JOURNAL *of* ETHICS
& SOCIAL PHILOSOPHY

ARTICLES

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.

THE NORMATIVE PLURIVERSE

Matti Eklund

Are there normative properties? Normative realists say yes; various opponents say no. An issue that has been much less discussed is whether there might be a certain kind of *multitude* of normative properties and what the metaethical consequences are of that kind of normative pluralism.

Here, intuitively, is what is at issue. If all properties are nonnormative, then there are no properties that, so to speak, tell us what to do. If there are normative properties, then some properties tell us what to do. But if there is a *multitude* of properties telling us what to do, providing *different* instructions, then we are back at square one.

This paper will be devoted to this kind of normative pluralism and the upshot of taking it seriously. I have just described briefly what this is about. But the aim of the paper is to provide a more careful characterization of the kind of pluralism at issue and of what its significance may be.

Throughout, I will be concerned with relating—and criticizing—what Justin Clarke-Doane has said about these matters.[1] However, the main aim is constructive: developing the relevant form of normative pluralism and gauging its significance. After having, in section 1, set out the issue, I turn to discuss how exactly we should conceive of the present kind of normative pluralism (in sections 2–5), and then how exactly we should think of the upshot of normative pluralism (sections 6–7). While Clarke-Doane does not much problematize normative pluralism, I think there are serious problems here. And I will argue that while consideration of normative pluralism has serious consequences for how we should think about the normative, it is a delicate matter just how to think of these consequences. More specifically, I will argue that—perhaps appearances to the contrary—the challenge is in effect the same challenge I have already presented in my book *Choosing Normative Concepts*. The normative pluralism at issue is the same kind of pluralism (although I did not use that label) as that

---

1 Clarke-Doane, "Objectivity in Ethics and Mathematics," "Objectivity and Evaluation," and *Morality and Mathematics*.

which is at issue in my book. And if there is an interesting philosophical upshot (as I believe there is), it is effectively the same as that discussed in my book.

1. MATHEMATICAL AND NORMATIVE PLURALISM

There is a certain kind of view in philosophy of mathematics—a kind of *pluralism*, I will call it, although the pluralist label has been used for many different kinds of views—according to which the mathematical universe is as densely populated as can be.[2] Roughly, for any logically coherent conception of some mathematical entities, there are entities that satisfy that conception. A careless formulation of the view is that every consistent mathematical theory is true. That cannot be exactly right, for one consistent mathematical theory says that sets satisfy the continuum hypothesis (CH), and another says that they do not, and these two theories cannot both be true. Rather, what the pluralist in question holds is that, for example, there are some set-like entities that make CH (meaning the version of CH that concerns them) true and some others that do not make CH (the version of CH that concerns *them*) true. I will ride roughshod over the problems in adequately formulating mathematical pluralism.[3] I trust that it will be clear enough for present purposes what mathematical pluralism amounts to—and in any case mathematical pluralism will not be my main topic.

In the works mentioned above, Clarke-Doane asks: What about a pluralism in metaethics corresponding to this mathematical pluralism? I will here critically discuss what Clarke-Doane says about the consequences of this kind of normative pluralism, and, more importantly, offer a constructive suggestion regarding the nature and import of a significant kind of normative pluralism.

Clarke-Doane asks us to suppose that all consistent ethical theories "are true," but "of different entities."[4] The mathematical pluralist believes in a *mathematical pluriverse*; the normative pluralist believes in a *normative* pluriverse. He then compares the consequences of mathematical pluralism with the consequences of normative pluralism. He thinks that mathematical pluralism in a certain way *trivializes* mathematics.[5] Given mathematical pluralism, there are genuine logical questions about what follows from what axioms, and questions about which theory, for example, best captures our concept of set, but no "peculiarly set-the-

---

2   See, e.g., Balaguer, *Platonism and Anti-Platonism in Mathematics*; Field, "Which Undecidable Mathematical Sentences Have Determinate Truth-Values?"; and Hamkins, "The Set-Theoretic Multiverse."

3   See sec. 6.2 of Clarke-Doane, *Morality and Mathematics*, for discussion of such matters.

4   Clarke-Doane, "Objectivity in Ethics and Mathematics," 103.

5   Clarke-Doane, "Objectivity in Ethics and Mathematics," 103.

oretic questions" are genuine. Turning to normative pluralism, Clarke-Doane asks whether ethics could be "trivialized similarly."[6] He responds as follows:

> Imagine that a philosopher convinces us that, contrary to all appearances, ethics too is like geometry—that every consistent ethical theory is true, albeit true of different entities. In addition to goodness, obligation, and so on, there is shgoodness, shobligation, and so on. Indeed, for every logically consistent ethical theory, there are corresponding properties, and all of them are instantiated "side by side." Knowing that there are logically … consistent formulations of both deontological and consequentialist ethical theories, we conclude that each is true (albeit of different entities). Is our deliberation as to whether we ought to lie when utility would be maximized thereby trivialized (and likewise for every question on which logically consistent ethical theories diverge)?
>
> It is hard to see how it could be. A general—even if not universal— rule is that if we conclude that we ought to X, then we cannot continue to regard the view that we ought to not-X as on a par. But given that that view is on a par with respect to truth, learning that "we ought to X" is true seems insufficient to resolve our deliberation. While knowledge that any consistent set theory is true, and knowledge that ZF + AC and ZF + ~AC are both consistent, frees us of the question of whether AC, something similar would not seem to hold in the ethical—and, more generally, normative—case. The fact-value gap appears to be even wider than Hume and Moore suggested. Even knowledge of the *normative* facts may fail to resolve a normative deliberation.[7]

Clarke-Doane makes two points here. One is that normative questions are not trivialized by normative pluralism in the way that certain mathematical questions are trivialized by mathematical pluralism. The second—which also complicates the first—has to do with the status of normative truths. He says that given normative pluralism, the view that we ought to X and the view that we ought to not-X are on a par with respect to *truth*, and given this, he thinks, knowledge of the normative truths does not suffice to resolve normative deliberation.

It is tempting to see Clarke-Doane as in effect saying that pluralism presents a challenge for a certain kind of *realism*, given what he says at the very end of the quoted passage. But one must be careful about what sort of realism is being challenged. Often "realism" is used as a label for all views given that normative propositions are capable of mind-independent truth and falsity, and some atom-

6   Clarke-Doane, "Objectivity in Ethics and Mathematics," 103.

7   Clarke-Doane, "Objectivity in Ethics and Mathematics," 103–4.

ic normative propositions are true; and that is also how I will use that label here. There is nothing in what Clarke-Doane says that even suggests that he takes normative pluralism to challenge realism in this broad sense. But there is another view that arguably tends to be endorsed by realists: echoing Clarke-Doane's formulation above, this view is that there are some facts knowledge of which properly resolves normative deliberation—facts that settle the normative question of what to do. This view is challenged by Clarke-Doane. The distinction between realism as first characterized and the view challenged by Clarke-Doane can seem elusive, or nonexistent: If there are objectively true normative propositions, are not the facts they state exactly facts that settle normative deliberation? But in my view, one lesson of the issues that come up is that Clarke-Doane is right, and there is reason to draw a distinction here. I will return to this below, in section 7.

There are other kinds of comparisons that can be made between mathematical pluralism and normative pluralism. It can be held that both sorts of pluralism have the consequence that certain classes of disputes are *merely verbal*: those who have a dispute "about sets" may simply be using "set" to denote different kinds of entities, and those with a dispute "about what is right" may simply be using "right" for different properties. And it can then further be held that whereas this consequence may be acceptable in the mathematics case, it is unacceptable in the normative case. The main thing I wish to say about the topic of verbal disputes is that this simply is not my topic here. I will focus on the sorts of things Clarke-Doane brings up—in his words, whether there are facts that resolve deliberation. But let me also briefly add that the connection between the pluralism at issue and disputes being merely verbal is by no means straightforward. Even if there are, for example, many different set-like entities, it can still be that disputants nearly always in fact think and talk about the same set-like entities and have real disagreements about them. Whether that is so depends on how it is determined what our thought and talk is about, and that is a vexed issue.[8]

2. WHAT IS NORMATIVE PLURALISM?

I agree with Clarke-Doane that attention to a normative pluralism modeled on mathematical pluralism may be of great significance for theorizing about the normative. But I have concerns about how Clarke-Doane conceives of the issue.

8   Balaguer also discusses a normative pluralism modeled on mathematical pluralism in "Moral Folkism and the Deflation of (Lots of) Normative and Metaethics." Much of Balaguer's discussion of this pluralism concerns metasemantic issues. He is not much concerned with the questions about realism and objectivity that Clarke-Doane and I focus on, beyond stressing that the pluralism is consistent with certain forms of realism.

As I will argue in this and the following sections, it is by no means clear what the normative pluralism at issue in Clarke-Doane's discussion might amount to. In section 5, I will present a constructive suggestion regarding how to understand normative pluralism so that it promises to be of significance. I then turn to the question of the exact nature of the significance of considering normative pluralism. Here too Clarke-Doane will be my foil. I will criticize Clarke-Doane's claims, and present an alternative view on why and how pluralism matters.

Before the substantive discussion, some preliminary remarks. First, I will freely go back and forth between different normative notions when illustrating the issues at hand. Sometimes I will talk about "ought," sometimes about "right." The background assumption is that the issues come up equally in either case, so it does not substantially matter which example we focus on. There is a good question regarding exactly which normative notions we could focus on and still raise essentially the same kind of problem. For example, could one use thick concepts to the same effect? Could one use an example with, say, aesthetically or epistemically normative notions to the same effect? But as a first approximation we can say that the notions in terms of which we raise the problem are the so-called thinnest notions, so those expressed by words like "ought" and "right" in their thinnest, so-called all-things-considered uses.

Second, I will disregard that the words focused on are arguably context-sensitive and express different things as used in different contexts, but will trust it to be clear enough for present purposes what uses of these words I focus on. Third, when speaking about "ought" I will sometimes speak of it as ascribing a property. This is simplified in a couple of ways. For one thing, it strictly ascribes a relation rather than a property. In response to that, I note that "property" is also sometimes used to cover relations. For another, one may hold that "ought" really is an operator. But all I need for present purposes is a simple way to speak of the worldly correlate of "ought."[9] I could in principle speak of what propositions p are such that "Op" expresses a truth, where "O" is an ought-operator, instead of speaking of what has the property *ought*.

What is the normative pluralism at issue in Clarke-Doane's discussion? Consider Clarke-Doane's talk of *being on a par*. How should this talk be understood? The view that we ought to X and the view that it is not the case that we ought to X cannot both be true. At most one of these views gets the facts about what we *ought to do* right. (Compare the set theory case: at most one theory gets the facts about *sets* right, even if other theories correctly describe other set-like entities.[10])

9 Where—following Pickel's usage of the phrase in "Naming, Saying and Structure"—the "worldly correlate" of an expression need not be an entity.

10 Given that there is the plurality of set-like entities postulated by mathematical pluralism, it

The normative pluralism at issue cannot be that the view that we ought to X and the view that we ought to not-X are both true. Instead it must be something like: there is a property *ought*, and there is a different but still in some way ought-like property, call it *ought**, such that there is the property of being what one ought* to do, and the properties of being what one ought to do and of being what one ought* to do are not coextensive. But once the normative pluralism is more carefully stated—no more flirting with the claim that there are different equally true views regarding what one *ought* to do—Clarke-Doane's further claims about what normative pluralism yields can seem questionable. Since pluralism is fully compatible with the view that one view is true and the other false, pluralism does not entail that the view that we ought to X and the view that we ought to not-X are on a par. This means that for all that has been said, knowledge of the normative facts can—to use Clarke-Doane's way of putting things—resolve normative deliberation even given normative pluralism as described. Knowledge of what one ought to do can resolve normative deliberation, even if there is also some *other* property, ought*, such that what one ought* to do is not always what one ought to do.

With these initial skeptical remarks as background, consider some more specific concerns along broadly similar lines.

Here is a first point. One dismissive reaction to Clarke-Doane's argument is the following. We simply care more about, and have more reason to care about, what is right, and what ought to be done, than about which mathematical entities are the sets. Given mathematical pluralism, the question "But which are the sets?" remains, and given normative pluralism the question "But what ought I to do?" remains. Insofar as there is a felt difference between the upshots of these kinds of pluralism, the difference is simply that the former question does not seem very important—who cares which ones, of equally real mathematical entities, are *the sets*?—but the latter question does seem important.[11] Even if there

---

is natural to speculate that set talk is referentially indeterminate: our thoughts and practices do not determine precisely which set-like entities we talk about when talking about "sets." This is a complication worth keeping in mind, but it does not conflict with the claim in the main text. A corresponding indeterminacy claim can be made regarding the normative. There are a number of ought-like properties and a number of right-like properties, and our thoughts and practices do not determine precisely which one of these is picked out by our "ought" and by our "right," respectively.

11   Set theory is often appealed to as a foundation for mathematics. There is a question, not in any way trivialized by mathematical pluralism, of which types of mathematical entities are such that a theory of them can serve as a foundation for mathematics. But the pluralist will be apt to hold that different set-like entities are apt to play this role.

are lots of properties for actions to have, of course it matters what I *ought* to do! That is all there is to it.

Second, consider the following property, *ought+* (or *being what one ought+ to do*), where an agent ought+ to ϕ if and only if: she ought to ϕ and ϕ-ing ≠ helping someone cross the street on a Thursday. If sometimes helping someone cross the street on a Thursday is what an agent ought to do, ought+ ≠ ought, but the property ought+ is still ought-like. Ought and ought+ are almost coextensive. But there being this ought-like property in addition to ought itself does not seem to problematize normative deliberation in the least. Normative deliberation concerns what ought to be done, and the possibility of gerrymandered properties like that of what being what one ought+ to do does not immediately do anything to lessen the sense that we are appropriately concerned with what one ought to do. Maybe the properties whose existence is entailed by the normative pluralism that Clarke-Doane describes present no more of a problem than ought+ does. Normative deliberation concerns what ought to be done, and there being in some sense nearby properties does not affect such questions or their significance. (The mathematical pluralism concerned there being mathematical entities corresponding to a variety of *mathematical theories*, and the notion ought+ was not introduced as corresponding to a normative theory. But this is easily remedied. Just consider a normative theory that in addition to whatever else it says includes a fundamental prohibition on helping someone cross the street on a Thursday.)

Third, a different reason to suspect that there is something off regarding Clarke-Doane's comparison is that, whereas the mathematical pluralism is a controversial hypothesis ("Are there really all these set-like entities?"), one can see the normative pluralism as somewhat trivial. One may reason: *of course*, say, the property of happiness-maximization that many utilitarians conceive of as rightness ("the utilitarian's property," as I will refer to it) exists, and of course the deontological property of being in accordance with such-and-such maxims ("the deontologist's property") exists, and I can believe in both these properties even if I will identify at most one with rightness. The mathematical pluralism does not seem correspondingly trivial or obvious. Now, the triviality of normative pluralism is not in itself and immediately a problem for Clarke-Doane. It is fully consistent with his discussion of the consequences of normative pluralism that normative pluralism is trivially true, and that mathematical and normative pluralism differ in whether they are trivial. It is rather that there is this difference between mathematical and normative pluralism, and this difference may make us *suspect* that something has gone awry somewhere.[12]

12   In a number of writings, David Enoch and Tristram McPherson have argued against Scan-

When considering the seeming triviality of normative pluralism, one must distinguish between two questions. One is whether it is uncontroversial that a given property exists; the other is whether it is uncontroversial that this property can be instantiated. It is possible, after all, that some properties exist despite not being uninstantiated, and even necessarily so. In the case of the utilitarian's and the deontologist's properties, the answers to both questions are affirmative (of course setting aside general skepticism about properties). It may be doubtful whether any actions we actually perform have these properties, but it cannot be in serious doubt that it is possible for there to be actions that have these properties. Moreover, the property ought+ is instantiated so long as the property ought is instantiated, and some things we ought to do fail to involve helping someone cross the street on a Thursday.

I have mentioned three ostensibly different concerns with what Clarke-Doane says about normative pluralism. There is what I dubbed the dismissive reaction, the objection from ought+, and the apparent difference in triviality between mathematical pluralism and normative pluralism. These concerns all turn on what exactly normative pluralism comes to. One thing that would deal with these concerns is a formulation of normative pluralism that promises to sidestep these objections. I now turn to the project of finding such a formulation.

3. NORMATIVITY AND NONNATURALNESS

Here is a general strategy for finding a formulation of the requisite kind. One can seek to identify a significant property of properties, X, and a form of normative pluralism that says that there is a plethora of properties that are X. Given the right X, this could in principle get around the problems identified. When it comes to the dismissive reaction: if ought, ought*, ought**, etc., are all X, the continued concern specifically with what *ought* to be done may start seeming unreasonable, for it may be objected to this continued concern that ought*, ought**, etc., are equally X. The objection from ought+ is avoided so long as ought+ and other seemingly irrelevant alternatives to ought are not X. So long as the X-ness of the properties is not trivial, this pluralism is not trivial. Even if the existence of the

lon's broadly quietist view (Scanlon, *Being Realistic about Reasons*) that Scanlon cannot legitimately claim there to be relevant differences between, e.g., reasons and schmeasons, where schmeasons are a non-coextensive alternative to reasons. See Enoch, *Taking Morality Seriously*, ch. 5; Enoch and McPherson, "What Do You Mean 'This Isn't the Question'?"; and McPherson, "Against Quietist Normative Realism." Enoch and McPherson undoubtedly raise an important problem for Scanlon. But I doubt that it is a problem for Scanlon's quietism specifically: for the alternative properties and relations threaten to exist on any reasonable metaphysical view.

properties like the utilitarian's property should be accepted on all hands, their *X*-ness need not be.

This description of the strategy is, to put it mildly, pretty abstract. So let me right away illustrate. One kind of normative pluralism may hold that corresponding to different theories supposed to be about rightness there are (not just different properties but) different *normative* properties. Here *X* = normative. Consider how this applies to the problems mentioned. First, the property ought+ perhaps need not be a normative property, even if the property ought is a normative property. Second, if this form of normative pluralism is true, then it becomes more problematic to say that it can still reasonably continue to care specifically about what we *ought* to do even if there are all these other properties. For if the other properties are normative as well, there may be nothing that distinguishes the property ought as especially worthy of our attention. Third, this would appear to be a less trivial form of normative pluralism. It can be held that it is not that trivial that there is this plethora of *normative* properties. It is one thing to claim that the utilitarian's property exists, another to claim that it is normative. If the utilitarian's property exists then the conjunctive property of having this property and being normative exists, but it will not be trivial that this conjunctive property can be instantiated.

Another illustration of the strategy might emphasize not the normativity of the properties but the supposed nonnaturalness. Some metaethicists are nonnaturalist realists, holding that normative properties are "nonnatural." Very roughly, this means that they are not part of the aspects of the world that can be studied by the sciences. One could focus on a form of normative pluralism according to which there is a plethora of *nonnatural* properties. Maybe ought but not ought+ is nonnatural. And if there are all these nonnatural properties, then again—it may be thought—there is nothing that distinguishes the property of being right as especially worthy of our attention. Other properties are metaphysically special in the same way. And it is by no means trivial that there is this plethora of nonnatural properties, or that they can be instantiated.

Both these illustrations of the strategy suggest themselves rather easily. But neither suggestion is in the end workable. Start with the appeal to normativity. Such an appeal invites the question: What is it for a property to be normative?[13]

A first possible answer to this question is that a property is normative if it can be ascribed by a normative predicate. This in turn immediately invites the question of what it is for a predicate to be normative. One cannot very well answer this question in turn by saying that a predicate is normative by virtue of ascribing

---

13 The discussion in the next few paragraphs is parallel to that in ch. 5 of Eklund, *Choosing Normative Concepts*.

a normative property, because of the circularity. And if one holds that a predicate is normative because it can be conventionally used to guide action, a problem is that the normative pluralism again threatens to be rather trivial. Even a die-hard deontologist may agree that a community of utilitarians can use "utilitarian right" to guide action, and that there can be a convention in that community to this effect.[14] Not only is the normative pluralism trivial: the other, more serious problems also still remain. The property ought+ is still a normative property, so long as some community can use "ought+" to guide action. And given how easy it is for a property to be ascribed by a predicate conventionally used to guide action, the point also remains, for all that has been said: Why should we not continue caring as we have about what *ought* to be done, even if there are other properties that are normative in the very liberal sense this characterization yields?

A second characterization of what it is for a property to be normative might be that a property is normative by virtue of standing in the right relation—metaphysical necessitation, to a first approximation—to (say) the property of a reason, or the property of what one ought to do.[15] But this is a nonstarter in the present context. The right-like properties postulated by the would-be normative pluralist will for trivial and irrelevant reasons not be normative in this sense. A right-like property ≠ right will not stand in this relation to the property of being a reason, but instead at best to some other reason-like property. But that by itself is insufficient to ward off reasonable concerns that rightness more properly warrants our attention than this right-like property.

A third suggestion regarding the normativity of properties is to say that this is a *primitive* feature, not amenable to further elucidation. But whatever we say about this primitivism, it does not seem particularly helpful in this context. For what we are after is a feature, X, of properties such that X satisfies the desiderata mentioned above. And if normativity is a primitive feature in the sense indicated, we get no handle on whether normativity satisfies these desiderata. Second, while primitivism might sound like a natural way to go, the primitivism at issue would have to be of a special kind. It is a familiar view that a property such as rightness may be primitive. But that is not the relevant kind of view. The relevant kind of view is that, for example, rightness has the primitive feature of normativity. This idea, that normativity is primitive, may again sound familiar. But it should be so taken only if understood as amounting to something like the

---

14 I will revisit this kind of reasoning later. Different things can be meant by appeals to "convention."

15 There may be reason to prefer instead appealing to what is part of the *nature* or *essence* of the property (see, e.g., Fine, "Essence and Modality"). The points made in the main text still stand.

view that rightness has the primitive property that it *ought* to be promoted. But this would just be a version of the previous strategy for explicating normativity. The relevant primitiveness idea is instead that there is a property, normativity, distinct from, and not analyzable in terms of, familiar normative properties like rightness and being what ought to be done. That should seem less familiar, and more like a philosopher's invention.

