I perform daily actions that are in role, and when I run into conflicts or novel situations, my understanding of these characters provides a framework within which to think things through. Just as important, and crucial to Lindemann's account of holding us in personhood, others understand what I am doing in these terms as well: Lance

---

4   Of course it is not only narratives that do this. The actual lives of people we grow up with—either family and friends or those regularly in the news—can play the same role in constructing an understanding of a type.

is off at a protest again; sure, that is what activists do. He missed the organizing meeting; yes, because he is a father and his daughter is sick.

2. CONSTRAINING CHARACTERS

There are a number of distinctions worth noting in the way we inhabit stock roles and the ways that they can harm us. Some of our roles we choose—philosopher, chess player—while others are largely outside our control, socially assigned on the basis of various observable features, with no input from the person assigned to that role—man, member of family *x*, perhaps in some cases religious affiliation.[5]

Socially imposed roles are particularly problematic. Even if not inherently oppressive, a role that one has no hand in authoring might simply not fit one's skills, self-understanding, or goals and even if it does, the idea that there is a dimension of identity that is outside one's autonomy might itself be problematic. Thus, even in masculinist societies, the imposition of the category *man*, and the associated stock character, could oppress someone who cannot live a flourishing life under that designation, for whatever reason. But having a role in the choice hardly guarantees that roles are unproblematic. Most obviously, one might have a forced choice between equally unpleasant alternatives. One might argue that being an employee in a capitalist system is inherently alienating. If so, then the fact that one can choose which capitalist to sell one's labor to does not remove the worry.

Some characters are subject to specific rules as a result of the stories that constitute them. Perhaps the Hippocratic oath is an essential element of our social stories of doctors, or a rule of fidelity of our stories of spouses.[6] Other times, there may be no explicit statement of the rule, but a norm emerges as a consistent aspect of a whole range of stories: "students should sit in their desks and raise their hand for permission to speak" (Lindemann's example), "women should be submissive," "fighters should be brave and protect the innocent."

5   There are also intermediate cases in which input into the social kind is neither purely voluntary nor socially given. Friendships, musical genres, and statuses like geek or hip, etc., have complicated entrance (and exit) conditions that involve a dialectical interaction between individual and social group. Cf. Lance and Kukla, "Intersubjectivity and Receptive Experience."

6   The idea is not that one must follow these rules to be an instance of the character. Any rule can be violated. The point is that to be of type t is to be *subject to* rule r. If one is not a doctor, there is no Hippocratic oath to violate, and if one is a doctor, then not following it is a violation.

Whether explicit or implicit in stories, these rules sometimes emerge as implications of seeing someone as a student/woman/doctor/fighter because they are central to all the relevant narratives.

Sometimes stories do not structure their characters as explicitly around rules as they do around characteristics: female lovers are blond and thin, male lovers are muscular and thin. Not that any social story is likely to suggest that such a character is *always* this way—if so, then we might as well treat that as a rule—but they can define a sort of paradigm or default. An instantiation of the character with the paradigmatic quality is "normal," unexceptional, not needing comment. Again, those who differ from the paradigm call for special explanation. This aspect is often remarked on in the context of race—the default framing of many characters is that they are white.

Such default assumptions can be their own sort of societal burden, even when it is possible to go against the default assumption. To always have to explain oneself, to be constantly subject to interrogation or even curiosity, to always be the atypical example of the type—all of these can be a form of social exclusion in the right context.

Similar default assumptions about characters apply to paradigmatic courses of action. Perhaps the standard story of the son is that he follows in his father's career footsteps, of the young woman that she will welcome flirtation and be on the lookout for a husband. Again, the imposition of such an assumption implies that anyone violating it becomes something of a rebel, or at least socially defiant. To occupy the social position without that course of action is to challenge existing normative assumptions. On top of that, deviant versions of a character may themselves become stock characters—the model minority, the perfect protestor, etc.—which can bring their own oppressive possibilities. By giving us standing default assumptions about the behavior of those defying a given role, they create a new role, often with a burden of double-defiance.

It is also possible to be harmed by the impossibility of inhabiting a particular character. This is most obviously the case when a role did exist and centrally structured the lives of real people only to have the material or social conditions necessary for its continuation forcibly removed. Jonathan Lear describes a particularly poignant and systematic removal of narrative possibility in *Radical Hope*, where he explains the utter destruction of the way of life—and possibilities for narrative understanding—of the Crow after the massacre of the buffalo and the people's imprisonment on reservations.[7] But even if what is at issue is not the destruction of existing narratives, I think we can call it a form of harm when a society simply does not allow certain stock characters to exist, if those

7   Lear, *Radical Hope*.

characters would facilitate the flourishing of people in the society. The emergence of various queer/genderqueer identities has been liberatory for people not comfortable identifying with the categories previously in existence in society, and arguably social resistance to the emergence of these stock identities was a harm. It seems to me that there is a positive social duty—defeasible of course and subject to all manner of trade-offs and resource issues—to promote the development of the practices inherent in potentially flourishing new identities.

Narratives can also frame stock characters within certain contexts or in conceptual terms that are harmful, even if the character itself is not, and sometimes that harm affects more than just the person who might inhabit the character. Think of the way that contemporary accounts of work/labor frame it as essentially commodified, thereby obscuring the very possibility of free productive labor. Or think of the many stories of "the racist." Repeated characterization of racists as explicitly vicious Klan members, even if they are presented as villains, reinforces the assumption that this is the only way to be racist. Thus, the more commonplace ways that we are complicit in, facilitate, or simply fail to challenge white supremacy become invisible through the very stories that institute the stock character of the racist. Or again, consider the very grammatical framing of a character such as "disabled person." We introduce the character via a one-place operator on persons. That is, "disabled" appears in typical narratives as a property of an individual, rather than, say, as a relation between an individual and an environment.