Fourth, the friend of appeal to normative properties might say something like: to call a property normative is to say that having the property *really matters* for purposes of normative deliberation. So to say that there are all of these different normative properties involves saying that all of these different properties *really matter*. But in the present context this is problematic. For example, while informal, the talk of "really mattering" is itself normative, and would naturally be thought to mean something like: is something that *ought* to be taken into regard. But thus understood, the present suggestion faces a version of a problem already mentioned. It will be trivial that *ought* but not *ought** really matters.

Turn then to the appeal to nonnaturalness. What might "nonnatural" be used to mean here? Common glosses are, like the label "nonnatural" itself, negative: to be nonnatural is to not be discoverable by empirical means, to not be part of the natural world investigated by the sciences. But to the extent that "nonnatural" just means this, I do not see that the property of being nonnatural plausibly could be our *X*. To be sure, the hypothesis that there is a plurality of instantiated nonnatural right-like or ought-like properties is nontrivial. So the third problem is avoided. But I do not see what the significance of such nonnaturalness pluralism could plausibly be for normative deliberation. If in my normative deliberation I tend to focus on questions about what ought to be done and you convince me that there are other nonnatural properties in the negative sense gestured toward, why should that bother me at all? Why can I not set this aside as irrelevant, telling myself that these other nonnatural properties are just different properties? Note, lastly, that the property ought+ is plausibly nonnatural in the negative sense if the property ought is. If the property ought is metaphysically nonnatural then ought+ plausibly is too since the latter is a construct of the former. Moreover, one finds out what one ought+ to do by the same kinds of means as one finds out what one ought to do (except to know whether one ought+ to do something one must also know whether it is an instance of helping someone cross the street on a Thursday).

4. BECOMING CONVINCED OF NORMATIVE PLURALISM

My discussion so far has focused on what an interesting form of normative plu-

ralism analogous to mathematical pluralism might be. Attention to how exactly Clarke-Doane introduces the issue reveals that what he is primarily concerned with is not the *truth* of some normative pluralist thesis. Rather, he is concerned with the consequences of us *becoming convinced* that some normative pluralist thesis is true. His reasoning in "Objectivity in Ethics and Mathematics" begins, "Imagine that a philosopher convinces us …" It may be thought that questions about the exact content of normative pluralism then are not crucial, in the way they would be if the *truth* of normative pluralism was at issue.

However, the problems regarding the formulation of normative pluralism do not go away, for there remains the question of exactly what it is that we are supposed to be convinced of.

Moreover, so long as the nature of the supposed normative pluralism remains obscure, pluralism seems not to be so central to the challenge. The point, in general terms, appears to be that the truth of a skeptical hypothesis regarding the normative—there are no normative facts knowledge of which is sufficient to resolve deliberation—would not trivialize normative deliberation. The point could equally well be made by appeal to (our having become convinced of the truth of) normative *nihilism*, the view that there are no normative facts; or, say, radical normative *indeterminism*, according to which the claims about what we ought to do are always indeterminate in truth-value—it is radically indeterminate what the normative facts are.

One reason not to focus on nihilism specifically might be that nihilism cannot actually be believed.[16] But even if nihilism cannot be believed, the general point stands: there are views distinct from any form of pluralism whose upshot is the same as the supposed upshot of normative pluralism. Appeal to indeterminism suffices to make the point.

5. NORMATIVE ROLE

I believe that the discussion in my *Choosing Normative Concepts* suggests a workable way of understanding normative pluralism in the context. That discussion is centered on consideration of possible scenarios like the following:

> *Alternative.* There is a linguistic community speaking a language much like English, except for the following differences (and whatever differences are directly entailed). While their words "good," "right," and "ought" are associated with the same normative roles as our words "good," "right," and "ought," their words aren't coextensive with our "good," "right," and

---

16   See Streumer, "Can We Believe the Error Theory?" and *Unbelievable Errors.*

"ought." So even if they are exactly right about what is "good" and "right" and what "ought" to be done, in their sense, and they seek to promote and to do what is "good" and "right" and what "ought" to be done in their sense, they do not seek to promote what is good and right and what ought to be done.[17]

Here is the philosophical point that I introduce by appeal to such scenarios. Assuming that there is such an alternative, it seems that there is a question to be raised with respect to whether to employ our concepts or their concepts when deciding how to act. At the same time—as I discuss—it is elusive what this supposed "further question" is. Speaking our language and using our "ought" it seems very plausible that one *ought* to use our concepts; but it seems equally plausible that in their sense of "ought" one ought to use *their* concepts. But how then can the further question be asked? If we ask the question using normative vocabulary, we must use some particular normative vocabulary or other and then there is the issue of what justifies using that rather than other vocabulary. If instead we ask it using descriptive vocabulary, we seem to have changed the topic in some problematic way. We were interested in what concepts to use when deciding how to act, not in what concepts ascribe properties with such-and-such descriptive features. I will return to the issue of the supposed further question and its non-statability later.[18]

The way I introduce the topic, the notion of normative role becomes of central importance. A concept's normative role is that aspect of its use by virtue of which it figures in practical deliberation in the way it does. The idea behind stating Alternative in terms of alternative concepts sharing normative roles is this. If the others' "ought" figures in normative thinking in the same way as ours does, and the only difference between these expressions concerns aspects of meaning not directly related to normative matters, then there is a clear reason to suspect that the others' "ought" is as much of a claim to be normatively relevant as our "ought."

It may be useful to compare the issue I raise with a well-known argument presented by P. H. Nowell-Smith:

17 Eklund, *Choosing Normative Concepts*, 18.

18 Given present purposes—getting clear on normative pluralism and its possible upshot—I will not here enter into a more in-depth discussion of the nature of my problem or what may be said in response. For relevant discussion, see Bykvist and Olson, review of *Choosing Normative Concepts*; Leary, "Choosing Normative Properties"; McDaniel, "Matti Eklund's *Choosing Normative Concepts*"; McPherson, "Ardent Realism without Referential Normativity"; Plunkett, "Normative Role, Conceptual Variance, and Ardent Realism about Normativity"; and Eklund, "Reply to Bykvist and Olson" and "Reply to Critics."

Learning about "values" or "duties" might well be as exciting as learning about spiritual nebulae or waterspouts. But what if I am not interested? Why should I *do* anything about these newly-revealed objects? Some things, I have now learnt, are right and others wrong; but why should I do what is right, and eschew what is wrong?[19]

This purports to be a challenge for familiar forms of realism, according to which facts about values and duties serve to guide action. And intuitively there is something to the challenge. But it is elusive what it is. I may actually fail to be interested in values and duties, as the "what if I am not interested?" alludes to; but what the realist is typically concerned with is rather what I *should* be interested in. Nowell-Smith asks a rhetorical question about this at the end of the quoted passage. But should this rhetorical question really be accorded any bite? One might say in response that the right things to do trivially are exactly the ones that should be done, because of how our notions of "right" and "should" are related, so the question does not get a grip. It is here that my way of raising the (or a related) issue comes in: even if it is true that I "should do what is right" there are other, nearby truths—I "should* do what is right*" and I am faced with a choice as to which truths to focus on and to let guide my actions. And if one tries to set aside this challenge by just saying that switching focus from what I should do to what I should* do just changes the topic, my response is to stress that if "should" and "should*" have the same normative role then there is a clear way in which the should*-facts promise to be as pertinent to action as should-facts are.

The central appeal to normative role suggests an alternative formulation of normative pluralism. The relevant normative pluralist hypothesis is that there are non-coextensive predicates all conventionally associated with the same normative role. (Alternatively put, there are different, non-coextensive properties all ascribed by predicates with the same normative role.) This pluralist hypothesis can threaten to have deflationary consequences of the kind Clarke-Doane is talking about. If R_1 and R_2 are two properties picked out by predicates having the normative role associated with "right," one might reasonably wonder what could warrant doing what has R_1 over doing what has R_2. Where ϕ-ing is some type of action, ϕ-ing may be R_1 while not-ϕ-ing is R_2, but, the thought would be, there is a clear sense in which ϕ-ing and not ϕ-ing are on a par. For even if we use "ought" and "right" to, for example, pick out properties under which ϕ-ing falls, we could have used corresponding normative predicates under which not ϕ-ing falls; and there is nothing that normatively privileges the properties we actually pick out using normative predicates over the properties we do not so pick out.

19   Nowell-Smith, *Ethics*, 41.

Return now to the three challenges I presented regarding Clarke-Doane's appeal to normative pluralism. Responses to these challenges are now available. Flat-footedly insisting on being concerned with what ought to be done seems dogmatic if other properties are picked out by possible predicates with the same normative role: What favors concern with ought over concern with those other properties? The property ought+ is relevant if, but only if, it is the semantic value of a possible predicate associated with a normative role associated with the concept RIGHT. Just because we can easily define such a property does not mean that it can be the semantic value of such a predicate. And it is not trivial that, for example, the utilitarian's property is picked out by a possible predicate associated with the normative role of our concept RIGHT. So what we have now found is a normative pluralist thesis that promises to be of the right kind for Clarke-Doane's purposes. It avoids the problems discussed earlier.

One potential complication is worth pausing on. It may seem obvious that the utilitarian's property is picked out by some possible predicate associated with the same normative role as that of our concept RIGHT: a community of utilitarians could use a predicate stipulated to stand for the utilitarian's property, and because these people are all convinced that this property is the rightness property that predicate comes to have, for them, the same normative role as our concept RIGHT. The response to this kind of complication is to distinguish between a predicate's being associated with a normative role as a matter of established semantic convention, and its being so associated simply by virtue of what beliefs are prevalent in the community that uses the predicate. It should be specified that it is the first kind of association with normative role that is at issue.[20]

A possible way of rejecting any challenge based on the kind of normative pluralism now at issue may be to say that normative role does determine reference, so that sameness of normative role guarantees sameness of reference.

In my *Choosing Normative Concepts*, I also consider some scenarios aimed to raise the same questions as Alternative does, but where the different communities' normative concepts do not have the exact same normative roles.[21] If those scenarios indeed raise the same questions, then the truth of the normative pluralist thesis now at issue is not *necessary* for the problems concerned to arise. But however that may be, the most urgent matter at hand is that of finding a normative pluralist thesis that promises to be *sufficient* for the problems to arise.

If the truth of the pluralist thesis now at issue is not in fact necessary for the

20 For discussion, see Eklund, *Choosing Normative Concepts*, ch. 3.
21 See Eklund, *Choosing Normative Concepts*, 39, 54.

problems to arise, then rejecting this pluralism is of course not sufficient to avoid these problems.[22]

In the next section, I will turn to the question of what exactly the problems raised by normative pluralism are.

6. CLARKE-DOANE ON THE UPSHOT OF NORMATIVE PLURALISM

We now have a possible appropriate formulation of normative pluralism, in the sense of a formulation of normative pluralism that promises to be philosophically significant and is not problematic in the way other formulations of pluralism have been. Let us now turn to what might be a significant upshot of (consideration of) this kind of normative pluralism. Let me start by discussing what Clarke-Doane has to say.

As Clarke-Doane states in "Objectivity in Ethics and Mathematics," the supposed upshot of consideration of normative pluralism is that "even knowledge of the normative facts may fail to resolve a normative deliberation."[23] In more recent work, Clarke-Doane states the upshot in similar ways: "settling the facts, *even the normative facts,* fails to settle the questions at the center of our normative lives"; "The question of what to do remains after all of our beliefs are all settled."[24]

Clarke-Doane clearly takes this to be a significant upshot. But might not what is claimed in fact be something trivial? Many philosophers, taking care to distinguish questions about normativity from questions about motivation, would already hold that we can fail to be motivated by knowledge of normative facts: and isn't that a way in which knowledge of normative facts can by itself fail to settle normative deliberation?

Clarke-Doane himself mentions this, and notes that any motivational externalist—holding that the connection between judgment and action is merely contingent—will agree that an agent may fail to be motivated to do that which she has concluded that she ought to do.[25] He remarks: "The point is that our *deliberation as to whether* to kill the one is not yet completed even once we conclude that we *ought* to, that it is the *thing to do,* that it would be *good,* that we

22 In "Matti Eklund's *Choosing Normative Concepts,*" Kris McDaniel develops a notion of concepts being *conflicting alternatives,* and argues that what matters to the challenge I present is not primarily that the alternative concepts at issue have the same normative role but that they are conflicting alternatives in his sense.

23 Clarke-Doane, "Objectivity in Ethics and Mathematics," 104.

24 From Clarke-Doane, "Objectivity and Evaluation," 111, and *Morality and Mathematics,* 173, respectively.

25 Motivational externalism often concerns moral judgments specifically. Here we are concerned with the thinnest, all-things-considered judgments.

have *reason* to, and so on, *for any normative properties whatever.*"²⁶ Clarke-Doane thus wants to distinguish his conclusion from that familiar from the debate over externalism. The question is whether he successfully does so. What is it for deliberation to be "completed"? It is natural to hold either that deliberation is completed when a motivation to act has been formed; or that it has been completed when one has arrived at the last conclusion in one's reasoning, if that is different. But if Clarke-Doane says the former, he has not said anything that serves to distinguish his position from that of the externalist. For the externalist will then agree with what the quoted passage says: no conclusion of the kind mentioned serves to complete a normative deliberation. And if Clarke-Doane says the latter, then deliberation may well be completed in the sense at issue even given normative pluralism. Even if there are alternatives to my actual normative concepts, the last conclusion I draw in my reasoning may well simply be one employing one of my own normative concepts. Clarke-Doane needs an alternative notion of the completion of deliberation, and it is hard to see what that alternative might be.

Both in "Objectivity and Evaluation" and in *Morality and Mathematics*, Clarke-Doane prominently describes the question that has not yet been resolved in terms of Allan Gibbard's notion of *what to do*.²⁷ Conclusions regarding what ought to be done (and what ought* to be done, etc.) do not settle the question of what to do. At least without further guidance with respect to "what to do" one can also reasonably wonder why the pluralist challenge could not extend also to "what to do." Even if ϕ-ing is what to do, maybe ψ-ing is *what to do**, where "what to do*" is the counterpart of our "what to do." Raising the pluralist challenge but exempting "what to do" seems unprincipled. Clarke-Doane stresses as important the noncognitivist nature of the "what to do," for example, saying in his *Morality and Mathematics* that the attitude that resolves deliberation—however exactly to think about it—is not belief.²⁸ But how, exactly, is this supposed to help? I only see that it helps if the attitude is or directly entails an actual decision to act, for if that is what the attitude is like then there is no room for thoughts employing alternative concepts to throw a wrench into the decision process. By contrast, if the attitude is anything less than that, then even if I have the attitude that what to do is to ϕ, I can still equally think that what to do* is to ψ and that can throw a wrench into my deliberations: that just is the pluralist challenge all over again. Of course, if the relevant feature of the attitude is its decision-likeness then the question of how to distinguish the position from that of the motivational externalist remains: for the motivational externalist too

26 Clarke-Doane, "Objectivity and Evaluation," 111.

27 See Clarke-Doane, "Objectivity and Evaluation," 110, and *Morality and Mathematics*, 166.

28 Clarke-Doane, *Morality and Mathematics*, 173.

distinguishes between conclusions regarding which normative propositions are true and actual decisions regarding how to act.

In Clarke-Doane's more recent work ("Objectivity and Evaluation" and *Morality and Mathematics*), the supposed upshot of normative pluralism is that realism and objectivity are "in tension."[29] He thinks there is a clear sense in which mathematical questions, or some of them, are not "objective" given mathematical pluralism. He thinks, for example, that given such pluralism, the question of the truth of the parallel postulate in geometry has no objective answer, since "there are different geometries, each consistent if the others are, and these give different answers to the Parallel Postulate question."[30] By contrast, practical questions are objective, and their objectivity lies in their very nature. They remain objective even given normative pluralism. Clarke-Doane says,

> Practical questions are highly objective in the sense in which austere relativists say they are not. We cannot answer them by disambiguating different notions of ought. Nor can we resolve practical disputes by saying "you take goodmoral and I will take goodmoral*." Only one answer to a practical question is possible, simply because coordinated action requires that we do exactly one thing.[31]

He continues,

> And while such questions do not answer to the facts, this is part of the reason *why* their objectivity is robust. If they did answer to the facts, then their objectivity would be hostage to how plentiful the facts turned out to be.[32]

If realism were true about practical questions, then such questions would answer to the facts, and their objectivity would be hostage to how plentiful the facts are. Since the objectivity of such questions is not thus hostage, realism about practical questions is false.

The notion of objectivity is central in Clarke-Doane's discussion, but he says very little about what he takes objectivity to be. Here is one sort of thing that might be meant by objectivity: a question is *objective* exactly if it has a *mind-independently correct* answer. There are, to be sure, good questions about what mind-independent correctness is in the first place. But that is a different story,

29 The very formulation "in tension" occurs in both works. See, e.g., Clarke-Doane, "Objectivity and Evaluation," 112, and *Morality and Mathematics*, 174.

30 Clarke-Doane, "Objectivity and Evaluation," 105.

31 Clarke-Doane, "Objectivity and Evaluation," 112.

32 Clarke-Doane, "Objectivity and Evaluation," 112.

and the issues I will bring up are different from other questions that may be raised about this.

One can readily envisage attempts to problematize the connection between realism and objectivity in this sense. Realism can be argued not to entail objectivity, for realism is compatible with pluralism, and pluralism can be thought to stand in the way of objectivity, in the way Clarke-Doane seeks to illustrate in his discussion of the parallel postulate. (Though see below where I criticize Clarke-Doane's reasoning in this case.) Conversely, objectivity can be argued not to entail realism, so long as "correct" is understood broadly enough that something may be "correct" even if it is not true. For example, one can in principle hold that normative discourse is not fact-stating and normative judgments are not truth-apt, but still think that normative questions have mind-independent correct and incorrect answers—so long as one does not understand correctness to imply truth.

However, even supposing that these points about lack of entailments are correct, that does not mean that the best way to describe the upshot is in terms of there being a *tension* between such objectivity and realism. That two theses fail to entail each other obviously does not mean that they are in tension.

Now, I very much doubt that Clarke-Doane uses "objective" in the sense of mind-independent correctness. Consider again his reasoning in favor of taking practical questions to be objective. He reaches this conclusion on the basis that only one answer is possible in practical deliberation. There is nothing about how any agent, or any rational agent, must arrive at the same answer. The point is only that in any given instance of an agent being engaged in practical deliberation and the agent arrives at an answer, the agent must arrive at one univocal answer. Call this *forced uniqueness*: the only kind of answer to a genuinely practical question that one can arrive at is univocal. Now, forced uniqueness is so remote from issues of mind-independent correctness that any argument from forced uniqueness to objectivity in that sense must be seen as a howler. Clarke-Doane cannot really mean mind-independent correctness by "objectivity." This dramatizes the question of what he might mean instead. One possibility is that by "objectivity" he just means forced uniqueness.[33] It is also forced uniqueness that is suggested by the appeal to Gibbard's notion of "what to do." As noted above, the appeal to a noncognitive attitude concerning "what to do" is immune to a pluralist challenge only if the attitude is decision-like: if it is anything less there is room also to consider what to do*. But if "what to do"-judgments are decision-like then they do not leave room for such further reflection.

The mention of "coordinated action" in the passage quoted above suggests

33   In *Morality and Mathematics*, he says, "unique or, as I will say, *objective* answer" (27).

that Clarke-Doane has in mind our arriving at the same answer—or at least co-ordinated answers—to the question of what to do. But this is not a theme that is developed in any way, beyond this one reference to coordinated action.

As mentioned, Clarke-Doane's discussion is centered not on the question of the truth of normative pluralism, but on the question of the consequences of becoming convinced of normative pluralism. He makes the point that certain questions about deliberation seem open even so. That is a point about how normative deliberation appears to us, and about what it is to deliberate. It does not speak to the question of whether there are mind-independently correct ways to deliberate, other than indirectly: it could for example turn out that certain views on deliberation rule out that a question of mind-independent correctness even arises.

Clarke-Doane says that mathematical pluralism rules out that certain mathematical questions have objective answers. One can question this, for reasons rather independent of whatever exactly is meant by "objective." Suppose that mathematical pluralism is true. Then there are many different kinds of geometrical entities for a question like that over the truth of the parallel postulate to be about. Now, either someone asking this question manages to ask a determinate question (the question is about $lines_1$, not $lines_2$), or it is indeterminate whether the question concerns $lines_1$ or $lines_2$. In the first case, there is a unique, and uniquely correct, answer to the question as posed, even given mathematical pluralism. It is hard to see how objectivity is in any way challenged. Turn then to the second case. What we say about that case might depend on the correct account of indeterminacy. But a first point to make is that on plausible understandings of indeterminacy, it is incorrect to say that we have some question that admits of different answers. One view is that there are many different questions, each admitting of unique, and uniquely correct, answers, and it is indeterminate which one is being asked. Another possible view is that while there is a respectable sense in which only one question is being asked, there is a correct answer and it is "indeterminate," or maybe "neither." None of these views on indeterminacy provides support for Clarke-Doane's account. On the first view, pluralism does not stand in the way of each question having an objective answer. On the second view, there is, for all pluralism entails, a unique and uniquely correct answer to the question asked: "indeterminate."