There is certainly much more to say in this regard. The forms and textures of oppressive social narrative are many, and underexplored. But I will leave it at that for now, and turn to an even more brief survey of varieties of resistance.

3. RESISTING, REFORMING, AND DESTROYING CHARACTERS

Lindemann focuses primarily on the process of creating counterstories and the ways that dominant narratives can resist that process. In this final section, I want to begin addressing the rich variety of ways that people go about countering oppressive or harmful stock characters.

Sometimes we try to reform the nature of a given stock character. Lindemann considers the #MeToo movement, which fits this category. The idea is to counter aspects of the stock character "woman," in particular the way such a character is expected to respond to unwanted sexual aggression. A number of aspects and behaviors come to be expected or interpretively privileged in our understanding of male-female sexual aggression: the well-meaning but awkward guy who goes a bit too far, or the vindictive woman who regrets sex, is out to get men, etc. As

Lindemann emphasizes, the movement functions as much to challenge these narrative conventions as it does to remedy particular wrongs by particular people. And in the fabric of stories that become public in this explicitly storytelling movement there emerge other versions of these characters—women asserting autonomy, men succeeding (or failing) in making a proper apology or appropriate restitution, new restorative and, indeed, punitive potentials, etc.

Related to this example, consider the early marches of ACT UP—thousands of angry, powerful gay men chanting "We're here; we're queer; get used to it!" This sort of public performance—and as public performative protest, I think it counts as a sort of narrative—took place in a context in which most Americans gave uptake to gay men as a stock character that is effeminate, passive, and closeted. This massive and directly challenging confrontation with that character—repeated dozens of times in dozens of cities, together with the thousandfold repetition of the process of coming out—was an attempt to destroy our ability to habitually render people according to the oppressive narrative.

Sticking with the same movement, we saw two powerful responses to Lindemann's counter-counterstory strategy of "making the language pretty." Early on in the AIDS epidemic, mainstream society took up one of two stances: the outright hostility of many religious and conservative political leaders who claimed that the disease was God's punishment or a predictable result of moral weakness; or on the other end of respectable opinion, those who counseled care and compassion for the dying, the victims of a "terrible tragedy." This latter response, emotive compassion, arguably made the response to systemic heterosexism pretty and acceptable in a way that diverted the radical potential of the movement. But ACT UP was having none of this, understanding that the death toll was as much a function of institutional decision as it was of natural process, and calling out the hypocrisy of caring for the dying, while encouraging the closet and systemic discrimination for the healthy. One concrete performative manifestation of this response to prettying the language was the much-used slogan of the movement: "Everyone loves a dying fag!"

Even more directly rejecting of liberal attempts to pacify and tame the movement: ACT UP members who had lost loved ones took to hurling their ashes into the faces of hostile politicians and religious leaders. It is very hard to continue conceptualizing someone as an instance of the passive, effeminate stock character after such an action.

Of course there are dangers to such tactics, and ways that the dominant narrative can react—the creation of new altered stock caricatures of the angry, intolerant gay man—but such is the dialectical nature of social change.[8]

8   The social enforcement of even the most absurd myths about marginalized groups is often

A stock characterization of pacifists—going back to the introduction of the word "Quaker" as a slur for the Society of Friends, indicating that they quaked in fear—has it that they are motivated by cowardice or lack of discipline. The civil rights movement worked hard to undercut this story of nonviolence, both in the writings of Dr. Martin Luther King Jr. and others that emphasized the courage necessary to confront violence, oppression, and racism without fighting, and in the practice of trained cadres that endured torture, imprisonment, and murder without breaking movement discipline. The portrayals of these brave men and women in the press constituted a counter-narrative that unsettled the dominant conception.

Evocative literature, then, often produces a different account of a negatively portrayed character. But it can also undercut the association of a negative trait with a particular character by associating it instead with a narrative opposite. Tim O'Brien's brilliant essay, "On the Rainy River," is a sustained reflection on his decision to go to Vietnam, rather than to resist the war by going to Canada. O'Brien recounts how he was afraid for his life and also morally opposed to the war—this was no opposition of principle and self-interest, for both pushed him toward Canada—but ultimately more driven by a fear of social disapproval, the mocking of friends and family. Embarrassment, he says, overcame the combined impetus of self-interest and morality. His essay ends with these words: "I passed through towns with familiar names, through pine forests, and down to the prairie, and then to Vietnam, where I was a soldier, and then home again. I survived, but it is not a happy ending. I was a coward. I went to the war."[9] By associating—and evocatively rendering in rich detail—the way that participation in warfare could be cowardly, O'Brien undercuts the association of that trait with the stock character of the war resister.

Humor and absurdity can also destabilize our understanding of stock characters. Think here of the actions of Yippies, performative, queer, street-theater groups like the Lesbian Avengers, who performed in what became known in the 1970s and '80s as "Temporary Autonomous Zones." In each case, the goal is not

---

not subtle. I came to Georgetown University in 1991. A professor in the philosophy department at the time wrote an article in the campus newspaper that year claiming that ACT UP was the moral equivalent of the Nazis, on account of its confrontational actions directed toward the New York Catholic Diocese, which had been a leader in promulgating antigay sentiments. I wrote a response piece that pointed out a few salient differences, such as that the Nazis launched the Holocaust and a world war, whereas ACT UP had not physically harmed anyone. His response to this was to refuse to speak to me for the remainder of his tenure at Georgetown and to tell students that I endorsed killing Catholics. Sadly, such hysterical reactions to the counternarrative actions of ACT UP were not unusual.

9   O'Brien, *The Things They Carried*, 58.

so much the creation of a concrete alternative to existing stock characters, but the unsettling of their place in the social psyche. When Yippies tossed dollar bills onto the trading floor of the New York Stock Exchange, the idea was less to construct a new image of capitalism and its various characters and more to simply break the spell of a particular narrative—to make traders look absurd.