There is a view on indeterminacy that would provide support for Clarke-Doane's claim that with regard to certain mathematical questions, mathematical pluralism rules out that there can be objective answers. This is a view on which only one question is being asked, albeit one that is in some sense indeterminate; and due to the indeterminacy in this question there are several different answers

that are in some sense correct. Such views are not unheard of. For example, Crispin Wright has defended the view that vagueness gives rise to "permissible disagreement": when a sentence is a borderline sentence due to vagueness one can permissibly judge the sentence to be true and permissibly judge the sentence to be false, even when knowing all the facts about the matter.[34] Clarke-Doane could use a view like Wright's in order to justify what he says about mathematical pluralism and the parallel postulate.

However, be that as it may, Clarke-Doane explicitly disclaims the suggestion that the reason the parallel postulate question lacks an objective answer has to do with indeterminacy. He allows that the geometrical expressions could be used with determinate meanings on a given occasion of use, and says the question posed using the expressions in this context would still lack an objective answer.[35] But that just makes it all the more elusive what his talk of objectivity might amount to.

7. REALISM AND OBJECTIVITY

In the previous section, I criticized what Clarke-Doane says about the upshot of consideration of normative pluralism for matters related to realism and objectivity. Here I will describe what I take the real upshot to be. Return to the issue of objectivity in the sense of mind-independent correctness. I outlined above how it could be argued from Clarke-Doane's perspective how it could be that the fact that statements within a given discourse are apt for mind-independent correctness does not imply that realism is true of that discourse, and how it could be that the converse does not hold either. The discussion in my *Choosing Normative Concepts*, briefly rehearsed above, in a different way illustrates how the converse may fail to hold in the case of normative discourse. Even if realism is true of normative discourse it can be that different possible communities can use different normative concepts, and different actions are, so to speak, favored by the different normative concepts used by the different communities. There is then in some sense an issue of which normative concepts to use. But that issue, if it exists, does not seem stable. Either we state it using only nonnormative concepts or we essentially employ some normative concepts. In the former case, we have changed the subject: we are interested in a normative question and not a merely descriptive one. In the latter case, we use some normative concepts or other: but if we use our concepts, we beg the question in favor of them and if we use concepts that are alternatives to ours, we beg the question in favor of those

---

34   See Wright, "The Epistemic Conception of Vagueness."

35   Clarke-Doane, *Morality and Mathematics*, 27.

other concepts. It may be that we ought to do what we ought to do; but it may equally be that we ought* to do what we ought* to do. One way to avoid this problem is to deny that the possibility can arise in the first place: to deny that there can be alternatives to our actual concepts, in the sense of concepts having the same normative roles as our concepts but different extensions. The viability of this strategy depends on two things. First, obviously, there is the question of how plausible it is that there can fail to be such alternatives. Second, there is the question of whether this is in the end sufficient. Suppose—to relate to one kind of example I discuss—that we come across a community whose normative concepts are so different from ours that there is no clear sense at all in which specific concepts of theirs are alternatives to specific concepts of ours, and the actions recommended by these concepts are not the same as the actions recommended by our concepts. The possibility of such a community raises the same questions as Alternative. It seems there is an issue of sorts regarding whether to use our concepts or the others' concepts; and as before the supposed issue threatens to be unstatable.

The considerations I present can be seen as an argument for why realism is compatible with the absence of mind-independent correctness. But one must be careful regarding what kind of mind-independent correctness is at issue. The claim is certainly not that there is anything in these considerations that shows that realism is compatible with rejecting the idea that there are mind-independently true claims about what normative propositions are true. Realism was characterized above in terms of mind-independent truth. What the argument shows, if successful, is rather something considerably more elusive: that even if realism is true of normative discourse, there is something—something about deliberation and action—we might have thought had mind-independent correctness conditions but it is unclear what this something might be, even if realism about normative discourse is true and normative sentences express propositions capable of mind-independent truth and falsity. For the structure of the problem I focus on is the following. Suppose realism about normative discourse is true. Normative pluralism can still be true. Given this normative pluralism, Alternative and similar scenarios are possible. A certain demand for mind-independent correctness regarding the normative demands that there be a mind-independently correct answer regarding a supposed further question of which concepts to use. The reason this is properly called a *further* question is that this supposed question is not immediately answered either by a statement employing our normative expressions nor by a statement employing the others' normative expressions. In my *Choosing Normative Concepts*, the label *ardent re-*

alist is used for the kind of realist who makes a demand for mind-independent correctness of the kind just described.

Whereas for Clarke-Doane it is not the truth of normative pluralism that matters for the pluralist challenge to arise—he considers instead our response to the supposed truth of normative pluralism—for me it is most definitely the truth of normative pluralism that matters. If there are not these alternative normative concepts, then there is not this further question that is crying out to be asked.

The problems in stating the upshot of my considerations are arguably similar to problems we have seen before when discussing Clarke-Doane. He speaks of there not being facts that settle normative deliberation. This is an evocative way of speaking. But, as already brought up, it raises immediate questions. No consideration even promises to show there are no facts such that recognition of them *as a matter of fact* settles normative deliberation. What is more nearly at issue is whether there are facts such that recognition of them *properly* settles normative deliberation; or such that recognition of them *ought* to settle normative deliberation. But these formulations, especially the latter, illustrate a problem regarding stating the upshot: one may think it is fairly trivial that what ought to settle normative deliberation is recognition of facts about what ought to be done. However, another community, using ought* instead, will find it equally trivial that what ought* to settle normative deliberation is recognition of facts about what ought* to be done.

However, even though the problems in stating the upshot are similar, there is a crucial difference. Clarke-Doane positively commits himself to the idea that there is something in normative deliberation—the question of what to do, as he stipulatively calls it—that remains unanswered by factual considerations. By contrast, in the context of my discussion, it is only claimed that there is a certain kind of realist, the ardent realist, who believes there is a further question there (and that this question has a mind-independently correct answer).

Suppose, for argument's sake, that considerations of either my or Clarke-Doane's kind do show that realism about normative discourse is incompatible with the relevant mind-independent correctness thesis (however this latter thesis is to be conceived of, exactly). As we may revert to putting it, for short: suppose that realism rules out mind-independent correctness. This is not yet sufficient for it to be non-misleadingly claimed that there is a *tension* between realism and mind-independent correctness. For it can be that mind-independent correctness can be ruled out given any view, realist or non-realist. And in fact, the argument against mind-independent correctness seems equally successful whether or not realism is adopted. Suppose you are a non-realist and want to affirm mind-independent correctness. You may think that whole normative discourse is not

truth-apt; you think normative judgments can be assessed for correctness and some such judgments are mind-independently correct. The pluralist argument can be raised against you too, so long as there are alternative normative concepts to use. Even if the normative judgment that what to do in situation S is to ϕ is mind-independently correct, maybe the normative judgment that what to do* in situation S is to ψ is *also* mind-independently correct.

An immediate concern has to do with whether there can really be this alternative "what to do*." Judgments about what to do are supposed to be special, in that such judgments somehow or other have a more direct connection to action than other judgments do. But first, if the non-realist can reasonably deny that there is such an alternative, the realist can in principle adopt that strategy as well. The realist can say that there is this unique "what to do," while reference-determination functions in such a way that "ϕ-ing is what to do" has a mind-independent truth-value. Second, once one starts to spell out the details regarding the "what to do," it does seem as if space opens up for possible alternatives. For example, the way Wedgwood describes his ought before action is in terms of what the thinker commits to preferring. But once the psychological notion of preference is employed, one can ask whether there are not relevant alternatives to it. To illustrate this in the most obvious way: "preference" is arguably vague, but then there are alternatives to it corresponding to different ways of precisifying it. The existence of such alternatives to the supposed "what to do" does not show that it is psychologically possible for us to use one of those alternative concepts. But mere psychological impossibility does not mean normative irrelevance. The fact that we psychologically inescapably use some concepts to guide action does not support the normative conclusion that these concepts are somehow normatively privileged over other possible normative concepts.

8. CONCLUDING REMARKS

I have had two main aims here: considering how normative pluralism, of the kind at issue, is best construed, and considering what is the upshot of normative pluralism regarding matters of realism and objectivity. In both parts of the discussion, I have used recent work by Clarke-Doane as my target. His work is important since he attempts to carefully lay out pluralism and its implications for realism and objectivity. I have criticized his discussion of what normative pluralism is, and I have separately criticized what he says about the upshot of normative pluralism. Both when it comes to how pluralism is best construed and when it comes to the upshot, I have instead concluded that the challenge presented is best construed as the exact challenge I present in *Choosing Normative*

Concepts. Of course, none of this is to provide a straightforward defense of the seriousness of the challenge I present.[36] While I of course believe the challenge is serious, the claim here is only the restricted one that insofar as pluralism presents a challenge of the kind considered, that challenge is best construed as the challenge that I have elsewhere presented.[37]

Uppsala University
matti.eklund@filosofi.uu.se

REFERENCES

Balaguer, Mark. "Moral Folkism and the Deflation of (Lots of) Normative and Metaethics." In *Abstract Objects: For and Against*, edited by José L. Falguera and Concha Martínez-Vidal, 297–313. Dordrecht: Springer, 2020.

———. *Platonism and Anti-Platonism in Mathematics*. Oxford: Oxford University Press, 1998.

Bykvist, Krister, and Jonas Olson. Review of *Choosing Normative Concepts*, by Matti Eklund. *Utilitas* 31, no. 3 (2019): 343–47.

Clarke-Doane, Justin. *Morality and Mathematics*. Oxford: Oxford University Press, forthcoming.

———. "Objectivity and Evaluation." In *Companions in Guilt Arguments in Metaethics*, edited by Christopher Cowie and Richard Rowland, 105–17. London: Routledge, 2019.

———. "Objectivity in Ethics and Mathematics." *Proceedings of the Aristotelian Society: Virtual Issue* 3 (2015): 100–105.

Eklund, Matti. *Choosing Normative Concepts*. Oxford: Oxford University Press, 2017.

———. "Reply to Bykvist and Olson." *Utilitas* 31, no. 3 (2019): 347–49.

———. "Reply to Critics." *Inquiry* (forthcoming). Published ahead of print, August 19, 2019. https://doi.org/10.1080/0020174x.2019.1656376.

Enoch, David. *Taking Morality Seriously*. Oxford: Oxford University Press, 2011.

Enoch, David, and Tristram McPherson. "What Do You Mean 'This Isn't the Question'?" *Canadian Journal of Philosophy* 47, no. 6 (April 2017): 820–40.

Field, Hartry. "Which Undecidable Mathematical Sentences Have Determinate

36 See again the discussions referred to in note 18 above.

Truth-Values?" In *Truth in Mathematics*, edited by H. G. Dales and Gianluigi Oliveri, 291–310. Oxford: Oxford University Press, 1998.

Fine, Kit. "Essence and Modality." *Philosophical Perspectives* 8 (1994): 1–16.

Hamkins, Joel D. "The Set-Theoretic Multiverse." *Review of Symbolic Logic* 5, no. 3 (2012): 416–49.

Leary, Stephanie. "Choosing Normative Properties: A Reply to Eklund's *Choosing Normative Concepts*." *Inquiry* (forthcoming). Published ahead of print, August 18, 2019. https://doi.org/10.1080/0020174x.2019.1656369.

McDaniel, Kris. "Matti Eklund's *Choosing Normative Concepts*." *Inquiry* (forthcoming). Published ahead of print, February 7, 2020. https://doi.org/10.1080/0020174x.2019.1656372.

McPherson, Tristram. "Against Quietist Normative Realism." *Philosophical Studies* 154, no. 2 (2011): 223–40.

———. "Ardent Realism without Referential Normativity." *Inquiry* (forthcoming). Published ahead of print, September 10, 2019. https://doi.org/10.1080/0020174x.2019.1656373.

Nowell-Smith, P. H. *Ethics*. London: Penguin, 1954.

Pickel, Bryan. "Naming, Saying and Structure." *Noûs* 51, no. 3 (September 2017): 594–616.

Plunkett, David. "Normative Role, Conceptual Variance, and Ardent Realism about Normativity." *Inquiry* (forthcoming). Published ahead of print, September 12, 2019. https://doi.org/10.1080/0020174x.2019.1656375.

Scanlon, T. M. *Being Realistic about Reasons*. Oxford: Oxford University Press, 2014.

Streumer, Bart. "Can We Believe the Error Theory?" *Journal of Philosophy* 110, no. 4 (April 2013): 194–212.

———. *Unbelievable Errors*. Oxford: Oxford University Press, 2017.

Wright, Crispin. "The Epistemic Conception of Vagueness." *Southern Journal of Philosophy* 33, no. s1 (Spring 1995): 133–60.

MORAL OUTRAGE PORN

C. Thi Nguyen and Bekka Williams

MOST ACADEMIC DISCUSSION about pornography has focused on the term in its, shall we say, classical use: pornography of a sexual nature. But right under the nose of the academic discussion, a secondary usage has evolved. Examples include, but are not limited to, "food porn," "closet porn," and "real estate porn." This usually refers to photographs (but also sometimes films and textual descriptions) of exquisitely prepared food, carefully arranged closets, and beautifully decorated apartments and homes. These representations are typically found and consumed via magazines and online sites dedicated to such images. Often, these images are of the kind of thing we would rarely be willing or able to get for ourselves, like photographs of meals at vastly expensive restaurants or airy Manhattan apartments far out of our economic grasp. Sometimes, these images are of things that we can obtain, but feel vaguely guilty about consuming—such as glistening, artery-clogging burgers. In other cases, they are perfectly ordinary things, but we enjoy looking at pictures of them anyway for some reason: close-up photos of a juicy steak or interior shots of other beautiful houses in our neighborhood that we could have purchased, but did not.

Perhaps this usage began as a metaphor or a joke, but it has quickly come to have a life and meaning of its own. Consider: we could introduce a new application of the term without further explanation and anybody who trafficked in modern colloquialisms would know exactly what we meant. For example, "I was up late last night looking at headphone porn," or "Have you seen that new site of high-end Japanese raw denim? Great fading porn," or "I'm feeling sad. Everybody please cover my Facebook with baking porn."[1] More importantly, we think this neologism captures something very important about the way that we sometimes relate to, and use, representations. The usage, we suggest, adapts a

---

[1] In a 2018 episode, *Saturday Night Live* opened with comic Alex Moffat (in the guise of Anderson Cooper) referencing "impeachment porn"—an utterly new usage, but one that was immediately comprehensible (season 43, episode 16, aired March 3, 2018).

part of the traditional concept of pornography—a part that is conceptually separable from sexuality. In using representations as sexual pornography, food porn, or real estate porn, we usually have no intention of engaging with the conveyed content of the representation. When we engage with pornography as such, we are not aiming to actually seek out sex with the porn star, actually go to that restaurant, or actually buy that house. Rather, we are using the representation itself for immediate gratification.

The first task of this paper, then, is to offer a conceptual analysis of this new use of "porn" in the generic sense. Our account will be, loosely, that a representation is used as *generic porn* when it is engaged with for the sake of a gratifying reaction, freed from the usual costs and consequences of engaging with the represented content. We can engage with sexual pornography without the need to find and engage with a sex partner; we can engage with food porn without worrying about the cost or health consequences; we can engage with real estate porn without having to clean and maintain all that spotless gleaming wood. Our claim is not specifically about the nature of sexual pornography, nor are we attempting to claim any new insight into that concept. Rather, we think the new generic usage has seized on a usefully exportable part of the cluster of ideas that surrounds sexual pornography, and cleaved it off. This conceptual analysis of generic porn is useful, we take it, because it draws our focus to a distinctive form of relationship that we have with certain representations.

The use of generic porn is not necessarily problematic, and many forms of such gratification are harmless. For example, C. Thi Nguyen's spouse has a particular affection for something she calls "organization porn," exemplified by the Things Organized Neatly page on Tumblr—a page full of an endless succession of images of pleasing organization, such as a thousand colored pencils arranged perfectly by shade, or a pile of oddly shaped pieces of wood stacked into a perfect square. She says that such images calm her down immediately when she is feeling overwhelmed by anxiety and the chaos of her life. We take it as a datum that her use of this site is, at least in the moral sense, unproblematic. But we think *some* specific types of generic porn are problematic—epistemically, morally, or both.

We will further demonstrate the usefulness of the concept of generic porn by using it to isolate another type of such porn, which has not yet been singled out: *moral outrage porn*. Moral outrage porn, as we understand it, is representations of moral outrage engaged with primarily for the sake of the resulting gratification, freed from the usual costs and consequences of engaging with morally outrageous content. The gratifications might include, among other things, a sense of moral superiority or smugness, the comforting sense of clarity that

arises from moral certainty, and the sheer pleasure of the feeling of outrage it-self.[2] We suspect that a significant amount of the activity on Facebook, Twitter, and other forms of social media might plausibly count as moral outrage porn, as does much of the content on many partisan news outlets. We will also argue that moral outrage porn is potentially more dangerous than other sorts of generic porn. Some kinds of porn are mechanistic—that is, they bring about their gratifications without requiring that their user engage in any sort of belief or belief-like states. Food porn, real estate porn, and many uses of sexual pornography are mechanistic in this way. But moral outrage porn is non-mechanistic; it is an essentially cognitive form of porn. One must engage in a belief, or belief-like state—a state of judging something to be morally bad, or something very much like this—in order to acquire the desired gratification. And this use, we will argue, is a bad thing, other things being equal. Let us be clear: our purpose here is not to condemn the use of moral outrage in moral and political discourse. Moral outrage is essential, when it proceeds from nuanced moral engagement, leads to moral action, and is aimed at the genuinely morally outrageous. Our goal here is to distinguish such authentic engagements with moral outrage from the use of moral outrage porn. Moral outrage porn, we will suggest, invites its users to seek simplified moral representations of the world, and to simplify their own moral beliefs in order to maximize the gratifications of outrage.

Finally, we offer a unified account of why some uses of porn seem benign while others seem deeply problematic. Using porn involves a particular form of instrumentalization of the porn itself. It may also encourage users to instrumentalize the kinds of real-world objects represented in the porn. Using porn is problematic when those sorts of instrumentalizations are problematic. When such instrumentalization is harmless, then using porn is (other things being equal) harmless. Using food porn is harmless because there is nothing wrong with instrumentalizing representations of food or food itself. On the other hand, there is something very wrong with instrumentalizing morally rich descriptions of the world.

---

2   One set of representations often described as "revenge porn" is like this. Consider, for ex-ample, the use of the term *revenge porn* in discussions of recent complaints against comic Aziz Ansari. (See, for example, Abcarian, "Is Aziz Ansari a Victim of 'Revenge Porn' or a Perpetrator?"; Flanagan, "The Humiliation of Aziz Ansari"; and Ham, "Ansari Isn't the First Victim of #MeToo Revenge Porn, but He Should Be the Last.") Please note, however, that we are not discussing another common use of the term *revenge porn*, where this describes dissemination of sexual images of a prior partner for purposes of revenge/embarrassment. (Instances of the two usages *can* overlap, but they typically do not.)

1. PORNOGRAPHY AND GENERIC PORN

Let us first establish some terminologies of convenience. We will reserve the term *pornography* for the traditional, sexual usage and use *porn* to refer to our new, generic sense.[3] Let us start by charting some uses of this newer sense of "porn." "Food porn" seems to be among the earliest usages of the generic sense of "porn," and the usage is now quite widespread. Consider, for example, the site FoodPornDaily—subtitled "click, drool, repeat."[4] The main page of the site consists of single large close-up shots of food, like juicy fruit cobblers or extra-cheesy nachos. The user stares at the picture for a while, and, at their leisure, clicks on it, which immediately refreshes the page with another randomly selected close-up shot of food. Though the use of "porn" in this case may contain a bit of self-mockery or a self-admission of guilt, the usage is not typically condemnatory, as can be gathered by the gleeful self-identification of many food porn fans as such. Similar openly acknowledged usages of porn, under that very term, can be easily found for real estate porn, closet porn, cabin porn, and various kinds of fashion porn.

Other uses are more condemnatory. For example, the term "poverty porn" has come to be used for a certain indulgent use of images and stories of poverty. Here is an example and explanation from recent journalism:

> In case you hadn't noticed, poverty is entertaining. "Poverty porn" refers to both Westerners' portrayal of global inequality, and also to the distorted presentation of disadvantage by the advantaged. Like mainstream sexual porn that produces sexualised images from the male gaze for male gratification, poverty porn produces objectifying images of the poor through a privileged gaze for privileged gratification.[5]

Similarly, pictures of urban decay have been called, in a critical mood, "ruin porn." Again, from recent journalism:

> "Ruin porn" is based purely on aesthetics and is almost always devoid of people. Employing the mismatched spoils of history, ruin porn ignores and overwrites the voices of those who still call Detroit home. When its ruins are fetishised as art, these injustices are, at best, ignored, and, at

---

3   We in no way mean to claim here that these usages perfectly track natural usages; we introduce them for the sake of writerly brevity.

4   FoodPornDaily, http://foodporndaily.com/.

5   Threadgold, "'Struggle Street' Is Poverty Porn with an Extra Dose of Class Racism."

worst, mimicked. They ignore the humanity of residents' current struggles, while replicating the history that created them.[6]

But, one might ask, why call it porn at all? The suggestions from both journalistic sources recorded above are remarkably harmonious. Ruin porn and poverty porn are exploitative presentations of impoverished people and ruined cities, used for some sort of gratification. The parallel to traditional accounts of sexual pornography is, at least in broad outlines, obvious.

Notice that the quoted discussion of "poverty porn" invokes the notion of objectification—well-familiar from feminist criticisms of pornography. The parallel between sexual pornography and generic porn is quite striking in the poverty porn case, because poverty porn also focuses on images of people. Here, the claim that they both share an objectifying gaze is plausible. But the notion of objectification, in its barest form, will not help us with developing an account of generic porn—since many of the subjects of such porn are, literally, objects. One cannot *reduce* a closet to an object because it already is an object, and any accurate representation of the closet will present it as such. Part of the usage of "porn" is to imply that it is a distinctive sort of representation. Not all pictures of real estate are real estate porn and not all pictures of closets are closet porn, though both porn and non-porn images can correctly present what they depict as objects. Thus, the brute notion of objectification will not help us isolate the distinctive quality of generic porn or explain why some forms of food photography count as food porn while others do not.

More help will come for our particular interests if we turn from the feminist discussion of pornography, with its particular focus on sex, sexuality, and human bodies, to the smaller and less traveled discussion of pornography from the aesthetics literature.[7] Here, the conversation has been one that attempts to distinguish between the concept of "pornography" and the concept of "art"—often focusing on trying to isolate the conceptual difference between an artistic nude (perhaps even an erotic one) and pornography. Much of that debate has concerned the question of whether the concepts of art and pornography are essentially incompatible. We will remain neutral on that debate for the purposes of this paper, but the conceptual territory that has been explored during the debate will be useful to us. As Anne Eaton and Hans Maes point out, artistic nudes can

---

6 Doucet and Philp, "In Detroit 'Ruin Porn' Ignores the Voices of Those Who Still Call the City Home."

7 Andrew Kania has provided a useful discussion of the different themes, framing issues, and results across the feminist discussion and the aesthetics discussion ("Concepts of Pornography").

also be problematically objectifying and misogynistic.[8] Thus, the conceptual distinction between erotic art and pornography must turn on something over and above objectification and misogyny, or at least the distinction must be more fine-grained.

One way to make the distinction, according to Maes's classificatory scheme, has been to distinguish pornography in terms of its prescribed response. George Steiner argues that pornography leaves nothing up to the imagination, while art invites the audience to enter into a jointly creative effort.[9] The thought goes something like this: art invites all kinds of complex responses. In fact, the freedom of our response is part of the value of art, whereas pornography is made for, and used for, one particular response. Pornography has a simple and mechanistic relationship to its users' desired response. From this observation, we can make a larger point. Pornography has purely instrumental value. It is there to provoke a response, and once this has been achieved, we discard it. Art, on the other hand, is intrinsically valuable. This is why, suggests Maes, we speak of *consuming* pornography and of *appreciating* art.

A recent variation on this approach pays special attention to the different kinds of relationship between a representation and its medium. With art, we care about the way that the content is presented—the technique, the use of the medium—but with pornography, we do not. This approach has many proponents and many fine-grained variations, but let us take Christy Mag Uidhir's analysis as our touchstone here.[10] Pornography, says Mag Uidhir, is valuable insofar as it achieves its purpose of sexual arousal in a "manner-inspecific" way, whereas art is valuable insofar as it achieves its purpose in a "manner-specific" way.[11] With art, we care about all the subtleties of how an artwork achieves its effects—about the delicacy of the brushwork, or the interesting framing—whereas with pornography, all we care about is the brute fact that we get that desired response. Thus, according to Mag Uidhir, even if there is erotic art whose purpose is sexual arousal, it differs from pornography in the following way: erotic art is valuable for the way in which it produces the sexual arousal—its usage of, say, photographic or painterly technique—where pornography is valuable just if it does the job. (Mag Uidhir is building from Jerry Levinson's account, according to which pornogra-

---

8   Eaton, "A Sensible Antiporn Feminism"; Maes, "Who Says Pornography Can't Be Art?" 22.

9   Steiner, "Night Words," 210.

10 For the variations, see see Davies, "Pornography, Art, and the Intended Response of the Receiver."

11 Mag Uidhir, "Why Pornography Can't Be Art."

phy simply presents its content and gets out of the way, whereas, with art, we care how that content is presented.)[12]

Perhaps here it is worth going back as far as Schopenhauer. As Alex Neill notes, Schopenhauer contrasts the much-vaunted aesthetic category of the sublime with what he calls the merely charming or attractive.[13] The charming was not a kind of content, but a style or manner for portraying contents. The charming cannot be experienced aesthetically, says Schopenhauer, because it is designed to excite desire in the beholder, and desire makes impossible the will-less, contemplative attitude that Schopenhauer takes to be essential to aesthetic experience. Strikingly, Schopenhauer picks out two subjects for his disdain—certain historical manufacturers of nudes who arranged their subjects precisely to excite lust, and certain reprehensible Dutch still life paintings of food that depicted the food in a manner that "necessarily excite[s] the appetite," and from which the spectator is "positively forced to think of [its] edibility."[14]

From both of these threads we can draw some useful lessons for our project. In general, what unifies these accounts is a sense that pornography offers something like a mechanistic or simplistically functional relationship to a representation. We use that representation to get a particular effect and we value it because it reliably gets that effect. And pornography is always contrasted with some other representative practice that supports, from certain lights, a fuller and richer relationship. Non-pornographic artistic representation, by varying accounts, leaves our imagination free, lets us be contemplative, or invites contemplation of the subtleties of the manner and form of its presentation. Roger Scruton, in fact, criticizes pornography, not for failing to be art, but for failing to work toward full personal relationships. Pornography gets us only sensations, whereas sexual desire, in its most mature and developed form, moves us toward deep interpersonal relationships.[15]

Perhaps one might wish to reject Scruton's claim that the purpose of sexual desire is always deep interpersonal relationships. Even then, a version of the point still holds. Using pornography involves a thinner and less rich interpersonal interaction than sex. Using pornography involves no interaction and no mutual responsiveness.[16] Something about this parallel is surely captured in the generic usage of "porn." Food porn merely stimulates us; it does not bring us

12   Levinson, "Erotic Art and Pornographic Pictures."

13   Neill, "The Pornographic, the Erotic, the Charming, and the Sublime," 49.

14   Schopenhauer, *The World as Will and Representation*, 207–9.

15   Scruton, "The Moral Birds and the Bees."

16   It is thus unsurprising that it is exactly this lack of mutual responsiveness in sexual cases that leads Nagel to classify such cases as "perverse" ("Sexual Perversion").

nourishment. Conversely, when my friend enthusiastically texts me pictures of the beautiful vaulted ceiling and the warm wooden hallways of the house that she has just bought, saying that, for the first time in her life, she feels at home, those pictures are most decidedly not real estate porn. Food porn, real estate porn, and ruin porn are unified in being shorn of context and consequence—in being used to satisfy some desire in a reliable, simple, and functional way. Our relationship to porn is sharply and straightforwardly instrumental. This observation will eventually help us to offer a unified account of why some sorts of porn are so troubling, and others are not—and help us to say why. But before we can do that, we need to provide a clearer account of what, exactly, porn is.

2. TOWARD A DEFINITION

We would like, now, to offer a definition of "porn," in the generic sense. We adapt our definition from Michael Rea's account of pornography.[17] We happen to think that this is a particularly good account of sexual pornography, but the reader need not share that view. What is most important is that Rea's account captures something central to what we have been discussing, and references to sexuality and obscenity are not fundamental to the account.

Rea's account is one that takes the *use of pornography* by its audience as the primary concept. He then treats "pornography" as a secondary concept, defined in terms of something's being used as pornography.

Rea defines use-as-pornography in the following way:

Part 1: *x* is *used (or treated) as pornography* by a person $S =_{\text{DF}}$ (i) *x* is a token of some sort of communicative material (picture, paragraph, phone call, performance, etc.), (ii) *S* desires to be sexually aroused or gratified by the communicative content of *x*, (iii) if *S* believes that the communicative content of *x* is intended to foster intimacy between *S* and the subject(s) of *x*, that belief is not among *S*'s reasons for attending to *x*'s content, and (iv) if *S*'s desire to be sexually aroused or gratified by the communicative content of *x* were no longer among *S*'s reasons for attending to that content, *S* would have at most a weak desire to attend to *x*'s content.[18]

To gloss this account, a piece of communicative material is used as pornography if its user's primary interest in engaging with that material is sexual arousal or gratification. Once the concept of something's being used as pornography is established, Rea defines "pornography" straightforwardly:

17   Rea, "What Is Pornography?"
18   Rea, "What Is Pornography?" 120.

Part 2: *x is pornography* $=_{DF}$ it is reasonable to believe that x will be used (or treated) as pornography by most of the audience for which it was produced.[19]

A virtue of Rea's account is that it makes sense of the relationship between pornography proper and the usage of repurposed pornographic artifacts. For example, if hackers break into a film star's computer, steal some medical pictures, and then distribute them for the purposes of sexual gratification, then those pictures are not pornography, but they will be used as pornography. Rea's condition (iii) is intended to distinguish between pornography and sexually explicit intimate communications between romantic partners. It is, we suggest, a relative of Scruton's thought that pornography is a shallower use of sexual desire than are relationships. It also allows us to finely differentiate between pornographic and non-pornographic uses of the same material. For example, suppose that one sends one's romantic partner revealing pictures of oneself. If they use those pictures for sexual arousal, but that arousal plays a part in building and furthering intimacy, then their use of the pictures is not pornographic on Rea's account. If, however, one's partner saved those photographs and used them for immediate gratification after the relationship had ended, then that use would count, on Rea's account, as *treating-as-pornography*.

We take Rea's account to be an excellent one from which to build; it will help us to crystallize the various associations we charted above. We make the following changes to generalize the account. We substitute for the specific notions of sexual arousal and sexual gratification the general notion of *gratifying reactions*. A reaction is gratifying when we take some positive enjoyment or pleasure from the reaction itself. Furthermore, Rea's condition (iii) is quite specific to the case of sexuality and its role in human relationships. To generalize the underlying idea, what seems to unite all the cases of porn is some disengagement from the usual complexities, entanglements, difficulties, and responsibilities of the represented content.

Finally, we shift from the notion of "communicative material" to "representations" because the notion of a representation is thinner than the notion of communicative material and will capture an appropriately wide array of cases. For example, if I take a picture with my phone to remember a delicious piece of cake for myself, that picture is clearly a representation, but not clearly a piece of communicative material. I can surely use that photo as food porn. Similarly, suppose I bump my phone and accidentally send a picture of the exquisite meal

---

19 Rea, "What Is Pornography?" 120.

I am enjoying to a friend. That is certainly a representation, but not obviously a communication, and it also could be used as food porn.

Thus, following Rea, we define treating-as-porn in the following way:

Treating a representation as Ω-porn $=_{DF}$ using the content of a representation of Ω primarily for the purpose of generating one's own gratifying reactions, freed from the typically attendant consequences and effort of engaging with Ω.

We include "and effort" because there are cases of Ω-porn where the effectiveness of the porn hinges at least in part on avoiding aspects of the process of engaging with Ω, rather than avoiding the consequences of engaging with Ω.

We then define generic porn in the following way:

x is Ω-porn $=_{DF}$ *x* is a representation where it is reasonable to believe that *x* will primarily be used (or treated) as Ω-porn by most of the audience for which it was produced or transmitted.

Our account adds to Rea's the demand that *x* will *primarily* be used as Ω-porn. This strikes us as a necessary amendment to capture the complexities of the generic use of the term "porn." For example, Cezanne's still-life paintings of fruit might primarily be interacted with as art, and secondarily inspire a gratifying appetitive response shorn from the entanglements of actually eating. It would be strange to call those paintings porn because their status and function as art is dominant. Note, however, that it is still open for particular audience members to primarily *use* those paintings as porn, without the paintings actually being porn.

Notice that we have added the notion of "transmission" to Rea's notion of production. This, we think, captures an additional possibility that has become particularly salient in the current socio-technological media context. In many cases, content aggregators have brought together representations that were not originally produced for the sake of being used as porn, but that have been aggregated or retransmitted for the sake of usage as porn. For example, some blogs collect particularly luscious real estate photos from sites that posted those photos for the sake of sale, but where the blogs have collected the most porn-worthy shots.

Finally, our account of *treating-as-generic-porn* drops an equivalent to Rea's condition (iv)—the requirement that if a user's sexual interest lapsed, the user would have little desire to continue to attend to the content. Even if (iv) were plausible in cases of sexual pornography, it is highly implausible in important cases of generic porn. As Stephanie Patridge notes, we can have multiple rea-

sons for attending to a representation.[20] Enjoyment of sexual representations may sometimes overlap with other motivations for attending to these representations, and individuals may still have non-pornographic reasons for attending to a representation in the (counterfactual) absence of the pornographic reasons. For example, suppose that we enjoy impeachment porn in large part because of the moral importance we attach to problems with the current administration. The very importance we attach to these problems likely grounds a reason for us to attend to newsfeeds reporting new problems for the administration, independent of the gratifying reactions that these reports facilitate. (We suspect that this grounding relation is especially common in cases of moral outrage porn, which we discuss below.) Notice, however, that even if a piece of news could be used as porn, it still might not count as porn, if it were not reasonable to believe that it would be so used by most of its intended audience.

After all these complexities, let us also offer a slightly more portable, albeit less exact, version of our account:

> *Portable version*: Ω-porn is representations of Ω used for immediate gratification, while avoiding the usual costs and consequences of actually engaging with Ω.[21]

A nice upshot of our account is that Rea's notion of pornography turns out to be a special case of generic porn, and it satisfies our definition of pornography in a way that highlights the parallels between sexual pornography and generic porn. On our view, sexual pornography is a representation of sexual content, primarily used for the purpose of a gratifying reaction, freed from the usual effort and consequences of sexual interaction.[22] Food porn is a representation of food, primarily used for the purpose of generating gratifying reactions like pleasurable hunger or culinary excitement, freed from the various efforts of making food or going to a restaurant, and consequences regarding price and nutrition. (This explains why food porn is so often a representation of very unhealthy, very expensive, and/or very difficult-to-prepare food.) Real estate porn is a representation

---

20 Patridge, "Exclusivism and Evaluation," 50.

21 Note that this version elides some of the linguistic complexity of the full definition of *porn*. This is meant only as a writerly convenience; we mean throughout to be invoking our full definition of *porn*.

22 Note that our definition demands only that the use-as-porn is in fact (or expects to be) freed from the usual effort and consequences, not that the user of porn would prefer to be so freed. Somebody using sexual pornography might have preferred to actually have sex and be entangled with the costs and consequences, but still counts as using-as-porn because, in fact, they have not been so entangled and have no expectation of being so entangled in connection with using the pornography.

of real estate, primarily used for the purpose of generating gratifying reactions, freed from the attendant efforts and consequences of purchasing and maintaining real estate. (This explains why real estate porn is so often of a representation not only of expensive real estate, but of hard-to-maintain surfaces and materials, like all-white upholstery.) Poverty porn is representations of impoverished people, primarily used for the purpose of various gratifying reactions—like the pleasures of sensations of sentiment and empathy—while freed from the attendant effort and consequences of actually having to morally engage with and face the prospect of relieving such poverty. (This explains why pictures of impoverished people published on charitable donation websites are likely not poverty porn, but those same pictures repurposed in a certain sort of travel magazine very likely might be.)

3. MORAL OUTRAGE PORN

We have identified what we take to be the concept underlying this new colloquial usage. And this new colloquial usage, we think, has come about for a very good reason. It picks out a clear, useful, and morally and practically relevant category. To demonstrate the account's conceptual fecundity, we will use it to identify a novel form of porn—one that, we hope, will help us to make sense of the world.

We suggest a new category of porn: *moral outrage porn*. Moral outrage porn is representations of moral outrage primarily used for the sake of the resulting gratification, where the user engages with the representation freed from the usual consequences and efforts of engaging with morally outrageous content. The term "representations of moral outrage" deserves some clarification. There are several forms. One is the representation of expressions of moral outrage—such as angry tweets expressing moral outrage at some event. Another is the representation of states of affairs *as* morally outrageous—such as a morally charged, condemnatory description of a political event. We use such representations as moral outrage porn when we engage with them primarily for the sake of a gratifying reaction, freed from the usual consequences and efforts. Such representations are moral outrage porn when it is reasonable to think that they will primarily be put to such uses. Most importantly, using such representations as moral outrage porn often involves engaging with them without applying the epistemic standards of veracity or worrying about the consequences of entertaining such representations.

One might now reasonably begin to suspect that a significant amount of the content of social media is, at present, moral outrage porn. A paradigmatic example of moral outrage porn is the content of certain politically partisan news sites

and social media feeds, which continually present inflammatory articles target-
ing some political enemy. Recall that one of the suggested accounts of sexual
pornography was that it is used for its capacity to reliably provoke some specific
effect—as opposed to art, which leaves the audience free to respond in a vari-
ety of ways and is sought out for that very capacity. Levinson suggests, in fact,
that the very complexities and subtleties that make for artful presentation get in
the way of the simple, reliable, mechanistic provocation of sexual responses of
pornography. When we want porn, we do not want artfulness. We just want for
pornography to do its job and then get out of the way.[23] We can find a similar
phenomenon with moral outrage porn. When one interacts with the news in a
non-porn way, one looks for the facts, in their full complexity, and engages with
them as purportedly true statements about the world. This sort of interaction
leaves the door open for one to encounter unsettling ideas—unexpected evi-
dence, challenging considerations. On the other hand, when one uses the news
as moral outrage porn, one comes to it hoping, consciously or subconsciously,
that it will reliably provide the gratifications associated with moral outrage.[24]
This would likely lead one to be largely disinterested in complex presentations
of morally ambiguous situations, because that would interfere with getting reli-
ably to that gratifying moral outrage. Such a user of moral outrage porn would,
then, seek out reliable sources of simplistic and predictable moral descriptions
of the world.[25] They would avoid morally rich representations, which might lead
to complex and unpredictable emotional and intellectual responses.

What makes something moral outrage porn then is the interest in using mo-
rality for gratification on the part of its users. A news item may indeed invite
moral outrage, but insofar as we reach that moral outrage through a full-blood-
ed and nuanced moral engagement—insofar as we are seeking moral truth and
not using our morality reactions for pleasure—then we are not using it as moral
outrage porn.

Notice, too, that our definition of "porn" includes things that are *transmitted*,
and not necessarily produced, for the sake of treatment as porn. Think about

---

23 Levinson, "Erotic Art and Pornographic Pictures," 232–33.

24 See empirical work by Green et al. suggesting a tendency for individuals to want to read
    moral-outrage-inducing articles *in order to feel good about themselves* ("Self-Enhancement,
    Righteous Anger, and Moral Grandiosity"). Also see Pizarro and Baumeister, "Superhero
    Comics as Moral Pornography," for a discussion of the gratification of moral judgment and
    categorization; as well as Rothschild and Keefer, "A Cleansing Fire," for evidence suggesting
    that expressing moral outrage can be effective in reducing feelings of personal guilt.

25 For an excellent discussion of superhero comics as a specific example of such moral simpli-
    fication for enjoyment purposes, see Pizarro and Baumeister, "Superhero Comics as Moral
    Pornography," 29–31.

moral outrage porn newsfeeds, which pick stories from more information-ly and morally rich resources. For example, a highly partisan newsfeed could simply repost news from more neutral news sites, but cherry-pick the most in-flammatory stories. That cherry-picked feed creates moral outrage porn through contextual aggregation. It might be that no particular news item on such a biased feed is moral outrage porn on its own, but the filtered, preselected assemblage may compositely be moral outrage porn.[26]

4. DANGERS OF MORAL OUTRAGE PORN

Though porn, in the generic sense, is not necessarily dangerous or morally prob-lematic, we think that moral outrage porn, in particular, is problematic. First, the use of moral outrage porn is a form of epistemic bad faith. If we adopt beliefs simply or primarily for the sake of their pleasurable resulting outrage, rather than for their putative veracity, then this is clearly a violation of any number of plausible epistemic standards. Second, in many cases when we indulge in using moral outrage porn, we are not adopting any new beliefs, but are instead using the moral values that we already have for the sake of generating personal gratifi-cation. We will argue that this, too, is a highly suspect and problematic practice.

First, we propose a distinction between two sorts of porn: mechanistic porn and cognitive porn. In *mechanistic porn*, one achieves gratification from experi-encing a representation without needing to take on any specific beliefs regarding the representation's truthfulness or accuracy. We take food porn to be an almost entirely mechanistic form of porn; much sexual pornography is similarly mech-anistic.[27] In order to be gratified by *cognitive porn*, however, one must seriously entertain the legitimacy of some belief relevant to the represented content. In the case of moral outrage porn, one must usually seriously entertain some moral belief arising from the represented content, or seriously entertain some moral belief that applies to the represented content.[28]

---

26 This raises the possibility of algorithmically generated moral outrage porn, in which auto-matic filtering effects from technological agents like Google Search algorithmically generate resources which it is reasonable to believe that its audience will use as porn. See Pariser, *The Filter Bubble*; Miller and Record, "Justified Belief in a Digital Age"; Watson, "Filter Bubbles and the Public Use of Reason"; and Nguyen, "Echo Chambers and Epistemic Bubbles." Fur-thermore, if it is not already clear, we think that there is moral outrage porn aplenty across the political spectrum.

27 We accept, however, that at least some sexual pornography is (at least to a significant degree) cognitive.

28 Consider, also, the recent coinage of the term "justice porn," which is representations of wrongdoers getting their immediate, and often violent, comeuppance. Justice porn is almost

Let us start with a paradigmatic case of moral outrage porn: a website whose readership is entirely of political party *A* posts detestable quotations from political leaders of political party *B*, where that readership is primarily interested in using those quotations for the sake of pleasurably stoking their moral outrage. In the usual case, there are two components: (i) a particular judgment concerning the represented incident, which is grounded in (ii) some moral value that readers accept. Intuitively, there seems to be something quite wrong with the use of moral outrage porn in this case. It seems like we are not so much inhabiting our moral beliefs as we are taking advantage of them for self-gratification.

But what, really, is the problem with that? It cannot just be that one is not taking one's moral beliefs sufficiently seriously. In fact, one typically needs to have a certain degree of moral belief and commitment in order to attain the desired gratifications. Holding one's moral beliefs with unswerving seriousness and commitment can actually make it easier to achieve the various gratifications of emotional security, clarity, comfort, and superiority.