Sometimes the creation of counter-narratives involves the creation of quasi-separatist communities. Such communities might produce novel understandings of a given character, or new characters entirely. In some cases, there is little effort to then integrate the new characters into the broader social understanding—e.g., Amish and other religious communities that seek only marginal integration—and other times local experiments burst out and demand broader recognition—for example the Stonewall riots, or various communes, anarchist communities, and collectivist squats that see themselves as prefiguring an anticapitalist way of life.[10]

As Lindemann says, there are no recipes here. For every counterstory tactic, there is a counter-counterstory retrenchment, dismissal, mischaracterization, etc. For every list of tactics on either side, it is possible to generate new ones. But the creative possibilities for dismantling oppressive narratives and the construction of new possible identities is enormously rich and an ongoing field of social contestation. To resist, reform, destabilize, and reconstruct the stock characters that provide the narrative grounding of our holdings of one another is part of what it is to be human. As such, it is a welcome subject for philosophical reflection.

Georgetown University
lancem@georgetown.edu

REFERENCES

Lance, Mark, and Rebecca Kukla. "Intersubjectivity and Receptive Experience." *Southern Journal of Philosophy* 52, no. 1 (March 2014): 22–42.
Lear, Jonathan. *Radical Hope: Ethics in the Face of Cultural Devastation.* Cambridge, MA: Harvard University Press, 2006.
Lindemann, Hilde. "Counter the Counterstory: Narrative Approaches to Narratives." *Journal of Ethics and Social Philosophy* 17, no. 3 (June 2020): 286–98.

10 For a theory and detailed case study of such processes of identity construction, see Shotwell, *Knowing Otherwise*, esp. ch. 5.

———. *Holding and Letting Go: The Social Practice of Personal Identities.* Oxford: Oxford University Press, 2014.

O'Brien, Tim. "On the Rainy River." In *The Things They Carried*, 37–58. Boston: Houghton Mifflin Harcourt, 2010.

Shotwell, Alexis, *Knowing Otherwise*, University Park, PA: Penn State University Press, 2011.

IDEOLOGICAL ABSORPTION AND COUNTERTECHNIQUES

COMMENTS ON LINDEMANN

Ásta

W HAT ARE MASTER NARRATIVES?[1] To use vocabulary that may be more familiar to some readers, they are part of ideology.[2] They are stories that are told in order to explain and justify arrangements. Accompanying them are policing mechanisms, or ideological apparatuses, to keep people in their assigned roles and to keep them from disrupting the arrangements and the justificatory story. Hilde Lindemann's focus in her essay is on what she calls "counterstories" and how the master narratives resist them by absorbing them. Counterstories are ideology critiques aimed at dismantling the justifications that are offered for a certain arrangement or phenomenon by offering alternative and more plausible explanations. The ideological absorption under discussion happens when counterstories are resisted. Lindemann identifies several ways that ideological absorption works, techniques that are used to neutralize the ideology critique. We should not be surprised. Ideology is very resistant. Despite that resistance, Lindemann is hopeful, for she thinks that the success of the #MeToo movement shows that ideology critiques sometimes manage to get traction and effect change. I am interested in the conditions of success.

Lindemann herself rightly points out that the *falsity* of master narratives is the crack in the wall that will, with enough onslaught, make it crumble. But that is only the first step in an account of what is required for a successful ideology critique. Drawing on Lindemann's insightful discussion of the various ideological-absorption techniques, I will venture to offer some further steps.

I begin by commenting on features and functions of four of the techniques

---

1   I am delighted to get a chance to comment on Hilde Lindemann's essay, "Counter the Counterstory." I have learned so much from Lindemann's work and been deeply moved by it.

2   Althusser, "Ideology and Ideological State Apparatuses"; Geuss, *The Idea of a Critical Theory*, 81.

Lindemann discusses: make the language pretty; play devil's advocate; play "What about Me?"; and require victims to be blameless.

1. MAKE THE LANGUAGE PRETTY

As an example of making the language pretty, consider calling attention to domestic violence as part of a critique of an ideology that sees women as for use by men. The ideological resistance involves redescribing domestic violence as "domestic disputes" as a way to absorb the criticism and deflect it: what happened within the walls of a private house has nothing to do with societal values and norms; it is not systematic. It is merely individuals disagreeing, or at most behaving badly.

The function of this instance of making the language pretty is to deflect from the systematic nature of the phenomenon and to write the injustice out of the story.

2. PLAY DEVIL'S ADVOCATE

An example of this technique is when we resist the ideology that has it that people who are poor are so because they are lazy by pointing out that many people who live below the poverty line are actually working more than one job and still cannot make ends meet. This is where the technique comes in. The interlocutor then offers an alternative explanation "just for the sake of argument" of why people may be poor, instead of engaging in the counternarrative. This alternative explanation, again, lays the blame for poverty on the individuals and their choices. As with *make the language pretty*, this technique functions to deflect from the systematic nature of the phenomenon and lay the blame on individuals and their choices.

3. PLAY "WHAT ABOUT ME?"

This technique shifts the conversation back to the dominant group. Consider men's groups organized around the possibility that a man might be falsely accused of rape. Instead of talking about sexual violence of women, they shift the conversation back to themselves and the possibility that they may be falsely accused. Another example of this technique is when a mansplainer feels guilty about having mansplained and all the energy goes to making him feel better instead of attending to what he did. I take the function of this technique to keep

the limelight always on the member of the dominant group and not give the member of the oppressed group equal standing.

4. REQUIRE VICTIMS TO BE BLAMELESS

Examples of this technique are when a woman is raped and someone says that she was asking for it because her skirt was too short. This sort of technique resists the charge that an injustice took place by pointing out features of the victim or their history that are seen as in some way blameworthy. Then, since they are blameworthy for those features, they deserved what they got. I see this technique as in line with an ideological defense that blames whatever happens on the victim's own choices and behavior and denies any systematic injustice. Moreover, since the focus is squarely on the victim and their behavioral history, it also prevents us from seeing the continuity between their predicament and ours. They are the bad apple. Bad things happen to bad apples, not to good ones. We are safe.