Let us start by considering the epistemic problems involved with using moral outrage porn. There are two possibilities here. First, the user of moral outrage porn could be adopting a new moral view in order to be gratified. Second, the user could be gratified by the exercise of a moral view they already accept. Either case is problematic.

Suppose one adopts a new moral view simply in order to be gratified. The problem here is clear: one is adopting a belief or set of beliefs for non-epistemic reasons.[29] The reasons of gratification urge one toward a different set of moral beliefs than the reasons of moral veracity. Imagine, for the moment, that one set out to develop a set of moral beliefs so as to maximize one's possibility for pleasing moral outrage. The moral system one would develop would likely be clear, strident, and demanding, so as to maximize the incidence of moral outrage. It would likely admit of few ambiguities or difficulties, for the purpose of maximizing the possibility of clear, undiluted outrage and its associated pleasures. This sort of procedure obviously violates any number of epistemic norms and

---

certainly another example of cognitive porn. In order to enjoy the representation of (say) a just punishment as just, it is plausible that one must accept, at least while enjoying the representation, that the retribution was justified. Thus, insofar as a representation is *used as* justice porn, the user must have, or at least entertain, a relevant belief. It may be that justice porn, in the sense indicated here, is actually a subset of moral outrage porn. We need not take a position on this issue, but if justice porn is simply an example of moral outrage porn, this fits nicely with the plausibility of the view that both moral outrage porn and justice porn are prime examples of cognitive porn.

29 Regarding using the "wrong kinds of reasons," see Hieronymi, "The Wrong Kind of Reason."

virtues.[30] And notice that this sort of unambiguous, strident morality does seem to be precisely the kind of moral system we often find in communities given to the endemic usage of moral outrage porn.

What, then, of those moral outrages that depend on one's antecedently accepted moral views? We already hold certain moral beliefs, and it is so tempting to indulge in those websites that hold up the enemies of our antecedent beliefs for outrage and disdain. There are two major problems with this behavior pattern. First, in most cases, independent confirmation is a good reason to increase our degree of belief. When we systematically seek out moral outrage porn, we are getting extra confirmations of our moral worldview. That confirmation works in several ways. One way relies on the fact that moral outrage porn can itself contain moral content. That is, much moral outrage porn does not simply represent some state of affairs neutrally, to which we apply our own moral beliefs—it instead presents the world already colored by moral judgment. Thus, it presents itself as a form of support for our moral beliefs. But when one is engaged with moral outrage porn, one is seeking out representations of moral outrage for the sake of the resulting gratification, and so one is incentivized to preselect those representations with which one agrees. This invites a problematic form of circularity—where one picks one's sources based on agreement with one's antecedent beliefs, and then goes on to use those sources to buttress one's antecedent beliefs.[31] Thus, the moral outrage porn user is tempted, through the logic of self-gratification, into epistemic relationships that can increase their degree of moral self-confidence without adequate epistemic justification.

Second, moral outrage porn may misrepresent empirical facts about the world in order to provoke more gratifying moral outrage. This may happen by simply presenting false "facts," or by cherry-picking outraging facts. For example, a moral outrage porn news site could easily operate simply by selectively picking out the most damning and awful single sentences said by members of some opponent political party. Such cherry-picking offers an easier and more reliable pathway to the gratifications of moral outrage than would a more complete presentation of the relevant facts. Thus, such a site would misrepresent the

30 We are assuming, for the sake of this paper, the kinds of epistemic norms that demand that beliefs arise from the evidence, as guided by the aims of truth (Hieronymi, "The Wrong Kind of Reason"). We suspect that most pragmatist accounts of epistemic norms would also yield norms forbidding this sort of bad faith belief, on long-term pragmatic grounds, but that issue is beyond the scope of this paper.

31 For a discussion of how moral agreement and disagreement are epistemically relevant to moral beliefs, please see Nguyen, "Autonomy, Understanding, and Moral Disagreement." For further discussion about the social circularity described here, see Nguyen, "Cognitive Islands and Runaway Echo Chambers" and "Echo Chambers and Epistemic Bubbles."

opponent party as a means of gratifying its audience. And the user of moral out-
rage porn, insofar as they seek that particular form of gratification, should prefer
such poorer-quality sources of information to more accurate, but less gratifying,
sources.

So far, we have been discussing nonfictional cases, which are the clearest
cases of problematic moral outrage porn. They represent purportedly real sit-
uations, and consequently engage one's genuine moral beliefs and/or morally
relevant empirical beliefs. But what about fictional cases? Surely we could have
representations of moral outrage in fiction, and surely we could engage with fic-
tion specifically for the gratifications of such representations. What could possi-
bly be wrong with that?

First, it is important to note that we may not be able to actually adopt fic-
tional moral attitudes. As the recent literature on imaginative resistance tells us,
there seems to be a striking difference between moral beliefs and other beliefs
in fiction.[32] We can easily imagine ourselves into a fictional world in which ships
travel faster than light or the Nazis won World War II. But we cannot easily imag-
ine ourselves into a fictional world where the Nazis' anti-Semitism was morally
correct or where killing innocents for sport is morally praiseworthy. This means
that fictional moral outrage porn will have the following character: it will pres-
ent us with fictional situations with which we engage on the basis of our actual
moral beliefs. Again, one might ask, what's wrong with that? After all, we will not
be changing our consumption of facts in the world, so we will not be engaged in
epistemic bad faith.

The worry about strategically shifting one's moral beliefs to maximize outrage,
however, remains. We engage with morally outraging fiction as non-porn when
we are appropriately entangled with its moral content—when we ask ourselves
if its moral vision is true, and, if it is, try to integrate that vision into our belief
system. We engage with morally outraging fiction as porn when we take gratifi-
cation from our reaction of moral outrage, while avoiding the further entangle-
ments of applying and integrating our larger epistemic and moral beliefs with an
eye toward the truth.[33] Furthermore, if a certain moral system is desirable, not
because it is accurate but because it maximizes the pleasures of moral outrage,
then that desirability will show up both in the fictional and the nonfictional cas-

---

32 Gendler, "The Puzzle of Imaginative Resistance."

33 We are not assuming here that the only content of fiction is moral content. For example, it
     seems to us that one can engage with fiction's artistic content in a way that is richly entan-
     gled with aesthetic and artistic engagement, or one can avoid such engagement, as Mag
     Uidhir and Levinson suggest.

es. When this happens, the use of fictional moral outrage porn can incentivize us to acquire moral beliefs for the wrong reasons.

5. THE MORAL STATUS OF MORAL OUTRAGE PORN

We have argued that the use of moral outrage porn is often epistemically problematic. It seems to us, however, that using moral outrage porn might also be morally problematic. In this section, we will sketch some possible frameworks for thinking about the moral problems of moral outrage porn. While we do not take ourselves to be presenting complete arguments here, we find these lines of inquiry plausible and worthy of further development.

First, reconsider the epistemic connection. As we noted, in the vast majority of cases of the use of moral outrage porn, an individual will seek out representations that are pleasurable because of that individual's antecedent moral views, thereby acquiring what is easily mistaken for further confirmation of these views. This is certainly problematic in the epistemic sense. But epistemic carelessness is not, in itself, always morally problematic. (For example, suppose that we form our belief about whether tomatoes are fruits or vegetables by consulting our Magic 8 Ball. This is bad reasoning, epistemically speaking, but not *morally* wrong.)

It is plausible, however, that epistemic carelessness specifically in the moral realm *is* morally problematic. Insofar as one's moral convictions are likely to inform how one behaves in morally relevant settings, epistemic carelessness in the moral realm exposes one to acting in morally wrongful ways. Increasing the likelihood that one will act wrongfully is clearly morally problematic for consequentialist reasons, and epistemic carelessness in the moral realm may also be morally problematic simply in virtue of expressing a lack of respect for moral reasoning.[34] Thus, because the use of moral outrage porn seriously risks violations of acceptable morally relevant belief formation, it is morally problematic.

Additionally, the use of moral outrage porn faces another, distinctively moral problem—one that is entirely independent of epistemic considerations. In discussions of moral outrage porn, we have noticed a common attitude, vaguely expressed as the concern that using moral outrage porn *misuses* morality—that such use "isn't what morality *is for*." This intuitive sense of wrongness, we think,

---

34 For our current purposes, it is enough that there is at least one serious moral strike against such epistemic carelessness. In an in-progress paper, however, Bekka Williams argues that nontrivial epistemic carelessness in the moral realm is morally wrong regardless of its consequences.

points the way to the distinctively moral problem. The use of moral outrage porn cheapens and undermines the role of moral experience.

This worry parallels, in some significant ways, Tosi and Warmke's complaint against moral grandstanding.[35] Moral grandstanding is morally problematic, they argue, in large part because grandstanders are treating moral discourse as a "vanity project":

> In using public moral discourse to promote an image of themselves to others, grandstanders turn their contributions to moral discourse into a vanity project. Consider the incongruity between, say, the moral gravity of a world-historic injustice, on the one hand, and a group of acquaintances competing for the position of being most morally offended by it, on the other.
>
> Such behavior, we think, is not the sort of thing we should expect from a virtuous person.[36]

Note, crucially, that the problem asserted by Tosi and Warmke in this instance is not that moral grandstanding has bad results.[37] Instead, the problem is that using moral discourse for self-promotion is problematically egotistical.

Tosi and Warmke focus on moral problems associated with using moral outrage for interpersonal jockeying. That is the essence of the notion of moral grandstanding—the use of moral expression for social signaling. Similarly, it seems plausible that the use of moral outrage porn in many cases involves a failure to respect the fundamental role of moral expression. Notice that, where the problem with moral grandstanding is essentially interpersonal and social, the problem with moral outrage porn is personal and hedonistic.[38] The problem of moral grandstanding is that we use morality for status; the problem of moral outrage porn is that we are using morality for pleasure. When one indulges in moral outrage porn, one uses what *by one's own lights* is morally outrageous for

---

35 As described by Tosi and Warmke, moral grandstanding involves expressing a moral view, attitude, etc., where the expression is significantly motivated by a desire to be recognized as "morally respectable" ("**Moral Grandstanding**," 200, 202).

36 Tosi and Warmke, "Moral Grandstanding," 215–16.

37 Although they also claim that moral grandstanding often has bad results.

38 For empirical evidence that experience of moral outrage serves to reduce the experience of individual guilt, see Rothschild and Keefer, "A Cleansing Fire"; and Rothschild et al., "A Dual-Motive Model of Scapegoating." For empirical evidence that individuals sometimes engage with moral-outrage-inducing articles *in order to* bolster their self-perceptions as morally virtuous (as "paragons of morality"), see Green et al., "Self-Enhancement, Righteous Anger, and Moral Grandiosity."

one's own enjoyment.[39] It is, loosely speaking, to make morality about oneself, when it clearly is not.

Furthermore, it is no accident, we think, that the features of moral outrage porn relevant to the "bad faith" problem mirror Michael Tanner's account of the problems of sentimentality.[40] In his discussion of Oscar Wilde and the sentimental, Tanner says, "the feelings which constitute [the sentimental] are in some important way unearned, being had on the cheap, come by too easily."[41] The use of moral outrage porn, if one accepts our definition, involves an attempt to be gratified by a representation of the end result of moral engagement without taking on the consequences or effort of actually engaging. This seems a paradigmatic case of getting a feeling on the cheap.

What we have sketched thus far are a number of considerations that weigh in favor of a serious moral strike against the use of moral outrage porn. There are also a number of consequentialist considerations that we might adduce. Tanner argues that the intrinsically sentimental tends toward passivity.[42] Sentimental emotions, Tanner suggests, can themselves *encourage* inaction.

> It also seems to me that some of my feelings are of a kind that inhibit action, because they themselves are enjoyable to have, but if acted upon, one would cease to have them, and one doesn't want to. Such a feeling does seem to me intrinsically sentimental.[43]

Just as sexual pornography can—although certainly need not—problematically replace real interpersonal sexual interaction, moral outrage porn runs the risk of mollifying its users into inaction.[44] Along similarly problematic lines, use of moral outrage porn could have the effect that Tosi and Warmke term "outrage exhaustion"—that is, those who regularly use moral outrage porn may, as Tosi and Warmke suggest, "find it increasingly difficult to muster outrage when it actually is appropriate."[45] But these final suggestions would require significant empirical investigation to substantiate.

---

39 See especially our above discussion of the difficulty (if not impossibility) of enjoying an instance of moral outrage porn without at least entertaining a relevant moral belief.

40 Tanner, "Sentimentality."

41 Tanner, "Sentimentality," 128.

42 Tanner, "Sentimentality," 134.

43 Tanner, "Sentimentality," 139.

44 See especially empirical findings by Rothschild and Keefer ("A Cleansing Fire") suggesting that the expression of moral outrage tends to decrease the experience of personal guilt, which thus tends to decrease motivation to act to remedy injustice.

45 Tosi and Warmke, "Moral Grandstanding," 211.

6. PORN AS INSTRUMENTALIZING

This analysis of moral outrage porn suggests a larger account of the moral status of porn in general. Our comments here will necessarily be a bit preliminary, since we are grappling with a large and diverse array of phenomena.

As we noted earlier, users are looking for pornography to provoke a mechanistic and simplistic response in them—unlike the complex, unpredictable, and rich responses we seek with art. That mechanistic attitude reveals something deeper: an instrumental attitude toward the pornography. This instrumentalization seems problematic in two ways. Insofar as art is intrinsically valuable, then pornography cannot be art, because our attitude toward it is strictly instrumental. And insofar as the frequent use of pornography encourages a strictly instrumental attitude toward actual people and sexual encounters, then it helps to undermine our capacity to treat humans and intimate human relationships with the dignity they deserve.

Moral outrage porn also involves various forms of problematic instrumentalization. First, with moral outrage porn, we are instrumentalizing the porn's content—though not through the kind of mechanistic pathway of sexual pornography. Insofar as I am interacting with the news as moral outrage porn, I am not looking to be informed by the facts, but am using the news for the sake of my own gratification. So long as we are supposed to be responsive to the genuine facts of the matter, using the news as moral outrage porn—and manipulating which facts I am exposed to, for the sake of gratification—is a violation of epistemic norms. Furthermore, the regular use of moral outrage porn encourages a further form of instrumentalization—one in which we instrumentalize our own moral beliefs. Moral outrage porn gives me an incentive to modify my own moral beliefs: I can be tempted to modify my beliefs to make it easier to access the gratifications of moral outrage. But that modification is an abuse of moral belief—so long as we think that moral beliefs are supposed to track something like moral truth.

This suggests a general account of the moral status of porn. Using porn involves making an instrumental use of a representation. Furthermore, using porn often encourages further instrumentalizations of the represented content.[46] Finally, using porn often encourages even more downstream instrumentalizations of various background beliefs and attitudes. For example: using sexual pornog-

---

46 Consider especially Green et al.'s evidence that those who have been manipulated to experience moral outrage ("righteous anger") had an increased likelihood of desiring to engage with moral-outrage-inducing articles ("Self-Enhancement, Righteous Anger, and Moral Grandiosity").

raphy certainly instrumentalizes representations of sex and bodies, may encourage the instrumentalization of sex and bodies themselves, and may encourage the instrumentalization of downstream associates, such as relationships and personalities.[47] Using the news as moral outrage porn certainly instrumentalizes the news, likely instrumentalizes the events in the world, and furthermore may encourage the instrumentalization of our moral beliefs and moral communities.

This helps us explain the variation in our attitudes toward the moral status of different kinds of porn. Other things being equal, the use of porn is morally problematic when such instrumentalizations run afoul of other norms (some of which we have discussed), and unproblematic when such instrumentalizations are themselves unproblematic. This offers a useful explanation for our asymmetric intuitions about the various forms of porn. Using food porn and real estate porn seems unproblematic. Our account suggests that this is because there is nothing wrong with instrumentalizing food or real estate or their representations. Sexual pornography and moral outrage porn strike many as problematic. Our account suggests an explanation: they are problematic insofar as they encourage the instrumentalization of things that ought not to be instrumentalized.

To be absolutely clear: our account is in no way intended to be a criticism of moral outrage, or some sort of general call for politeness and civility. Moral outrage can be crucial to proper moral action and the quest for social justice. Our worry is, in fact, something of the opposite: our worry is that using moral outrage porn can dilute genuine moral outrage or lead it astray. It is precisely because moral outrage is so important that we must not instrumentalize it. We certainly should not retune our sense of outrage for our own pleasure.

It might be useful here to note that there can certainly also be "civility porn": calls for civility and politeness, used for the sake of pleasurable feelings of, say, smugness and superiority at one's own maturity and high-mindedness.[48] The problem is not with moral outrage or civility themselves; it is with instrumentalizing the representations of either. Both moral outrage and civility are vital, which is exactly why moral outrage porn and civility porn are so potentially undermining.

7. CONCLUSION

We think that our definition captures, to a significant degree of accuracy, the natural usage of the term *porn* in the generic sense. It describes the regular way in

---

47   For a sophisticated discussion of the potential harms of sexual pornography, see Eaton, "A Sensible Antiporn Feminism."

48   We owe this idea to a suggestion by Aaron Rabinowitz.

which the term has come to be used and understood in recent discussions, and it naturally categorizes sexual pornography as an instance of generic porn—which is both plausible on its face and helpful in explaining the genesis of the generic usage. Furthermore, our account of generic porn is useful in that it highlights and provides a framework for discussing a widespread and rapidly increasing phenomenon: moral outrage porn. The use of moral outrage porn, we have argued, is epistemically problematic. It may also be morally problematic. It is plausibly objectionable in the way that cheapens and undermines moral experience. Moral outrage porn invites us to instrumentalize something that ought not to be instrumentalized.

We have tried to unpack the conceptual insight contained in a recent, naturally evolved sort of neologism. The term "porn" once had a clear and restricted meaning, concerning certain sexual representations. That term has recently gained a new, secondary colloquial usage—the generic usage of "porn." That usage may have started as a metaphor or a joke. But the reason that this secondary usage has caught on so well is because it identifies a common thread in our usage of representations—that sometimes we use representations for self-gratification, freed from the usually attendant consequences and worries that might accompany actual interactions with what was represented. That kind of usage is starkly obvious in the sexual cases, but thinking seriously about the use of sexual pornography, as we have argued, provides a very useful framework for understanding the new and widespread use of the term "porn" in the generic sense.[49]

University of Utah
thi.nguyen@gmail.com

Minnesota State University, Mankato
bekka.williams@mnsu.edu

REFERENCES

Abcarian, Robin. "Is Aziz Ansari a Victim of 'Revenge Porn' or a Perpetrator?" *Chicago Tribune,* January 17, 2018. http://www.chicagotribune.com/news/opinion/commentary/ct-perspec-victimhood-aziz-ansari-revenge-sexual-aggression-weinstein-lauer-charlie-rose-0118-story.html.

Davies, David. "Pornography, Art, and the Intended Response of the Receiver." In Maes and Levinson, *Art and Pornography*, 61–82.

Doucet, Brian, and Drew Philp. "In Detroit 'Ruin Porn' Ignores the Voices of Those Who Still Call the City Home." *The Guardian*, February 15, 2016. https://www.theguardian.com/housing-network/2016/feb/15/ruin-porn-detroit-photography-city-homes.

Eaton, A. W. "A Sensible Antiporn Feminism." *Ethics* 117, no. 4 (July 2007): 674–715.

Flanagan, Caitlin. "The Humiliation of Aziz Ansari." *The Atlantic,* January 14, 2018. https://www.theatlantic.com/entertainment/archive/2018/01/the-humiliation-of-aziz-ansari/550541/.

Gendler, Tamar Szabó. "The Puzzle of Imaginative Resistance." *Journal of Philosophy* 97, no. 2 (February 2000): 55–81.

Green, Jeffrey D., Constantine Sedikides, Daryl R. Van Tongeren, Anna M. C. Behler, and Jessica M. Barber. "Self-Enhancement, Righteous Anger, and Moral Grandiosity." *Self and Identity* 18, no. 2 (December 2017): 201–16.

Ham, Mary Katharine. "Ansari Isn't the First Victim of #MeToo Revenge Porn, but He Should Be the Last." *The Federalist,* January 17, 2018. http://thefederalist.com/2018/01/17/aziz-ansari-isnt-the-first-victim-of-metoo-revenge-porn/.

Hieronymi, Pamela. "The Wrong Kind of Reason." *Journal of Philosophy* 102, no. 9 (September 2005): 437–57.

Kania, Andrew. "Concepts of Pornography: Aesthetics, Feminism, and Methodology." In Maes and Levinson, *Art and Pornography*, 254–76.

Levinson, Jerrold. "Erotic Art and Pornographic Pictures." *Philosophy and Literature* 29, no. 1 (April 2005): 228–40.

Maes, Hans. "Who Says Pornography Can't Be Art?" In Maes and Levinson, *Art and Pornography*, 17–47.

Maes, Hans, and Jerrold Levinson, eds. *Art and Pornography: Philosophical Essays*. Oxford: Oxford University Press, 2012

Mag Uidhir, Christy. "Why Pornography Can't Be Art." *Philosophy and Literature* 33, no. 1 (April 2009): 193–203.

Miller, Boaz, and Isaac Record. "Justified Belief in a Digital Age: On the Epis-

temic Implications of Secret Internet Technologies." *Episteme* 10, no. 2 (June 2013): 117–34.

Nagel, Thomas. "Sexual Perversion." *Journal of Philosophy* 66, no. 1 (January 1969): 5–17.

Neill, Alex. "The Pornographic, the Erotic, the Charming, and the Sublime." In Maes and Levinson, *Art and Pornography*, 48–60.

Nguyen, C. Thi. "Autonomy, Understanding, and Moral Disagreement." *Philosophical Topics* 38, no. 2 (Fall 2010): 111–29.

———. "Cognitive Islands and Runaway Echo Chambers: Problems for Epistemic Dependence on Experts." *Synthese* 197 (2020): 2803–21.

———. "Echo Chambers and Epistemic Bubbles." *Episteme* 17, no. 2 (June 2020): 141–61.

Pariser, Eli. *The Filter Bubble: What the Internet Is Hiding from You.* London: Penguin Books, 2011.

Patridge, Stephanie. "Exclusivism and Evaluation: Art, Erotica and Pornography." In *Pornographic Art and the Aesthetics of Pornography*, edited by Hans Maes, 43–57. Basingstoke: Palgrave Macmillan, 2013.

Pizarro, David A., and Roy Baumeister. "Superhero Comics as Moral Pornography." In *Our Superheroes, Ourselves*, edited by Robin S. Rosenberg, 19–36. New York: Oxford University Press, 2013.

Rea, Michael C. "What Is Pornography?" *Noûs* 35, no. 1 (March 2001): 118–45.

Rothschild, Zachary K., and Lucas A. Keefer. "A Cleansing Fire: Moral Outrage Alleviates Guilt and Buffers Threats to One's Moral Identity." *Motivation and Emotion* 41, no. 2 (April 2017): 209–29.

Rothschild, Zachary K., Mark J. Landau, Daniel Sullivan, and Lucas A. Keefer. "A Dual-Motive Model of Scapegoating: Displacing Blame to Reduce Guilt or Increase Control." *Journal of Personality and Social Psychology* 102, no. 6 (June 2012): 1148–63.

Schopenhauer, Arthur. *The World as Will and Representation.* Translated by E. F. J. Payne. New York: Dover Publications, 1969.

Scruton, Roger. "The Moral Birds and the Bees." *National Review*, September 15, 2003. https://www.catholiceducation.org/en/controversy/marriage/the-moral-birds-and-the-bees.html.

Steiner, George. "Night Words: High Pornography and Human Privacy." In *The Pornography Controversy: Changing Moral Standards in American Life*, edited by Ray C. Rist, 203–16. New Brunswick, NJ: Transaction Books, 1975.

Tanner, Michael. "Sentimentality." *Proceedings of the Aristotelian Society* 77 (1976–77): 127–47.

Threadgold, Steven. "'Struggle Street' Is Poverty Porn with an Extra Dose of

Class Racism." *The Conversation,* May 5, 2015. https://theconversation.com/
    struggle-street-is-poverty-porn-with-an-extra-dose-of-class-racism-41346.
Tosi, Justin, and Brandon Warmke. "Moral Grandstanding." *Philosophy and Public Affairs* 44, no. 3 (Summer 2016): 197–217.
Watson, Jamie Carlin. "Filter Bubbles and the Public Use of Reason: Applying
    Epistemology to the Newsfeed." In *Social Epistemology and Technology: Toward Public Self-Awareness Regarding Technological Mediation*, edited by Frank
    Scalambrino, 47–57. London: Rowman and Littlefield, 2015.

THE TROUBLE WITH FORMAL
VIEWS OF AUTONOMY

Jonathan Knutzen

THERE IS a deep theoretical rift between formal and normative-capacity accounts of personal autonomy. According to formal accounts, personal autonomy consists in conditions that can be specified in purely structural or procedural terms.[1] According to normative-capacity accounts, personal autonomy consists, at least in part, in the possession of capacities for recognizing and responding to the norms that apply to one's choices and attitudes.[2] The first type of view denies that there are any substantive constraints on autonomously formed preferences and attitudes; the second affirms such constraints. In deciding on an account of personal autonomy, the choice between formal and normative-capacity accounts represents an important fork in the road: it is one of

---

1   Christman, "Autonomy and Personal History," "Liberalism and Individual Positive Freedom," "Procedural Autonomy and Liberal Legitimacy," and *The Politics of Persons*; G. Dworkin, *The Theory and Practice of Autonomy*; Ekstrom, "Autonomy and Personal Integration"; Frankfurt, "Freedom of the Will and the Concept of a Person" and "Autonomy, Necessity, and Love"; Friedman, *Autonomy, Gender, Politics*; Killmister, "Autonomy and False Beliefs"; Meyers, *Being Yourself* and "Decentralizing Autonomy"; Westlund, "Rethinking Relational Autonomy."

2   Benson, "Freedom and Value" and "Feminist Second Thoughts about Free Agency"; Kauppinen, "The Social Dimension of Autonomy"; McDowell, "Autonomy and Its Burdens"; Sayre-McCord and Smith, "Desires … and Beliefs … of One's Own"; Sher, *Beyond Neutrality*; Stoljar, "Autonomy and the Feminist Intuition." For a similar view about moral responsibility, see Wolf, *Freedom within Reason*. A brief note on terminology: in the recent literature on personal autonomy, these views are frequently called *normative competence accounts*. I prefer *normative-capacity account*s instead because of the broader connotations of the label. Normative competence suggests psychological infrastructure that is relatively cross-situationally stable and purely internal to the agent. Normative capacity, by contrast, suggests something potentially broader, referring either to normative competence or to the ability to deploy that competence on this or that occasion. Since thinking of the normative abilities involved in personal autonomy in a wide, situation-inclusive sense is plausible, it is important to leave room for this interpretive option. Normative capacity therefore seems like a broader and potentially less misleading label for the family of views.

the deepest and most consequential choice points for our understanding of the nature of autonomy.

Formal accounts of personal autonomy represent the dominant type of view in the existing literature.[3] It is not difficult to see why formal accounts have seemed attractive to many philosophers: they seem to avoid controversial assumptions about normativity and metaphysics and steer clear of undesirable political implications like perfectionism and paternalism. Notwithstanding their attractions, formal views have troubles of their own—troubles that are rarely noticed. As I will explain below, such views have difficulty making sense of the idea that autonomy entails a fairly robust form of responsibility, seem committed to an arbitrary asymmetry between the relevance of facts and values, and cannot properly vindicate the thought that autonomy is reason-giving in roughly the way we take it to be. By contrast, normative-capacity accounts have the resources to deliver a conception of autonomy that secures a robust form of responsibility, treats factual and evaluative information symmetrically, and shows why autonomy plays the normative role it does.

A full-scale comparison of the relative merits of formal and normative-capacity accounts would require considering all relevant theoretical costs and benefits of the competing views. Such comprehensive assessment is beyond the scope of this paper. What I offer instead is a *prima facie* case for reconsidering mainstream views of personal autonomy, highlighting some doubts about formal accounts and briefly indicating why the normative-capacity alternative need not be as troubling as it is sometimes taken to be. While I will not try to settle which type of account is ultimately superior, I do hope to contribute to clarifying what is at

3   Cf. Mackenzie and Stoljar, "Introduction," 13; Westlund, "Rethinking Relational Autonomy," 26. Note that formal views often go by other labels in the literature. Two common labels are "procedural" and "structural." Perhaps the most common label for the type of view I have in mind is "procedural." But there is precedent in the literature for calling the kind of view I have in mind "formal" (cf. G. Dworkin, *The Theory and Practice of Autonomy*, 12; Westlund, "Rethinking Relational Autonomy"), and I think doing so has certain advantages. The label "structural" is frequently associated with ahistorical views, like Harry Frankfurt's, on which only the structure of an agent's attitudes matter, whereas the label "procedural" is frequently associated with historical views, like Gerald Dworkin's and John Christman's, on which it matters how an agent's attitudes come about and on which autonomy is centrally characterized in terms of actual or counterfactual psychological *procedures*. Some views, like Diana Meyers's skill-based account, fit neither model but are nevertheless formal in the sense in which I intend. Since the labels "structural" and "procedural" have these more specific connotations that not all views in the family share, and since "formal" is a natural contrast to "substantive," the term "formal" seems fitting for the broader family of views. However, I want to be clear that nothing I say below rides on this terminological choice. I will define formal views more carefully below. If one prefers the label "procedural" or "structural" for such views, then that is fine; the objections will be the same.

stake in the choice between these two alternatives. Any comprehensive assessment will need to begin by getting clear about the liabilities and advantages of each type of view. There are, as I will argue below, important and underappreciated systematic pressures favoring normative capacity over formal accounts of personal autonomy. If that is right, then appreciating these pressures will be an important step in assessing the merits of competing conceptions of personal autonomy.

The paper is in four sections. The first begins with a brief characterization of formal views. The following three sections articulate several *prima facie* objections to these views: that they cannot furnish an adequate account of responsibility, that they introduce *ad hoc* asymmetries between the importance of facts and values for autonomous agency, and that they are ill-equipped to vindicate the normative role played by the idea of autonomy. At the end of the paper, I briefly address why normative-capacity accounts need not have the troubling political implications some philosophers think they have.

1. WHAT IS A FORMAL VIEW OF PERSONAL AUTONOMY?

Formal views of personal autonomy come in many different shapes and sizes. The most popular variants offer some twist on the idea that to be autonomous one must be in some way identified with one's preferences and attitudes, for example, through taking reflective ownership of them or satisfying the condition that one would not disavow them if one became aware of their source.[4] Some of these views also incorporate external conditions that must be met, e.g., that the formation of preferences occurs in the absence of coercion, manipulation, domination, and so on. Gerald Dworkin's classic statement of a formal view combines these two elements.[5] On his view, personal autonomy consists in "a second-order capacity of persons to reflect critically upon their first-order preferences, desires, wishes, and so forth and the capacity to accept or attempt to change these in light of higher-order preferences and values."[6] Choices must issue from this capacity, but as Dworkin makes clear, they must also do so in conditions

---

4   Christman, "Autonomy and Personal History"; G. Dworkin, *The Theory and Practice of Autonomy*; Frankfurt, "Freedom of the Will and the Concept of a Person"; Friedman, *Autonomy, Gender, Politics*. Frankfurt's classic essay ("Freedom of the Will and the Concept of a Person") centers on the concept of freedom rather than autonomy. However, in a later essay ("Autonomy, Necessity, and Love"), Frankfurt extends his structural view of freedom to the concept of autonomy.

5   G. Dworkin, *The Theory and Practice of Autonomy*.

6   G. Dworkin, *The Theory and Practice of Autonomy*, 20.

of "procedural independence," that is, in the absence of autonomy-undermining conditions like coercion and manipulation.[7]

The defining feature of formal views is that they exclude substantive elements by design. They do so at two levels.[8] First, they place no *direct* constraints on the contents of choice: *any* choice can in principle be autonomously made. Second, they exclude *indirect* constraints on choice in the form of substantively defined attitudes or capacities featuring in the background. This means, for example, that on formal views autonomy cannot require that substantively defined attitudes, like self-esteem or self-respect, feature in the background of choice, as some philosophers have proposed.[9] It also means that they cannot make substantively defined capacities, like the ability to appreciate and respond to genuine values and reasons, a condition of autonomous choice, as normative-capacity accounts maintain.[10]

To be sure, on many formal views, autonomy does require some kind of rational capacity. For example, it may require at least thin, procedurally defined, rational capacities like the ability to be sensitive to coherence constraints on beliefs and desires.[11] According to another popular suggestion, autonomy requires agents to be able to treat considerations *as* reasons. This thought can then be cashed out in functional terms. On a psychological version of the sugges-

---

7   G. Dworkin, *The Theory and Practice of Autonomy*, 16, 18, 20.

8   Cf. Benson, "Feminist Intuitions and the Normative Substance of Autonomy."

9   Benson, "Feminist Intuitions and the Normative Substance of Autonomy."

10 What I am here calling *direct* versus *indirect* substantive constraints is meant to be the same as Benson's contrast between *strongly* and *weakly* substantive accounts of autonomy. The direct/indirect labeling is more intuitive. Suppose your view is that autonomy requires only choice in accordance with some narrowly defined substantive content: perhaps Adam and Eve are autonomous only so long as they do not eat the fruit on that one tree. And suppose my view is that autonomy requires capacities for reasons-responsiveness: to be autonomous, Adam and Eve must enjoy capacities for appreciating the reasons that bear on their choices. Then both of our views are substantive, but in different ways. Is one stronger or weaker than the other? That depends. One might very well interpret your view as imposing stronger conditions on autonomy than mine, or mine as imposing stronger conditions than yours. There is no obvious sense in which your view is substantive in a stronger sense than mine. The *direct* versus *indirect* distinction is clearer. What defines Adam and Eve's autonomy on your view is that they make (or not make) certain choices. The relevant normative constraints thus bear a direct relation to the content of agents' choices. What defines Adam and Eve's autonomy on my view (and equally on views like Benson's on which the agent must have certain self-regarding attitudes) does not depend on what they choose. Anything can in principle be autonomously chosen so long as certain background conditions upstream from choice are met. On this view, the relevant normative constraints bear an indirect relation to the content of agents' choices.

11 Christman, "Autonomy and Personal History," 14.

tion, considerations are treated as reasons when they play a specified role in the agent's psychic economy.[12] On a social version, considerations are treated as reasons when agents are prepared to answer for themselves in the interpersonal exchange of reasons.[13] Crucially, however, on formal views there is no requirement that agents be, or have the capacity to be, attuned to *genuine* reasons. Indeed, there *cannot* be such a requirement consistent with the strictures of formalism. Since they are committed to doing without substantive commitments in specifying the criteria of autonomous agency, formal views are unhitched from objective values and reasons by design. Fidelity to the core commitments of formal accounts therefore requires that whatever rational facility is criterial of autonomous agency, it cannot be reasons-responsiveness in any sense that requires being hooked up to genuine and substantive normative features independent of the agent.

A caveat about this definition is in order. Formal views, I have suggested, rule out direct *and* indirect substantive constraints on choice. In some of the literature on personal autonomy, however, the focal contrast concerns only the first level: it is between views that are *directly* substantive and those that are not. For example, Dworkin contrasts his own view with a substantive account of autonomy, which he understands as placing *direct* constraints on what can be autonomously chosen.[14] Similarly, Marilyn Friedman characterizes her view as "neutral with regard to the content of what a person must choose in order to be autonomous," and contrasts this with a substantive view according to which "someone is not autonomous unless she chooses in accord with certain values."[15] The contrast invoked by Dworkin and Friedman does not perfectly align with the more demanding, two-level definition I have given.

Defining formal views in the more ambitious way is nevertheless appropriate for several reasons. First, it represents a trend internal to theorizing about autonomy by formal theorists themselves.[16] This trend is in the spirit of non-substantive accounts of autonomy like those developed by Dworkin and Friedman, making explicit what these authors left implicit, or at any rate articulating a more thorough version of formalism. Second, the more ambitious two-level definition of formal views I have given follows more recent efforts at taxonomizing. The binary contrast invoked by Dworkin and Friedman does not adequately cap-

12　Bratman, "Intention, Practical Rationality, and Self-Governance."

13　Westlund, "Rethinking Relational Autonomy."

14　G. Dworkin, *The Theory and Practice of Autonomy*, 12.

15　Friedman, *Autonomy, Gender, Politics*, 19.

16　Christman, "Liberalism and Individual Positive Freedom"; Westlund, "Rethinking Relational Autonomy."

ture the space of interesting possibilities. For example, Paul Benson and Susan Wolf give accounts of autonomy on which it partly consists in the possession of reasons-responsive capacities.[17] Benson has subsequently abandoned his earlier view and now defends the idea that autonomy requires certain substantively defined attitudes, like self-respect or a sense of self-worth.[18] Both types of view have a claim to being substantive in an interesting sense, even if they do not require that an agent make specific kinds of choices. For good reason, therefore, both types of view are now routinely classified as substantive.[19] This means formal views are best seen as those that exclude substance at the second and not merely at the first level, i.e., at the level of capacities and attitudes as well as the content of choice. Third, by focusing on a more thorough formalism, the two-level definition presents us with a sharpened and more interesting contrast. Few accounts in the literature are directly substantive in the way envisioned by Dworkin and Friedman—and for good reason. So far as I can see, such accounts do not seem highly compelling. By contrast, indirect substantive accounts of the kind proposed by Benson and Wolf have a good deal going for them. The binary contrast invoked by Dworkin and Friedman risks obscuring the most interesting and relevant alternatives.

Before moving on, it is worth clarifying that the critical upshot of some of the arguments presented below has relevance for views that are not strictly formal in the sense just defined. My criticism targets views that exclude normative capacity (or reasons-responsiveness) as a condition on autonomous choice. Consequently, insofar as the problems identified below are genuine, they will affect all views that exclude (or do not include) such capacities. Hence, views like Benson's, which require attitudes like self-trust or self-respect but do not require substantive normative capacities, are vulnerable to many of the criticisms identified below.[20] I nevertheless focus on formal views because these are popular and represent the starkest, most thorough alternative to thinking of personal autonomy in terms of the possession of normative capacities. They therefore

---

17 Benson, "Freedom and Value" and "Feminist Second Thoughts about Free Agency"; and Wolf, *Freedom within Reason*. Strictly speaking, these views are couched in the language of freedom rather than autonomy. But they nevertheless offer a model for a certain way of thinking about autonomy and are cited as such with some frequency in the literature on autonomy.

18 Benson, "Free Agency and Self-Worth" and "Feminist Intuitions and the Normative Substance of Autonomy."

19 Cf. Benson, "Feminist Intuitions and the Normative Substance of Autonomy"; Christman, "Autonomy in Moral and Political Philosophy"; Mackenzie and Stoljar, "Introduction"; Stoljar, "Feminist Perspectives on Autonomy."

20 Benson, "Feminist Intuitions and the Normative Substance of Autonomy."

constitute the most natural paradigm with which normative-capacity accounts can be contrasted.

With these clarifications in place, let us turn to some problems with formal views of autonomy.

2. THE RESPONSIBILITY CHALLENGE

A central problem with formal views of autonomy is that they cannot deliver an adequate conception of responsibility. In this section, I argue that there is a strong conceptual link between autonomy and responsibility and that reflecting on how best to interpret the notion of responsibility speaks against pure formal views and in favor of normative-capacity accounts.

Begin with the link to responsibility. The association between autonomy and responsibility is widespread in the literature.[21] This is no accident; it reflects important conceptual connections between the two ideas.[22] Whether it is made explicit or not, a basic assumption in much theorizing about autonomy is that autonomy is responsibility-entailing in roughly the following ways:

1. An autonomous agent is a responsible agent.
2. An autonomous agent is responsible for her choices and actions insofar as they issue from relevant autonomy-supporting capacities and circumstances.
3. All else equal, the greater one's autonomy in respect of choices and actions, the more one is responsible for those choices and actions.

Why think autonomy is responsibility-entailing in the sense expressed by these three claims? Let me highlight several pieces of evidence in support of this conclusion.

Performance respect. The exercise of autonomy capacities typically merits

---

21 For a small sampling: Arneson, "Mill versus Paternalism," 475; Buss, "Autonomous Action," 648; G. Dworkin, *The Theory and Practice of Autonomy*, 20; R. Dworkin, *Life's Dominion*, 224; Friedman, *Autonomy, Gender, Politics*, 21–22; Westlund, "Rethinking Relational Autonomy," 30–36.

22 Philosophers do frequently distinguish between autonomy and *moral* responsibility (e.g., Berofsky, *Liberation from Self*; Christman and Anderson, "Introduction"; Killmister, "Autonomy and False Beliefs"; McKenna, "The Relationship Between Autonomous and Morally Responsible Agency"). This shows that one can conceptually distinguish personal autonomy from moral responsibility. It does not show that there are no deep and interesting conceptual pressures linking personal autonomy and moral responsibility, much less that there are no deep and interesting pressures linking personal autonomy and a broader (less domain-centered) notion of responsibility.

a kind of appraisal respect related to the quality of an agent's performance.[23]
When an autonomous agent enjoys circumstances conducive to the exercise of
her autonomy capacities, she merits our esteem or disesteem on the basis of how
she exercises her autonomy. I use the language of esteem and disesteem here
because it has fewer moral connotations than praise and blame and leaves open
the precise connection to moral praise and blame. (On some views, what I am
here calling esteem and disesteem will turn out to be a kind of moral praise and
blame.) What I have in mind is a credit-implying reactive attitude that tracks
the exercise of capacities. While there are forms of appraisal that do not assume
any notion of responsibility (e.g., appraising someone's physical attractiveness),
other forms do, and it is quite plausible to think that exercises of autonomy are of
this kind. At least in this context, esteem and disesteem are crediting responses,
and they suggest that we see the agent as in some meaningful sense responsible
for her choices and behavior. When we respond to a person with esteem or dis-
esteem on the basis of how she exercises her autonomy, we plausibly see her as
meriting such responses *via* her exercise of responsible agency.

   The ethics of paternalism. It is plausible to think that at least part of what makes
paternalism presumptively wrong is that it in some way violates autonomy.[24] In
a suggestive metaphor due to Joel Feinberg, autonomous agents enjoy a kind of
self-sovereignty.[25] Somewhat like the inappropriate meddling by one nation in
the internal affairs of another, paternalistic interventions are thought of as ille-
gitimate incursions into a person's proper sphere of choice. While one might in-
terpret the sovereignty metaphor as suggesting that (hard) paternalism directed
at competent adults can never be legitimate—that sovereignty sets an absolute
side constraint—a weaker claim seems at least as plausible: competent adults
are entitled to strong presumptions against paternalistic interference, making it
difficult to justify warranted interferences for their own good. (Note that in the
international arena, sovereignty is not plausibly absolute either.) The idea that
competent agents are entitled to a sphere of choice is the idea of autonomy as a
right. As I said above, how people exercise their autonomy capacities is associ-
ated with performance respect; by contrast, their right to make choices as they
see fit is associated with recognition respect. To treat persons as little sovereigns

23 See Darwall ("Two Kinds of Respect") for the distinction between appraisal respect and
     recognition respect. Note that what I am here calling performance respect is only *part* of
     what Darwall calls appraisal respect. According to Darwall, appraisal respect includes as-
     sessments both of how agents perform in various roles/practices *and* of their characters.

24 Christman, "Autonomy in Moral and Political Philosophy"; Darwall, "The Value of Auton-
     omy and Autonomy of the Will"; Feinberg, *The Moral Limits of the Criminal Law*; Groll,
     "Paternalism, Respect, and the Will."