5. #METOO

Why has the #MeToo movement had the success it has? Lindemann herself suggests that what accounts for the success is the number of shots fired, given that the ideology is false. And that is certainly part of it, but more is required.

Naming. Having *hermeneutical resources* to describe the phenomena is key, as all consciousness-raising groups know. People get together and discuss their experiences and collectively develop linguistic and conceptual tools to identify phenomena that are harmful to them. The #MeToo movement relies on feminist work and activism around sexual harassment and sexual violence that has taken place since the seventies. This work has involved coming up with linguistic and conceptual tools, educating others about the phenomena, and fighting for legislative and social improvements.

Acknowledging systematicity. Recognizing the pervasiveness and the *systematic* nature of sexual harassment and sexual violence is another essential element. A precondition for that is having the words and concepts to name and describe the phenomenon, of course, but then we need to be able to spot the phenomenon when we encounter it, and acknowledge it. Here, facing shame is essential. And we face shame by realizing the systematic nature of the phenomenon and that we, individual people, are not the ones to blame. Great risk can be involved in stepping forward and owning that a certain harmful thing has happened to you. This is why small signals of solidarity, whether they be pink ribbons to show solidarity with breast cancer survivors or hashtags such as #MeToo, play a large

role: they lower the risk in coming forward and, therefore, enable more people to do so. Only then does the systematic nature of the phenomenon become visible. And only then do we see the continuity between our predicament and those around us. Victims are not bad apples. Othering victims cannot keep us safe. We can only be safe if we join together.

The above steps, naming and acknowledging systematicity, are important. But what do we do next? We need to resist the ideological apparatuses and absorption techniques in action. We need to identify the phenomena when they happen, and armed with an understanding of how these techniques function, disable their functioning. Lindemann's essay is part of that work. The techniques she identifies give us tools to resist: we recognize that a certain maneuver is being made and understand the function and effectiveness of that maneuver. Feminist work, both popular and philosophical, offers us tools to identify ideological apparatuses of this sort. For example, recently Rebecca Solnit has made the concept of *mansplaining* part of popular culture in the United States.[3] It is not only a concept that helps us make sense of our experiences, but it also makes us spot certain power differentials at work around us (think of observing a heterosexual couple on a first date where the man talks at the woman for hours). Similarly, Kate Manne has offered us the concept of *down girl moves*, which are ideological apparatuses to put women in their ideological place.[4] Both of these are cases of informal apparatuses, but laws or regulations can also keep people in their place, as can their material effects.[5]

Resistance requires not only alternative explanations, or counterstories, for the oppressive arrangements we live with. It also requires that we attend to the ideological apparatuses and techniques that help maintain the status quo and find ways to resist them. Lindemann's discussion of the various techniques of ideological absorption are part of that work. Drawing on that work, we can develop *countertechniques*. For example, when someone calls a case of domestic violence a "domestic dispute," we can insist that mere disputes do not involve being beaten unconscious. When someone plays the devil's advocate, we can say, "That is an interesting theory, but let's linger on the fact that people working two or three jobs cannot make ends meet." When someone wants to turn the light back onto the dominant group by considering the possibility that a man may be falsely accused of sexual harassment or rape, we can insist on keeping the light on the victim. And when the victim is required to be blameless, we can point

---

3   Solnit, *Men Explain Things to Me*.

4   Manne, *Down Girl*.

5   The interplay between the ideological and the material cannot be ignored, but I am not attending to that here.

out that what happened to them is completely out of proportion to whatever they supposedly did that could be blameworthy (we can also contest the blame-worthiness of some of those things as well, such as wearing a short skirt). Even an eye-for-an-eye account of just punishment has the thief who steals a pack of Marlboros suffer the equivalent of a loss of a single pack of cigarettes, not three rounds of bullets to the stomach.[6]

Using *countertechniques* such as the ones above is an act of ideology resis-tance and can, in turn, result in sanctions. It can be uncomfortable to behave badly (ideologically). And it certainly is unpopular with those invested in the status quo. But there are areas where such deliberate ideology resistance is, and has been, effective. For a current example, we can think of *bystander training* as training in countertechniques to fight oppressive ideologies. The point of by-stander training is not so much to convince the person who is spouting false ideological explanations or policing others, but to dismantle the effectiveness of the narrative or policing mechanism. We may not be able to change everyone's mind, but we can empower those around us to resist the ideology with us. And there is strength in numbers.[7]

San Francisco State University
asta@sfsu.edu

REFERENCES

Althusser, Louis. "Ideology and Ideological State Apparatuses." In *Lenin and Philosophy and Other Essays,* translated by Ben Brewster. London: New Left Books, 1971.

Geuss, Raymond. *The Idea of a Critical Theory: Habermas and the Frankfurt School.* Cambridge: Cambridge University Press, 1981.

Lindemann, Hilde. "Counter the Counterstory: Narrative Approaches to Nar-ratives." *Journal of Ethical and Social Philosophy* 17, no. 3 (June 2020): 286–98.

Manne, Kate. *Down Girl: The Logic of Misogyny.* Oxford: Oxford University Press, 2017.

Solnit, Rebecca. *Men Explain Things to Me.* Chicago: Haymarket Books, 2014.

---

6   These are all examples of countertechniques. The reader may think of others that may be more effective.

7   I thank the editor, Mark Schroeder, for helpful comments on this essay, and Hilde Lindemann for an inspiring piece. Naturally, I am the only one responsible for the views herein and any remaining errors.

IT'S COMPLICATED

THE COMPLEXITY AND POWER OF LINDEMANN'S NARRATIVE FRAMEWORK

Marya Schechtman

IT IS NO SURPRISE that it is difficult to get oppressors to take up a counterstory, that is, to replace a picture of the world in which the social organization that generates their privilege is natural and justified with one that recognizes its injustice. While the *motivations* for those in privileged positions to hold on to an oppressive master narrative may be obvious, the *mechanisms* by which these narratives are maintained are less so. It is this timely and important matter that Hilde Lindemann explores with characteristic force and clarity in "Countering the Counterstory," which considers in detail how master narratives keep the counterstories that challenge them from getting traction.[1]

Lindemann begins by laying out some of the reasons it is difficult for counterstories to emerge as challenges to master narratives in the first place—the fact that master narratives are organic ensembles, that they constitute a worldview, and that they are epistemically rigged, naturalized, and/or privatized. The bulk of her paper, however, is devoted to explaining why it is that even when a counterstory is able to emerge as a challenge to a master narrative, it is likely to flounder on the master narrative's ability to "assimilate opposition," and so faces an uphill battle in finding uptake.[2] Lindemann discusses several tools of assimilation: making the language pretty, playing devil's advocate, playing "What about Me?", requiring victims to be blameless, and changing the subject. In each case she supplies a compelling account of how the mechanism works and convincing examples of what it looks like in action. Despite the formidable obstacles for those trying to launch a counterstory, however, Lindemann ends with a ray of hope. Counterstories can and do get taken up, and change does occur. The #MeToo movement is her example of this possibility. While the battle is by

---

1 Lindemann, "Counter the Counterstory."

2 Lindemann, "Counter the Counterstory," 291.

no means won, Lindemann says, there is reason to think that the counterstory about women that #MeToo offers is taking hold.

The analysis offered here is forceful and inspiring, and I have little to add to what Lindemann says about the mechanisms for assimilating counterstories. I will therefore use my commentary to drill down a bit on the message of hope she offers at the end, which is somewhat less fully developed. While the possibility of finding uptake for a counterstory is urged with great conviction, and the example of #MeToo is convincing, the description of the assimilation mechanisms available to master narratives is so powerful that one does wonder just how they can be overcome. In providing her concluding words of inspiration, Lindemann tells us that the "reason why any counterstory has a chance of succeeding, despite whatever defenses a master narrative might set up against it, is that oppressive master narratives are always *false*. That means there are always cracks in the story" and that "if you shoot enough counterstories" at such false narratives the cracks will show and the narrative will ultimately break apart.[3] This sounds right, but what we have seen in the pages before is a powerful account of all of the mechanisms master narratives have at their disposal for smoothing over and patching these cracks as they appear, thereby keeping the truth at bay. There is thus some need to consider how having "enough" counterstories will ultimately allow the truth to prevail. What I will offer here is not any kind of answer to this question, but only preliminary reflections on some of the complications involved in trying to provide one.

One obvious way to approach the question of how a counterstory can succeed is to look at the example of the #MeToo movement offered in the paper and consider how it managed to gain traction. The story about women it counters was *always* false, after all, and it is not as if counterstories to this narrative had not been offered before, so why did this one get taken up? There are undoubtedly important lessons to be learned by looking at the particular circumstances surrounding this movement. A great many things might be said about the political moment in which it occurred, or about celebrity involvement, the use of social media, and many other factors. Crucially important as I take this project to be, here I am going to take up the different, and somewhat more abstract, question of just what uptake of a counterstory consists in, in the case of #MeToo but also more generally.

This turns out to be a complicated matter, worthy of far more attention than I can give it here. If we are to understand how the assimilation mechanisms outlined in the paper can be overcome, it is necessary to know what it means to overcome them, what success looks like. Thinking even briefly about this ques-

---

3   Lindemann, "Counter the Counterstory," 297.

tion reveals that there are many facets and kinds of success, and that keeping this in mind is likely to be crucial in appreciating all of the tools that might be deployed in getting traction for a counterstory. In what follows I will merely note some of the questions and complications that arise without taking a position on any of them. The idea is to point to a fruitful research program implied in Lindemann's paper, rather than to begin carrying it out.

One natural way of thinking about what it is for a counterstory to get taken up is for a significant number of those who hold the master narrative to come to see the world through the lens of the counterstory instead. But if we try to think about what, exactly, that would mean it immediately becomes evident that there is not going to be a simple and straightforward answer to this question. Those who hold the master narrative do not constitute a monolithic group, and the counterstory is not a single thing. To begin, the ways in which those who subscribe to the master narrative opposed by #MeToo do so is going to be diverse. There will be those who are actively and deeply invested in its picture of the world, those who are at some level uncomfortable with the privilege it affords them but distract themselves from these worries using methods of assimilation, and those with infiltrated consciousness who are harmed by it but hold it anyway.

Lindemann makes it clear that the notion of a narrative as it is used in this context is also diffuse. Narratives, she says, are as "capacious, as cluttered and untidy as a Victorian attic."[4] She tells us, further, that

> talk of *a* master narrative is really just a manner of speaking, because the term does not designate a single story with a specific plot and a fixed cast of characters. Instead, they are ensembles of repeated themes that take on a life of their own. Fragments of history, biography, film, fables, jokes, and similar narrative forms ring changes on the theme, as do proverbs, music, advertising slogans, and other cultural artifacts.[5]

Something similar is presumably true of counterstories. Indeed, this seems clearly the case when we think about the #MeToo movement, which is made up not of a single story, but of the many stories of the individual women who use the platform to share their experiences. These are stories that can and do differ in sometimes radical ways. It would be difficult and distorting to extract a single narrative about women and their circumstances from these individual stories, and I know of no one who has suggested that we should.

Given these degrees of freedom, it seems clear there is no single, simple outcome that constitutes the uptake of the #MeToo counterstory by those who had

4   Lindemann, "Counter the Counterstory," 289.
5   Lindemann, "Counter the Counterstory," 288.

held the master narrative. This suggests that the strategy for getting those who hold this narrative to change their perspectives will likely vary depending upon which part of this space of possibilities we find ourselves in. From this recognition several further observations arise.