25 Feinberg, *The Moral Limits of the Criminal Law*, 47.

is to treat them with due regard for their status as the kind of agents that merit protections against paternalistic interference.

This story seems to me hard to get off the ground without a background assumption about responsibility. Ideas about responsibility are plausibly implicated both in the scope and ground of the presumptive claim against paternalism. First, it is responsible agents who merit special protection against paternalistic interference. This is presumably why we think it is presumptively wrong to paternalize adults but not children. The difference is that adults are responsible agents whereas children are not. Second, facts about responsibility affect the case for and against paternalism, so that (all else equal) paternalism becomes harder to justify as responsibility increases and easier to justify as responsibility decreases. This is presumably part of the reason why it is much easier to justify soft paternalism. Think of Mill's classic example of a man about to walk over a bridge he does not know is unsafe. Paternalism is easier to justify in such a case than it is to justify in the case where the man, knowing the bridge is unsafe, intends to walk on it. This is so *whatever* one thinks about the all-things-considered justification of paternalistic intervention in the two cases. The case for paternalistic intervention is stronger in the first case than in the second, and a natural explanation for this is that facts about responsibility are salient: given his ignorance, the first man is less responsible for his choice (and the outcome of that choice) than the second.

If part of what makes paternalism presumptively wrong is that it violates autonomy, we have here a powerful reason to think autonomy is responsibility-entailing. To be sure, one could coherently accept that paternalism violates autonomy and that the case against paternalism is sensitive to facts about responsibility while denying any connection between the two. One might, for example, think paternalism involves the double wrong of violating autonomy *and* being inappropriately sensitive to facts about responsible agency. But this needs motiving. Antecedently, the simpler explanation connects the two: paternalism violates autonomy and autonomy entails responsible agency; it is the fact that agents are capable of being responsible for their own lives and choices that (in part) makes them autonomous; it is this very same fact that grounds a strong claim to being left free to pursue their lives and choices as they see fit. This picture is elegant in its simplicity, and it forges a straightforward connection between autonomy and responsibility.

Two brief caveats about the picture are in order. First, autonomy can imply responsibility without entailing that responsibility is sufficient for autonomy. There might well be additional elements to autonomy and, therefore, additional wrong-making features to paternalism. It is common, for example, to distinguish autonomy as a right from autonomy as an agency ideal. Plausibly, autonomy as

an agency ideal is more demanding than mere responsible agency. But it never-theless entails responsible agency. For the ideal to be in the offing, one has to be a responsible agent, capable of being responsible for one's life choices in some suitably rich and meaningful sense. Second, a well-justified regime of anti-pa-ternalist norms presumably has multiple sources of justification.[26] My point is not that considerations of responsibility are the exclusive source of justification for anti-paternalist norms, but that they are one important plank in the ethics of paternalism. Moreover, it is only fair to acknowledge that justifications of an-ti-paternalist norms are conceivable that make no appeal whatsoever to respon-sible agency (e.g., that agents typically know best what is in their interest). To fully defend the claim that strong anti-paternalist norms are best justified by a background assumption of responsible agency would require showing that al-ternative justifications, which do without the assumption of responsible agency, are not sufficient. That is more than I can do here. I will therefore content myself with making the bet that these alternative explanations fail. They may contribute to partial justifications for anti-paternalist norms, but it is doubtful that they can deliver complete and adequate justifications.[27]

When someone decides to smoke or climb dangerous mountains, that choice plausibly merits respect *as* the choice of a responsible agent. Hence, as noted above, when someone is adequately informed about the risks, there is also stronger reason to desist from interfering. This is a backward-looking responsi-bility rationale: all else equal, there is more reason to allow persons to reap the consequences of their actions when they are undertaken responsibly than when

---

26 Thanks to Richard Arneson for helping me to see this more clearly.

27 For example, epistemic arguments are fragile. While it may be true that people tend to be better judges of what is in their interest because they have access to a richer base of relevant information, that is not always so. Moreover, many of us want to say that paternalism can wrong someone even if we grant that the paternalizing agent knows better. Another candi-date explanation for the presumptive wrongness of paternalism, in many ways elegant in its simplicity, is that people just *do not like* being forced to do things against their will. (Thanks to an anonymous reviewer for *JESP* for raising this possibility.) As with the epistemic argu-ment, this type of explanation may contribute to a general anti-paternalist rationale, but I doubt it can be the whole story. Strong-willed agents do not merit greater anti-paternalist protections than docile agents; very small and non-responsible stubborn children do not merit stronger anti-paternalist protections than equally non-responsible but agreeable chil-dren. Moreover, the reasons given by an agent's desires seem not to ground the full-orbed recognition respect that seems appropriate to an adult chooser. To be sure, honoring peo-ple's wishes so far as possible is plausibly a way of respecting them. However, the recogni-tion respect appropriate to an adult chooser seems to require more: the recognition not only that here is an individual with desires and a perspective on the world (a child is an agent of this sort), but that here is a *responsible* agent—an agent whom it could make sense to let live with the potentially momentous consequences of her choices.

they are not. One might prefer a more forward-looking responsibility rationale instead.[28] Perhaps a regime of anti-paternalist norms can be partly justified by its proleptic or educative effects, tending to cultivate capacities for responsibility of roughly the kind it appears to assume. Either way, without the idea that persons are, or can become, responsible choosers, it is very difficult, I think, to support quite robust and general anti-paternalist presumptions of the sort most people in liberal societies subscribe to. The point here is not that people always live up to this picture of responsible agency or that facts about responsibility entirely settle issues about the ethics of paternalism. Rather, it is that our commitments to anti-paternalist norms plausibly depend on a deep background assumption of responsible agency.

Options. The third line of evidence comes from the persistent attractiveness of the idea that options matter for personal autonomy.[29] Raz gives memorable examples.[30] The man who falls into a pit and can only decide when to nap or which direction to move his head is not very autonomous. Nor is the woman who is trapped on an island with a hungry beast and who spends her every waking moment trying to avoid being eaten by it. Something similar goes for the slave, who lacks options and cannot choose his own course through life, as well as for the many more prosaic forms of impoverishment that may not involve domination but nevertheless involve restricted options.[31] For example, it seems natural to describe refugees trapped in refugee camps as suffering diminishment of autonomy.[32]

Lack of options is constraining; it leaves agents less free to choose their course. It thereby also tends to diminish responsibility. Those who lack adequate options will tend to be less responsible for their choices and for the consequent shape of their lives.[33] The man in the pit is responsible for a few things—for whether he naps now or later, for which way he turns his head. But he is not responsible for much else about his life. His constrained circumstances change how it is appropriate to appraise the man. Before he fell into the pit, it might have

---

28 Cf. Vargas, *Building Better Beings*.

29 Hurka, "Why Value Autonomy?"; Kauppinen, "The Social Dimension of Autonomy," 284–86; Mackenzie, "Three Dimensions of Autonomy," 28; Mackenzie and Stoljar, "Introduction," 22, 26; Oshana, "Personal Autonomy and Society," 94, and *Personal Autonomy in Society*; Raz, *The Morality of Freedom*, 373–77; Terlazzo, "Conceptualizing Adaptive Preferences Respectfully."

30 Raz, *The Morality of Freedom*, 373–74.

31 Oshana, "Personal Autonomy and Society"; Nussbaum, *Creating Capabilities*; Sen, *Development as Freedom*.

32 Betts and Collier, *Refuge*, ch. 6.

33 Cf. Hurka, "Why Value Autonomy?"

been appropriate to feel some disesteem for him because, while enjoying signif-
icant talent and opportunities, he spent most of his days playing video games.
Now that he is trapped in the pit, however, it would be absurd to feel disesteem
for him on the basis of his unambitious choices. Because he lacks opportunities
to exercise his agency capacities in a meaningful way, such performance-based
assessments would be out of place. The impact of limited opportunity on *moral*
accountability is familiar from fair-opportunity accounts of moral responsibil-
ity.[34] Something similar seems plausible in the case of personal autonomy. In
general, lack of options tends to spell diminishment of autonomy. If personal au-
tonomy implies responsibility, we can make sense of this. Limited opportunity
undermines or threatens autonomy because, all else equal, it makes persons less
responsible for their choices and lives.

Self-authorship/self-creation. A final piece of evidence for the link between
autonomy and responsibility is to be found in widespread appeals to tropes of
self-authorship and self-creation throughout the autonomy literature.[35] These
metaphors express something deep and important about what it means to be
autonomous, yet they are hardly intelligible without some background idea that
persons are responsible for their lives. Creators and authors, after all, must be
more than merely *causally* responsible for the products they create or author.
To be self-authors or self-creators in any meaningful sense, persons must enjoy
the right kind of responsibility-conferring relationship to their choices and lives.

Together, these four lines of evidence suggest significant connections be-
tween personal autonomy and responsibility. In particular, they suggest that it
is plausible to think of autonomy as responsibility-entailing in roughly the way
suggested: that to be an autonomous agent, one must be a responsible agent;
that an autonomous agent is responsible for her choices and actions when they
issue from favorable circumstances; and that, all else equal, greater autonomy
in respect of choices and actions implies greater responsibility for them. If the
connections we have noticed are genuine, it is little wonder that the idea of re-
sponsibility crops up with some frequency in discussions of autonomy. There
are systematic pressures supporting the idea that autonomy is responsibility-en-
tailing.[36]

34 Brink and Nelkin, "Fairness and the Architecture of Responsibility."

35 E.g., Benn, "Freedom, Autonomy and the Concept of a Person," 125, 127; R. Dworkin, *Life's
Dominion*, 224; Enoch, "Hypothetical Consent and the Value(s) of Autonomy," 27; Griffin,
On Human Rights, 150; Raz, *The Morality of Freedom*, 369–70, 390.

36 An anonymous reviewer from *JESP* asks a helpful clarifying question: Does the entailment
run only in one direction or in two? And what exactly is the intended order of explanation
here? Answer: what I am suggesting is that responsible agency is a necessary but not a suf-
ficient condition of autonomous agency. If we imagine a Venn diagram, the circle labeled

Now for the trouble. Many formal theorists accept that autonomy comes with responsibility. Consider a representative quote from Gerald Dworkin: "By exercising [their capacities for autonomy], persons define their nature, give meaning and coherence to their lives, and take responsibility for the kind of person they are."[37] The question is whether they have the resources to make sense of this commitment. More specifically, the question is whether they can deliver a notion of responsibility that is adequate to the task. Formal views spell out the conditions of autonomy in terms of properties like structural mesh between attitudes of higher and lower orders, agential coherence, actual or counterfactual reflective endorsement, forming temporally extended plans, treating considerations as reasons in the evaluation and adoption of plans, and so on. Such properties do seem well-suited to furnishing the basis for *some* ascriptions of responsibility. In particular, they seem to support judgments of attributability, according to which agents are related to their actions in such a way that their actions manifest their character and commitments.[38] In the case of moral conduct, responsibility-as-attributability typically means that actions express an agent's quality of will. But the idea of attributability can be generalized to cases not limited to moral matters. An agent will be attributively responsible for her life choices (even purely self-regarding ones) if they reflect on her—on what kind of person she is, on her sense of self, on her character, priorities, commitments, and values. Many formal views of personal autonomy are preoccupied with authenticity conditions. These aim to tell us when some choice or attitude is the agent's own in a special sense. Such accounts therefore seem to be well-equipped to capture

---

"autonomy" would be entirely within the circle labeled "responsibility"—the one is a subset of the other. On this picture, autonomy entails responsibility but not vice versa. Note that this is a *logical* claim. On a plausible interpretation, the supporting metaphysics that would make this logical claim true would involve a claim about constitution: autonomy is constituted, at least in part, by responsible agency. This makes responsibility (at least in one sense) *explanatorily* prior.

37   G. Dworkin, *The Theory and Practice of Autonomy*, 20.

38   Following Gary Watson ("Two Faces of Responsibility"), many discussions of moral responsibility distinguish two senses of responsibility: attributability and accountability. Roughly, one is responsible in the attributability sense if one's actions reflect one's quality of will, and one is responsible in the accountability sense if one's actions meet whatever control conditions are required for being held morally accountable. A plausible specification of the control conditions involved in moral accountability is that they consist (at least in part) in the possession of normative capacities (Brink and Nelkin, "Fairness and the Architecture of Responsibility"; Fischer and Ravizza, *Responsibility and Control*; Nelkin, *Making Sense of Freedom and Responsibility*; Wolf, *Freedom within Reason*). By contrast, a plausible specification of attributability-relevant conditions requires only that an action reflect something like the agent's genuine or authentic self—her character, perspective, or will.

the sense in which people can be attributability-responsible. When agents meet the requisite authenticity conditions, they stand in the relation of ownership to their choices and attitudes such that those choices and attitudes reveal where the agent stands, what she is about, and so on. Such views can therefore yield an important sense of responsibility: the kind that reveals something of the agent's inner life, putting her on display and opening her up to certain forms of appraisal.

The crucial question is whether this conception of responsibility is the right kind. Is it adequate for an account of personal autonomy? Two considerations suggest it is not.

First, as we have seen, the exercise of autonomy capacities typically merits performance respect. Since formal views of autonomy can plausibly make sense of attributability-responsibility, they plausibly have the resources to make sense of certain forms of appraisal respect: character-grading, aretaic judgment, assessment of motive, and so on. But performance respect requires something more specific. When an agent merits our respect for the exercise of her autonomy, her actions must meet a kind of credit condition such that the agent can earn our esteem or disesteem on the basis of how she exercises her autonomy. Some crediting responses are quite weak: they amount only to something like approval or disapproval. Attributability responsibility suffices for making this weak class of responses apposite. Other crediting responses, however, are stronger: they amount to something like performance criticism. It is not clear attributability responsibility suffices for this stronger class of responses.

Consider that the facts determining choice-worthiness are normative. Since the paradigm of personal autonomy is often taken to be self-regarding choice, consider for simplicity the domain of prudence. On all of the most widely held and plausible views of welfare, there are facts about what is good for agents that is independent of their momentary desire and whim. Choices in this domain can be better and worse, right and wrong, wise and unwise, and so on.[39] It is hard to see how the stronger class of crediting responses could be apposite in the absence of sensitivity to the very facts that determine choice-worthiness. To be a suitable target of performance criticism on the basis of how an agent exercises her autonomy capacities, she must enjoy the right kind of control. But it is hard to see how the agent could enjoy such control in the absence of normative capacities. Agents who satisfy formal autonomy criteria but lack normative capac-

---

39 This is true on hedonist, objective-list, and perfectionist views, but it is true on the most compelling versions of the desire-satisfaction view as well, which add counterfactual and idealizing conditions as a filtering mechanism on desires, the fulfillment of which count toward a person's welfare.

ities seem a bit like blindfolded dart throwers attempting to hit a target.[40] Why should an agent merit our disesteem if she is completely insensitive to the facts in virtue of which she ought to choose one way or the other or cannot suitably regulate her conduct in light of this sensitivity? Unhitch agents from the relevant normative facts, either because they are blind to them or incapable of acting on them, and it becomes very hard to see how they can be responsible for their choices in the way that is characteristic of the kind of performance respect we associate with the exercise of autonomy.

Second, as we also have seen, on a plausible interpretation of the ethics of paternalism, both the scope and grounding of anti-paternalist principles are sensitive to facts about responsible agency. Does being attributability-responsible suffice to ground robust anti-paternalist norms? It is hard to see how it could. The same facts that make it difficult to see how an agent who lacks normative capacities could have the kind of control needed to render performance criticism apposite also make it hard to see how it could ground a strong claim against paternalistic interference: it is precisely because children lack such capacities that they do not have a strong claim against intervention by parents and educators.

Formal views of autonomy do frequently posit reflective ownership capacities. Would such capacities suffice to merit anti-paternalist protections? It is hard to see how. Again, the domain of choice is governed by practical norms. Stipulate that the agent is insensitive to these norms and it becomes difficult to see why she merits strong protection against intervention by third parties. To be sure, since facts about responsibility are not the only relevant facts for determining the appropriateness of paternalistic intervention, there may still be all-things-considered reasons to protect her choice even if she is not sufficiently responsible. But, as I argued above, facts about responsible agency are a huge pillar in the anti-paternalist case. Once this is acknowledged, we need an interpretation of the relevant notion of responsibility. What we need is a kind of responsibility that is robust enough to ground strong anti-paternalist norms and (as a corollary) puts agents on the hook for the upshots of their own choices. Mere identification with, or ownership of, attitudes seems much too weak. What matters is not that a choice is authentically yours; what matters is that the choice is one for which you are robustly responsible, in such a way that you can be on the hook for its consequences, and that I have strong presumptive reasons to let you make your choice even if I could do better. To secure this result, something more is needed. Adding reflection does not do the trick. Perhaps reflective endorsement increases authenticity, such that you then own the choice in a special and

---

40 For the metaphor of blindness, see Kauppinen, "The Social Dimension of Autonomy," 281; Wolf, *Freedom within Reason*, 92.

deeper way. This may say something about you—about your character, perspective, and values. In that minimal sense, it constitutes a kind of responsibility. But reflective endorsement does not amount to enjoying a relation to your choices that would explain why other agents have presumptively decisive reasons to let you have your way, even when they know better. The fact that a choice is *authentically yours* just does not seem to have the right kind of normative relevance to ground such reasons. By contrast, if you have normative capacities for appreciating and responding to the values and reasons bearing on your choice, then that does seem to do the trick. If you can appreciate and respond to the normative features relevant to your choice, then that gives you a deeper kind of control over your choices and actions, putting you on the hook for their upshots, and giving me reasons to desist from paternalistically interfering with your choice.

To return to the difference between children and adults: What is the salient difference between children and adults, such that paternalism of the former is generally more acceptable than paternalism of the latter? It is plausibly a difference in their status as responsible agents. But what kind of responsibility is relevant here? I have argued that strong anti-paternalist protections would be better supported by a form of responsibility that puts agents on the hook for the upshots of their choices than a form of responsibility that merely reveals what kind of person they are, where they authentically stand, etc. If that is right, the relevant difference between children and adults seems to be that the former have more fragile normative capacities than the latter, not that they have a less well-developed sense of self. This has some intuitive plausibility. Think of it this way. You are about to meet a seven-year-old child who is a stranger to you. All you know is that the child is extremely precocious and wants to undertake a dangerous activity. You have the power to stop her. Which set of facts would ground a stronger claim against you not to interfere with her choice? The fact that she is reflectively mature and seems to have crystalized a perspective and stance on the world that is genuinely her own? Or the fact that she is reflectively mature in such a way that she seems sensitive to normatively relevant features of her choice? I suspect you will agree that normative capacities ground stronger claims against intervention than mere authentic ownership of choices. While children generally have more fragile normative capacities and a less developed sense of self, it is the first of these properties that seems more important in considering whether paternalistic treatment is warranted.

Together, these considerations put enormous pressure on formal views of autonomy. What we need, I have argued, is an interpretation of autonomy that can deliver a robust conception of responsibility. The worry is that formal views cannot deliver such a conception. They can give us a conception of responsibil-

ity that shows us where the agent stands and thereby reveals something good or bad about her. But they cannot give us a conception of responsibility that shows the agent to be in the kind of relationship to her attitudes and choices that seems to be required by our treating exercises of her autonomy as meriting positive or negative performance respect; nor can they give us a conception of responsibility robust enough to ground strong claims in favor of allowing the agent to live with her choices and against others that they not paternalistically interfere.

3. THE FACT/VALUE ASYMMETRY

Perhaps formal theorists will succeed in giving us a rich and convincing story about responsibility. Even so, a further problem looms. Responsibility requires adequate non-evaluative information. This idea is familiar from discussions of *moral* responsibility, where ignorance is typically taken to be exculpatory. Ignorance, of course, does not *always* excuse, as in the case of willful or negligent ignorance, but ignorance can and frequently does serve as an excusing condition in assessments of moral responsibility. The analogous thought is plausible in the case of personal autonomy as well: just as inadequate information can diminish moral responsibility, inadequate information can diminish personal autonomy. Someone who smokes in complete ignorance of the risk this poses to her health is plausibly less autonomous with respect to that choice than someone who is apprised of the facts and chooses to smoke anyway; a lover who marries her beloved ignorant of his true character is less autonomous with respect to that choice than someone who knows her lover in greater depth; and so on. All else equal, more choice-relevant information means more autonomy; less choice-relevant information means less autonomy.

It is possible to deny that non-evaluative information is relevant to autonomy. Michael McKenna commits himself to this bold thesis in an effort to discover some interesting differences between moral responsibility and personal autonomy.[41] In his central example, Tal attempts to help his sick friend, Daphne. Pulling mislabeled medicine from the cabinet, Tal gives Daphne poison and thereby accidentally poisons her. According to McKenna, Tal acts autonomously (though he is not morally responsible) in poisoning his friend. This is because, explains McKenna, there is a sense in which Tal rules himself by acting in accordance with self-chosen principles. The principle on which Tal acts is: always attempt to help those who suffer innocently. And Tal's action conforms to this self-chosen principle because, in administering the drug, he does attempt to save his friend.

41 McKenna, "The Relationship Between Autonomous and Morally Responsible Agency."

The resulting picture is a fairly stark form of internalism on which autonomy is compatible with sweeping ignorance of relevant facts.