To begin, we can probably assume that there is some group of those currently holding the master narrative who will never take up the counterstory no matter how many times the truth is spoken to them or how forcefully. In her paper, for instance, Lindemann mentions Paul Elam, founder of A Voice for Men, who has said publicly that he would always acquit if he were a jury member for a rape trial, no matter what the evidence.[6] If one's goal is to get as many people as possible who hold the master narrative to take up the counterstory instead, it seems reasonable not to expend too much effort on this group.

Those who are guiltily enjoying privilege and rationalizing to enable themselves to keep holding the master narrative seem a more promising target for change. Nevertheless, some difficult and fraught questions arise about the best strategy for addressing this group, questions that are in part possible because of the imprecise nature of the counterstory itself. On the one hand, we might think that the project of getting those in this group to take up the counterstory should proceed along the lines of an intervention, in which the defense mechanisms of assimilation are challenged through repeated confrontation with undeniable truths that speak against the narrative. If this is the strategy, it might seem best to focus on stories without ambiguity, those that make the mechanisms of assimilation most awkward to apply. The cases of Aziz Ansari and Al Franken, for instance, may be easier for many people who hold the master narrative to assimilate than those of Jeffrey Epstein or Bill Cosby or Harvey Weinstein. Some have therefore argued that in launching a counterstory the movement should focus on the more extreme cases, which are especially difficult to explain away using the mechanisms described in Lindemann's paper. Bringing in cases that are easier to assimilate, it is suggested, risks backlash and disagreement that will only make the master narrative more entrenched.

There is, however, also a great deal to be said against this strategy. Plausibly, focusing only on cases that many of those holding the master narrative will have to recognize as egregious and downplaying those that might be perceived as ambiguous or assimilable under the category of "boys will be boys" in fact makes maintaining the bulk of the master narrative all too easy. It allows those who hold the master narrative to denounce obviously criminal behavior, believing that they have listened to reasonable challenges to their worldview, without forcing them to confront the widespread existence of arguably more subtle

6   Lindemann, "Counter the Counterstory," 293.

but equally damaging forms of oppression. The whole point of a *counterstory*, as opposed to a mere list of bad behaviors, one might argue, is that it shows that the behaviors that many of those who hold the master narrative acknowledge as unacceptable are not only more numerous and widespread than they might have thought them to be but are also, crucially, directly connected to attitudes and behaviors they are willing to dismiss as relatively harmless. Disagreement between these two strategic positions has been expressed in many venues as the #MeToo movement has gained steam, and while these disputes are messy and painful, there is reason to hope that in the end they will bring important issues to the fore and lead to real progress.

Things are perhaps more straightforward in the case of those with infiltrated consciousness who have bought into the master narrative despite being disadvantaged by it. For many in this situation, seeing their own circumstances described in ways that resonate clearly and make vivid to them what they are already experiencing might be all that is needed to quickly change their gestalt. A vocabulary for thinking about one's experience and validation that others see things in the same way may well be enough to get this group to take up the counterstory.

There is thus a great deal of complexity concerning the question of what it means to get those who currently hold the master narrative to take up a counterstory. Beyond this, however, it is plausible to assume that a counterstory does not get traction only, or even primarily, through this kind of change. Another way of thinking about what it means for a counterstory to get traction is that it does so by making the voices of those who already see the world through its lens more audible and impactful. What makes the master narrative the master narrative, after all, is its dominance and influence. A movement like #MeToo, by collecting and telling individual stories in a highly visible way, ensures that these stories and the more general picture they paint will play a role in the public discourse that a smattering of isolated stories without this kind of platform likely would not. Here it is not in the first instance a matter of directly convincing anyone who does not now accept the counterstory to see things differently, but rather of diluting the voice of the master narrative with an alternative. Especially important, perhaps, is the way in which the audibility of counterstories can impact those whose worldviews are just being formed. Those who grow up hearing the counterstory robustly expressed are, it is to be hoped, less likely to be easily led to see the world exclusively in terms of the existing master narrative.

While this picture of what it means for a counterstory to get uptake is somewhat simpler than that of convincing those who currently hold the master narrative to change their minds, questions still arise. One concerns precisely how

we are to think about the role of truth in this dynamic. Lindemann makes it clear that the ultimate destruction of master narratives rests on their falsity and, presumably, counterstories prevail in the end because they are true. In the mechanism just described, however, the uptake of counterstories is a matter of having the narrative sufficiently present in the public sphere to influence conduct, judgments, and the outlook of future generations. This seems to imply that it is the frequency and centrality of the telling of the counterstory, rather than its more legitimate claim to truth, that gives it traction.

This description of the situation is, of course, overly simple. There are many ways in which questions of truth could play into this general picture of how counterstories are taken up. One might argue, for instance, that the reason there are so many instances of the counterstory is precisely because it is true, and so that it is the truth of the counterstory that allows it eventually to drown out the master narrative. Another place in which questions of truth might enter into this picture is in explaining the impact of the counterstory on future generations. Here the idea would be that, once the story is out there to be heard in a way that allows real comparison with the counterstory, the master narrative's falsity will be obvious to those who have not already been socialized into seeing it as true via the mechanisms outlined in Lindemann's paper. Still, there seems to be an underlying question about exactly what falsity amounts to on this picture. We have an intuitive idea of what it is to have a false narrative about others that is natural to apply here, but it is not evident just how it interacts with the strong role of social factors in constituting identity that is at the heart of Lindemann's view. There are many possibilities, of course. A true narrative might be one that maximizes flourishing, or one where there is optimal harmony between one's narrative about her own identity and the narratives of others, or it might be one that is most internally consistent or, most likely, some combination of these and other factors. There are many important resources in Lindemann's works and beyond for making the relevant notions of truth and falsity more precise, and this seems to me important work well worth undertaking.

These, then, are some of the questions and complications that arise as we try to think about what it means for a counterstory to get traction and how this is achieved. As advertised, I have drawn no actual conclusions about what uptake involves. I hope to have provided a sense of just how complicated this framework is and, in particular, how much remains to be investigated in thinking about how counterstories can be allowed a fair hearing. This is by no means intended as a criticism of the view presented here, or a complaint that it is incomplete. To the contrary, it is meant to display its power. What seems evident is that there is no one thing that is *the* counterstory to a master narrative, nor a single circumstance

that counts as its uptake, and this is exactly what Lindemann's view suggests. It is precisely the diffuse, dynamic, and organic nature of master narratives that makes them so pervasive and difficult to counter. But these same features provide counterstories with a wide range of tools for combating these narratives, and different places and ways for them to start taking hold. This is what those who would counter these narratives need to recognize and develop, and this is what Lindemann displays so beautifully. Her message of sober realism infused with hope is timely and welcome.

University of Illinois at Chicago
marya@uic.edu

REFERENCES

Lindemann, Hilde. "Counter the Counterstory: Narrative Approaches to Narratives." *Journal of Ethics and Social Philosophy* 17, no. 3 (June 2020): 286–98.

REPLY TO MARK LANCE, ÁSTA,
AND MARYA SCHECHTMAN

Hilde Lindemann

MANY THANKS to my commentators, who expand on my ideas, fill in some missing gaps, and gently correct my blunders in "Counter the Counterstory."[1] They all, in one way or another, press down on my overoptimistic assessment of reasons to hope that counterstories will prevail in the end. Mark Lance focuses on the institution of stock characters, highlighting some of the pressure points and vulnerable joints where there is a possibility of social pushback against invidious master narratives. Ásta helpfully argues that more is needed to dismantle a master narrative than pointing to its falsity—she rightly points out that hermeneutical resources to describe the harmful phenomena are necessary, as is the need to recognize the systematicity of the oppression, to attend to the ideological apparatuses and techniques that maintain the status quo. Marya Schechtman, too, reflects on the difficulties of creating the conditions for successful counterstories. She points out that those who accept a master narrative are not a monolithic group. Some buy into it wholeheartedly, others are made uncomfortable by their privilege but distract themselves using methods of assimilation, and others with infiltrated consciousness believe it anyway. So, successfully dismantling the master narrative depends on to which group those whom the counterstory is aimed at belong.

Let me start my reply with Lance's paper, as he takes up the concept of stock characters that must begin any discussion of master narratives and counterstories. He highlights some "vulnerable joints and pressure points" where resistance to oppressive narratives employing stock characters has the best chance of succeeding, and then illustrates some of the forms of resistance.[2] His point that stock characters are not taken up literally is exactly right—when we consistently or momentarily identify people via stock characters we do it inexplicitly, out of habit. Also worth noting is Lance's observation that some stock characters

---

1 Lindemann, "Counter the Counterstory."

2 Lance, "Counterstories, Stock Characters, and Varieties of Narrative Resistance," 300.

are important because of the rules they are supposed to act by, while for others what matters is their characteristics. Let me add that the strength of still other characters is their relationships—the Virgin Mary, for example, or Romeo and Juliet—and for yet others, like Judas, what matters is a single deed.

Important to note here is that any stock character depends on, cannot exist without, a specific form of life. It is within that form of life that the practice of holding and letting go takes place and so, as Lance points out, if the form of life disappears, that particular character can no longer be inhabited. Where the form of life is damaging—think of the antebellum South—it is good that the character can no longer exist. This, I think, is one of the pressure points Lance is referring to: here, chipping away at the form of life may be a better form of resistance than merely telling counterstories, although such narratives may and do assist in that chipping.

I was particularly taken by a throwaway line in Lance's discussion of ACT UP. He is absolutely right that public performance counts as narrative. The suffrage movement in England, for instance, often employed the figure of Joan of Arc on a white horse to lead its parades, while suffragettes chained themselves to the fence outside Parliament as they enacted the bondage that disenfranchisement creates. Indeed, it has been argued that feminist theater itself has its roots in suffragist demonstrations. The idea was always to shift understandings of what the narratively constructed identities of the players are, whether as gay men, black people, disabled people, Native Americans, or women.

Of equal interest in the ACT UP discussion is the possibility of clashing counterstories, as his examples of the prettified story of compassion for the dying as opposed to ACT UP's story of "Everyone loves a dying fag!" attest. Certainly some counterstories are better than others. The ones that see certain people in a group as exceptional (the Tuskegee Airmen, for example) might liberate those people but do nothing for the rest of the group. Counterstories that pathologize specific individuals rather than portray them as moral perverts, as has happened to gay and trans people, nevertheless belittle them and take away their dignity (hence the "dying fag" criticism). Lance has offered a thoughtful reflection on all this, and I thank him.

Ásta's redubbing the countering of counterstories as "ideological absorption" is an apt reminder of how the *policing* mechanism of oppressive ideologies works.[3] I, ever the optimist, contented myself with placing hope on the falsity of oppressive master narratives as the reason why they are vulnerable to attack. To this Ásta, more gently than I deserve, suggests that pointing out such falsity is only a first step, and that much more, presumably because of the policing, is

---

3   Ásta, "Ideological Absorption and Countertechniques," 310.

needed if the counterstory is ultimately to succeed. She is right, of course. Merely pinning one's faith on the belief that unmasking the falsehood is enough is naive. Perhaps only an academic would be so foolish.

It is in her discussion of the #MeToo movement that Ásta lays out what more is required. Having the hermeneutical resources to name the injustice is certainly, as she says, key. For example, it was not until Catharine MacKinnon and others hit on the term "sexual harassment" for the repeated, relatively minor sexual liberties men were taking with women in the workplace that people had a way to speak of and understand the injustice inflicted in this way. Marital rape is another such term, as is mansplaining. But this naming is not confined merely to the original injustice—it also extends to techniques of ideological absorption, as Kate Manne deftly shows by popularizing the term "himpathy."

Acknowledging the systematic nature of the oppression is also important. In Ásta's discussion of four of the techniques for ideological absorption I identify, she repeatedly hammers home how they each paper over, ignore, or deny the systematicity of the injustice. The concept of "face shaming" she employs here is useful, because it allows for systematic resistance to systematic injustice. As an example, a black woman approached my white daughter in a parking lot recently to tell her how much it meant to her to see my daughter's Black Lives Matter bumper sticker displayed on her car. She said it helped her feel less alone to know that at least some white people cared about her people. And Ásta is right to point out that in banding together to resist, the systematic nature of the oppression becomes visible.

It is definitely not enough to unmask the falsity of an oppressive master narrative. We need countertechniques that firmly resist the techniques of ideological absorption by refusing to allow the systemic nature of oppressions to be ignored. The real comfort here is not that master narratives are false, but that for resisters, as Ásta concludes, there is strength in numbers.

Schechtman considers how having "enough" counterstories will ultimately allow the truth to prevail.[4] Such considerations as the political climate, social media, and celebrity involvement play a role, but Schechtman is more interested in the theoretical question of what uptake of a counterstory consists in—what success in this regard looks like.

Well, one measure of success would certainly be that a significant number of people who bought the invidious master narrative started identifying people via the counterstory instead. The difficulty is that master-narrative believers are not a homogeneous group, nor is there only one counterstory.

And here is where Schechtman makes an interesting and helpful move. First,

4   Schechtman, "It's Complicated," 316.

she notes that some people are "actively and deeply invested" in the master narrative, some subscribe to it but are uncomfortable with the privilege it yields them and distract themselves by the use of ideological absorption, and still others are harmed by it but buy into it anyway.[5] Next she points out that counterstories too vary in different ways. So success depends on who is listening, and how hard, to what counterstory.

Some people, of course, will not listen to any counterstory no matter how shrewdly it is pitched, so there is no use wasting time on them. It therefore makes better sense to work on the uncomfortable group. The obvious strategy to employ here might seem to be to hammer away at them with repeated confrontation of undeniable truths that refute the master narrative, and in this way break down their defenses. These counterstories would target black-and-white villains, such as Jeffrey Epstein or Bill Cosby, rather than more ambiguous cases such as Al Franken, since Franken-type cases would be easier to assimilate into the master narrative by characterizing them as "not so bad," or "He's a comedian whose joke got out of hand." The idea is to stay focused on what is clearly terrible, rather than what is more easily explained away. We don't want to risk backlash, do we?

Actually, says Schechtman, we do. She quite rightly points out that resisting the Jeffrey Epsteins leaves a great deal of the master narrative untouched. It lets those who have bought into the master narrative denounce outrageous behavior without having to confront the subtler but still harmful practices of oppression. In fact, the focus on the blatant turns those bad actors into the exception, making the counterstory, like the one about the Tuskegee Airmen, stand-alones that leave everything else in place. The only difference is that here the protagonist of the counterstory is not the good guy but the bad guy. What is needed instead is to show that the bad guy's behavior is connected quite directly to attitudes and behaviors that are dismissed as relatively harmless.

Schechtman thinks a different strategy might be employed for those with infiltrated consciousness whose identities are damaged by the very master narratives they embrace. Here, hearing their circumstances described in ways they can easily identify with might be all that is needed to switch out at least part of their worldview. I recall, for instance, chatting with a dear friend in the early 1970s who, like me, was a stay-at-home mom. She remarked in passing over our morning coffee that she did not see why, just because she took care of her son at home, she should always be the one to clean the bathroom. I had never thought about it before, but that simple observation flipped the switch for me, and from then on, my husband cleaned the bathroom too.

Schechtman also insightfully notes that to get a counterstory to succeed may

---

5   Schechtman, "It's Complicated," 317.

not be so much a matter of directly convincing those who do not accept it to see things in a different light, as of gradually and generally diluting the master narrative itself. This can take several generations, as those of us telling counterstories in the seventies taught them to our children, who are now passing them on to their own children, so that by now they are accepted as matter-of-fact truths. Here, Schechtman agrees with Ásta that it is the frequency and ubiquity of the counterstory, not the truth of it, that gives it traction, although the truth of it contributes to its ubiquity when compared to the diluted master narrative.

It is then the diffuse and dynamic power of the counterstory that provides the resources for combating the diffuse and dynamic power of the master narrative. These dynamics are going on all around us, and may give us guarded hope that resistance is *not* futile, whatever the Borg may say to the contrary. Counterstories *can* be heard. Master narratives *can* be dismantled, despite their tremendous power to assimilate. I am grateful to my commentators for saying more clearly, and developing more powerfully, the ideas I have been working out about this for years.

Michigan State University
hlinde@msu.edu

REFERENCES

Ásta. "Ideological Absorption and Countertechniques: Comments on Lindemann." *Journal of Ethics and Social Philosophy* 17, no. 3 (June 2020): 310–14.

Lance, Mark. "Counterstories, Stock Characters, and Varieties of Narrative Resistance: Response to Lindemann." *Journal of Ethics and Social Philosophy* 17, no. 3 (June 2020): 299–309.

Lindemann, Hilde. "Counter the Counterstory: Narrative Approaches to Narratives." *Journal of Ethics and Social Philosophy* 17, no. 3 (June 2020): 286–98.

Schechtman, Marya. "It's Complicated: The Complexity and Power of Lindemann's Narrative Framework." *Journal of Ethics and Social Philosophy* 17, no. 3 (June 2020): 315–21.