This is not compelling. On its own, the example seems to lend intuitive support to the thought that Tal's autonomy is undermined or threatened by his ignorance. So do similar examples discussed by Al Mele, like the example of Connie, who chooses an investment plan but is systematically deceived by the company offering the plans, and King George, who rules his kingdom contrary to his deepest commitments because his staff systematically distorts the information arriving at his desk.[42] If anything, it seems intuition antecedently favors the verdict that these agents suffer some impairment of autonomy by being informationally cut off. McKenna acknowledges the intuitive pull of Mele's examples, but he insists that the intuitive pull tracks moral responsibility rather than personal autonomy. If we stipulate that autonomy is acting in light of self-chosen principles, McKenna suggests, Tal and Connie and King George can all be seen as autonomous. But this is not obvious. Even granting McKenna's stipulative definition of autonomy, no strong form of internalism follows, for it is plausible to suppose that acting in light of one's principles imposes success conditions on action that are not met in the examples.[43] McKenna avoids this problem by describing Tal's action (and by implication, Connie's and King George's) as an *attempt*. Tal's principle is: *attempt* to help your friends. This is something he succeeds in doing. But suppose his principle were: help your friends. This is not something he succeeds in doing. The intuitive force of the examples as instances of autonomy thus depends on an artefact of description. Tal and Connie and King George would, of course, not be identified with their actions under an informationally enriched perspective. Once we shift the act description to a more objective frame, it becomes much less compelling to think of them as autonomous. Consider another possible fix. One might describe all principles of action in evidence-relative terms. Tal's principle might be: act to fulfill your goals and values as indicated at the moment of action by your subjective evidence base. He might then, for example, grievously harm Daphne while counting as exemplary in his autonomy simply because he acts on his (misleading) evidence about what helps and harms her. But what would serve the values and principles of agents is not typically evidence-relative or subjective in this way: Tal cares about his friend, Connie cares about her future in retirement, King George cares about the flourishing of his kingdom, and so on. This outward focus imposes objective success conditions that require being suitably well-informed if one is to promote the relevant values and principles. Antecedently, as I said, intuition seems to fa-

42   Mele, *Autonomous Agents*, 179–81.

43   For similar points of criticism, see Killmister, "Autonomy and False Beliefs."

vor Mele's verdict about the cases, that autonomy is threatened by deprivation of decision-relevant information. One can try to deflate some of the intuitive force of these examples, as McKenna seeks to, but only by redescribing the principles and values from which agents act in terms of implausibly unambitious success conditions—as attempts or evidence-relative respondings. If one sticks with a realistic interpretation of what agents actually care about, then their being autonomous plausibly does depend on being adequately informed.

As we have seen, there are strong conceptual and theoretical pressures to preserve the association between personal autonomy and responsibility. McKenna arrives at his conclusion precisely in an effort to locate some interesting notion of personal autonomy that *comes apart* from responsibility. But he offers neither theoretical motivation nor robust conceptual anchor points for this strong internalist suggestion. If, as I argued above, autonomy is a form of personal freedom in virtue of which agents are responsible, then we have good reasons to reject the kind of extreme informational hermeticism on which an agent can be completely ignorant or deceived about factual information relevant to her choice.

Theoretical pressure is increased by noticing the connection between being informed and having control. Ignorance threatens an agent's control.[44] The examples of Tal, Connie, and King George exemplify this. By being significantly ignorant, these agents have impoverished control over their actions. And control seems fairly clearly relevant to autonomy. Think of a case involving complete absence of executive control. Perhaps in the inner sanctum of my mind I endorse normative principles and aim to conform my actions to them. It seems utterly implausible to think that I enjoy autonomy if I have no power whatsoever to conform my actions to my principles. But if lack of control threatens autonomy on the "active" side, why would it not do so on the "receptive" side as well? After all, both executive capacities and representational capacities can be thought of as aspects or dimensions of control. In the absence of reasons to posit an asymmetry in control conditions, it seems arbitrary and unmotivated to insist that one dimension of control matters while the other does not. In short, McKenna's proposal is an interesting suggestion about how to divide conceptual labor between moral responsibility and personal autonomy, but we have few independent reasons to accept it and some independent reasons to reject it.

Must we go to the other extreme and hold that only those who act in light of all relevant information are autonomous? This would entail the absurd conclusion that almost no one ever acts autonomously. We need not accept this extreme conclusion. An intermediary view is available, namely that *sufficient* information is necessary for autonomy. This is plausible if we distinguish scalar and

---

44 Cf. Mele, *Autonomous Agents.*

threshold assessments of autonomy. Many of the properties relevant to autonomy (like being informed) are a matter of degree. Deploying a scalar conception of autonomy, we can say that agents are more autonomous the more they satisfy the relevant scalar property. Switching to a binary, threshold conception, we can say that some threshold level of the property must be reached to qualify as autonomous. The two conceptions can be combined. On such a picture, autonomy "kicks in" above the threshold but one can be more or less autonomous (perhaps with no upper bound) above that point. When it comes to being informed, the combined conception seems plausible. Below some threshold of understanding, agents may not be autonomous with respect to a choice at all. Above that threshold, being more informed tends to enhance, and being less informed tends to diminish, autonomy. Citing Columbus's ill-informed decision to sail west, Nomy Arpaly doubts "that anyone wishes to claim … that an ill-informed decision cannot be an instance of autonomous agency."[45] But she also accepts that giving someone more information might make the person more autonomous. Once we distinguish scalar and threshold verdicts, both of these claims seem plausible. Being ill-informed can be autonomy-impairing while not rendering one entirely non-autonomous.[46]

Now for the challenge. Suppose formal views accommodate the idea that non-evaluative information is relevant to autonomy. By parity of reasoning, it seems plausible to suppose that autonomy likewise requires *normative* information. It is hard to see why someone who is completely normatively "blind" would be any more autonomous than someone who is ignorant of non-evaluative information. Suppose someone smokes, knowing the risk this poses to her health but in total ignorance of what is good for her or the reasons this gives her to make one choice rather than another. Such a person seems just as blind, in the relevant sense, as someone who is ignorant of the non-evaluative facts. There

---

45   Arpaly, "Responsibility, Applied Ethics, and Complex Autonomy Theories," 175.

46   So far as I can see, the *source* of misinformation is irrelevant to autonomy. Tal's autonomy is not lessened more if his misinformation is the result of intentional manipulation than if it is the result of accidental labeling. Similarly, if Connie and King George are informationally impaired due to a fluke of circumstance, this is no less an impairment of their autonomy than if they are the victims of campaigns of disinformation. To be sure, we are likely to see social sources of autonomy deficits as having special moral significance. We can, for example, condemn (unjustified) manipulation and blame those who engage in it. But ignorance is ignorance and seems threatening to autonomy no matter what its source. If my autonomy is threatened by your intentional deception of me, there is little reason to think it would be less threatened if my ignorance is the result of impersonal forces. Of course, this is holding all else equal. Social sources of misinformation may be worse threats to autonomy insofar as one person is in the power of another. If your manipulation is a way of also dominating me, then the way in which my ignorance originates seems more significant.

is, then, a simple parity argument for treating factual and normative ignorance alike.[47] If autonomy implicates responsibility, and if both factual and normative ignorance can defeat or attenuate responsibility, we have (in the absence of further considerations) no more reason to credit autonomy in the absence of the one than in the absence of the other.

Extreme internalist conceptions of autonomy are implausible. They suggest that complete factual ignorance does not in any way threaten autonomy. Formal theorists therefore do well to accept that non-evaluative information can make a difference to autonomy.[48] But once this much is accepted, there is pressure to accept that evaluative information is relevant to autonomy as well. If one accepts that autonomy is responsibility-entailing, there is a principled rationale for taking this further step. Sensitivity to evaluative information is just as relevant as purely factual information in constituting an agent as responsible. The domain of choice is one in which norms apply: choices can be better or worse, right or wrong, prudent or imprudent, and so on. Truths about choice-worthiness are a function, not of descriptive facts *per se*, but of descriptive facts *plus* relevant evaluative or normative truths. Hence, truths about choice-worthiness are partly normative. But truths about choice-worthiness also furnish the basis for critical assessments of agents. Factual and normative deficits alike tend to be responsibility-diminishing: below some minimal threshold level, agents are not responsible for their choices at all; above that level, they are more or less responsible, depending on their sensitivity to the relevant features.

Formal views, however, must reject parity. Since they sever the connection between autonomy and substantively defined evaluative capacities, such views must also deny that normative information matters for autonomy.[49] This creates an explanatory burden. On the face of it, the exclusion of evaluative information seems *ad hoc*. This puts pressure on formal accounts to explain why evaluative and non-evaluative information should be treated in an asymmetric fashion. The answer cannot be: because that is what is predicted by formal theories. In the absence of some salient difference we have independent reasons for accepting parity. This speaks against formal theories precisely because they predict an asymmetry. It is therefore not satisfactory to point to this implication of formal

47 On the symmetry of facts and values in moral responsibility, see Rosen, "Culpability and Ignorance"; and Wolf, *Freedom within Reason*. For the parallel claim about personal autonomy, see Kauppinen, "The Social Dimension of Autonomy," 280; and Savulescu, "Rational Non-interventional Paternalism," 330.

48 Cf. Berofsky, *Liberation from Self*; Killmister, "Autonomy and False Beliefs"; Mele, *Autonomous Agents*.

49 Killmister does just this, arguing that false factual beliefs tend to impair autonomy but not false principles or values ("Autonomy and False Beliefs," 527).

theories in reply to the challenge. Perhaps formal theorists can ultimately give us some principled, non-question-begging story about why we should treat normative and factual information differently. In the meantime, we have a *prima facie* case for thinking autonomy requires sensitivity to evaluative information. This speaks against formal views of autonomy.

4. AUTONOMY'S NORMATIVE ROLE

The final worry about formal views is that they cannot make adequate sense of autonomy's normative role. We recognize autonomy's normative role in the kinds of reasons it supplies. Autonomy is reason-giving in roughly two ways. On the one hand, we think it good, all else equal, for people to live autonomous lives and make choices autonomously. It is therefore the sort of thing we have reason to aspire to ourselves and promote the realization of in others. On the other hand, we think autonomy marks out a sphere within which individuals are free to choose and that their autonomous choices carry authority or gravity in certain contexts of decision-making to which we must often give greater weight than the choices would merit on the basis of their direct consequential value or other forms of choice-worthiness.[50] When an agent or her choices meet the conditions of autonomy, we must take her decision with special seriousness. Even when it is trumped by other considerations, autonomy places the bar of interference higher than it otherwise would be: it ratchets up the demands for warranted intervention.

An adequate conception of autonomy should be able to make sense of this twofold normative role. In other words, an adequate conception of autonomy needs to vindicate the thought that autonomy is worthy of promotion and worthy of respect. But there are reasons to doubt that formal views provide ingredients sufficient to meet this demand. Let us consider each of these normative roles in turn.

What are the kinds of autonomy-relevant conditions we generally have reasons to promote? The most obvious is perhaps this: to ensure that people have sufficiently valuable options to choose from. Some autonomy theorists, like John Christman, deny that valuable options matter for autonomy.[51] But often the

---

50   One might object to putting the point in terms of the weight of reasons. According to Groll, an autonomous will is to be taken as "structurally decisive" ("Paternalism, Respect, and the Will," 699–706). However, Groll suggests that paternalism is only "presumptively wrong" (710–11). So even on a Raz-style view like Groll's, where certain considerations are shielded from entering deliberation, the shield is not necessarily absolute.

51   Christman, "Procedural Autonomy and Liberal Legitimacy," 282. Christman subsequently argues that if valuable options matter to autonomy this should be understood in terms of a

motivation for this denial is heavily theory driven, for example, by the desire to avoid perfectionist implications. Pre-theoretically, it is quite natural to describe people with limited valuable options as suffering a diminishment of autonomy.[52] Consider refugees living decades of their lives in a camp. These people typically have a dearth of valuable options and it is natural to think of them as suffering from an autonomy-relevant impairment as a consequence.[53]

Valuable options are an external good. We plausibly also have reasons to promote an internal good to go along with it. Think of what parents want for their children. Parents do not just want their children to face a lush banquet of valuable options; they want their children to possess the capacities to appreciate and respond appropriately to those options. This pattern of concern seems appropriate more generally. It would seem a bit odd to care that persons enjoy valuable options but not to care that they enjoy capacities for appreciating and responding appropriately to those valuable options. Some formal theorists argue that autonomy requires valuable external options but that the internal capacities should nevertheless be understood in terms of purely procedural conditions.[54] This is an unstable position. Once one accepts that valuable options matter, why not also accept that internal competencies for tracking and pursuing those valuable options matter?

There are, of course, complicated questions about who may promote whose autonomy and how this may be done. Some liberal theorists, for example, insist that the state may play no role in promoting the autonomy of its citizens. Moreover, there are plenty of cases where we have reasons not to promote, and even to curtail, autonomy—for example, prospectively, when people's exercise of autonomy will likely bring about significant and unjustified harms to others, and retrospectively, for purposes of punishment. But the present point does not depend on denying such qualifications. What it depends on is only the broad generalization that people ordinarily have robust reasons to promote their own autonomy and often also the autonomy of others. The exceptions are important, but they should not obscure the fact that there are general standing reasons for anyone to promote anyone else's autonomy. A plausible interpretation of what people generally have reason to promote includes (i) valuable options and (ii) normative competence over those options. It is consistent with this to think that

---

subjective conception of value, i.e., the options need only be valuable from the perspective of the agent (*The Politics of Persons*, 170).

52 Cf. Nussbaum, *Creating Capabilities*; Oshana, "Personal Autonomy and Society" and *Personal Autonomy in Society*; Raz, *The Morality of Freedom*.

53 Betts and Collier, *Refuge*, ch. 6.

54 Terlazzo, "Conceptualizing Adaptive Preferences Respectfully."

there are secondary considerations excluding states or other agencies from the role of autonomy promoter and that in some cases there is most reason *not* to promote autonomy.

This specification of what people often have reasons to promote fits elegantly with a normative-capacity account. It does not fit well with formal accounts. When we think about the kinds of properties identified by formal accounts—reflective acceptance, ability to treat a consideration as a reason, answering for oneself in the social exchange—it is at least not obvious whether and why we have reasons to promote these things. Perhaps we do have reason to promote these things; much will depend on how the details are spelled out. But it surely is not obvious that we have quite general and powerful reasons to promote these properties. Contrast this with our confident commitment to promoting autonomy. Barring complications about special secondary reasons some agents might have not to be autonomy promoters, we think there are standing agent-neutral reasons to promote *anyone's* autonomy. This confidence is readily vindicated if autonomy turns out to require (i) valuable options and (ii) normative competence. We can readily appreciate why these twin goods would be valuable and worthy of promotion. Perhaps formal views can ultimately rise to the challenge of explaining why the properties they posit as constituents of autonomy are worthy of promotion. But the case needs to be made. There is at least a *prima facie* challenge here: normative-capacity accounts are well-positioned to make sense of the idea that we generally have reasons to promote people's autonomy; formal accounts, by contrast, are not so obviously well-positioned—whether they can make sense of the reasons we have to promote autonomy is more of an open question.

Perhaps, however, this is an unfair assessment of the situation. Consider the following problem. There is an ambiguity in the idea of reasons-responsiveness: Does it mean merely *having capacities* for responding to reasons or *actually exercising* those capacities? Which of these does the normative-capacity account of autonomy appeal to? Is the mere capacity for responding to reasons enough for autonomy or must one also exercise one's capacities in such a way as to conform to one's reasons? The label—normative-capacity account—certainly suggests the former. There are also systematic pressures encouraging this interpretation. Intuitively, it seems that choosing autonomously is not the same thing as choosing rightly or wisely or well.[55] And this intuition is underwritten by one of autonomy's central normative roles: if part of what makes paternalism presumptively wrong is that it conflicts with autonomy, it is hard to deny the possibility of autonomous bad choices; and if autonomous bad choices are

---

55 Cf. Raz, *The Morality of Freedom*, 412.

possible, autonomy cannot consist in appropriately exercising one's capacities for reasons-responsiveness. Moreover, the successful-exercise-of-normative-capacity interpretation of autonomy seems to imply that only the virtuous are really autonomous. If one wants a view of autonomy that squares with standard liberal commitments, the pure-capacity interpretation looks far more promising. But suppose the pure-capacity interpretation of autonomy is right. Then it is no longer so clear why autonomy is the sort of thing we generally have reasons to promote.[56] Merely having *capacities* for reasons-responsiveness, after all, is not all that valuable; what *is* valuable is having the capacities *and exercising them well*, i.e., actually succeeding in responding to one's reasons. The normative-capacity theorist cannot have her cake and eat it too: if she wants an account that makes sense of *one* of autonomy's key normative roles—being a bar to paternalism—she has to give up on being able to account for its *other* normative role—being the sort of thing we have reasons to promote.

There are a variety of ways the normative-capacity theorist might respond. Here I want to focus on two. The first is to deny the objection's presupposition and to insist that mere normative capacity *is* valuable and worth promoting after all. How so? Put briefly, normative capacity constitutes persons as responsible for their lives in a deep and meaningful sense—and *that* is good. To be sure, being responsible need not always and invariably be good: maybe being responsible for very bad decisions can make someone's life go worse. However, this qualification is consistent with the general and prospective value of autonomy as a thing worthy of promotion. Think of it this way. Persons are responsible agents. It is because of this fact that persons are capable of accessing or realizing special forms of value. Not all value in life is conditioned by being responsible, but some of it is. When autonomy is diminished, so is the opportunity for realizing these special forms of value. Hence, autonomy has prospective value as a generic life asset because it gives one the chance to live the most valuable kind of life. In general, and considered apart from what people make of their freedom, the interests of people who are robbed of the opportunity of being autonomous, and hence responsible, for their choices and lives are set back. The suggestion, in other words, is that we think of normative capacity as an *opportunity good*. Opportunities, of course, can be misused and wasted. But we often have reasons to provide people with valuable opportunities all the same. This vindicates, at least in a general way, the claim that we have reasons to promote normative capacity.

The second response accepts the objection's presupposition. It agrees that merely having normative capacities is not valuable and, hence, not the sort of thing we generally have reasons to promote, but insists that this only shows that

we should reject a purely capacitarian account of autonomy. For the reasons I have given, this response may look unpromising. But I do not think this is necessarily so. Because I think we need to preserve the idea of autonomy as a bar to paternalism and hence the possibility of misguided autonomous choices, I think the response would be unpromising if it simply collapsed autonomy into a form of virtue without remainder. It need not, however, do that. T. H. Green distinguishes between two kinds of freedom: responsibility-entailing freedom and perfection-entailing freedom.[57] Putting this in terms of normative capacity, the former is the idea of having the ability to detect and pursue norms and values—reasons-responsiveness. This ability is plausibly at the basis of responsible agency, so the corresponding idea of freedom is a responsibility concept: one is responsible and can, in whatever way is suitable, be held responsible for one's choices. This ability can be had even if it lies dormant or is exercised poorly. The latter is the idea of realizing normative capacity—of actually successfully tracking relevant norms and values and then successfully *conforming* behavior accordingly. This ability is plausibly a kind of perfection of our rational natures, so the corresponding idea of freedom is a virtue concept: one realizes an important human excellence and merits approbation and esteem on that basis. Now one might plausibly hang on to *both* the responsibility concept and the virtue concept as essential to our thinking about autonomy. If that is right, then the second response can be put as follows: once we distinguish the responsibility concept from the virtue concept, we can see that what it is we have reasons to promote is the property corresponding to the virtue concept—perfection-entailing freedom. As long as we also accept that respecting responsibility-entailing freedom is important, we have not collapsed autonomy into virtue; instead, we have come to see that our thinking about autonomy is more complex than we might have initially thought.

I leave open which of these two responses is more compelling. Readers may find one or the other more compelling depending on how they think a normative-capacity account should be developed. For my part, I like them both. It seems to me that the first is partly right: there is *some* positive value to normative capacity as a generic opportunity good and this accounts for some of the reasons we have to promote autonomy. But it seems to me the second is partly right too: by itself, unexercised normative capacity is *not very* valuable; we also, and perhaps especially, have reasons to promote the fulfillment or appropriate exercise of normative capacity. This is not the place to develop a normative-capacity account in any detail, so I will just confine myself to the following brief and mostly suggestive remark. I think it is important to recognize different strands in our

---

57   Cf. Brink, *Perfectionism and the Common Good*, 81.

thinking about autonomy. Making use of the distinction introduced by Stephen Darwall between two forms of respect, we can recognize distinct normative statuses associated with our thinking about autonomy: recognition respect goes with being a responsible chooser, appraisal respect with how capacities are exercised.[58] Moreover, once we spell out autonomy's normative role, I think we will see that it commits us to both the capacity concept and the virtue concept, each one associated with a different status. A full and adequate specification of what we have reasons to promote will include the virtue notion; at the same time, we cannot dispense with the idea of negative autonomy rights attaching to responsible agency. The ambiguity between *mere capacity* and *fulfillment of capacity* is present in many of the normative-capacity accounts that have been offered in the literature. Once the ambiguity is noticed and the alternatives are clarified, there is an important intramural debate to be had about how best to develop such a view: Should we go for a pure capacity view or a pure virtue view? My sense is that we should go for neither: we should reject the choice as a false dichotomy and give up instead on the assumption that autonomy is a simple unitary thing. On closer inspection, I believe our thought about autonomy reveals two distinct ideals—the ideal of responsibility-entailing freedom and the ideal of perfection-entailing freedom—neither of which are dispensable. The challenge for such a hybrid account, of course, is to spell out the details in such a way that each strand of our thinking is part of a larger, complex whole, rather than merely two entirely separate and unrelated things. This is not the place to tackle that challenge. But I believe the hybrid option is worth investigating. For those who are suspicious of this option, believing that the virtue notion must be treated as a separate value and not confused with autonomy, but who nevertheless find formal views of autonomy wanting, I commend the pure capacity interpretation along with the first response above. If I'm right about the viability and attractiveness of the hybrid option, the pure capacity interpretation will turn out to be an impoverished conception of autonomy, though still superior to formal views.

Turn now to reasons of respect. Respecting a person's autonomy means at least two things. First, it means respecting the person's right to make self-regarding choices as she sees fit, including (perhaps up to some threshold) bad choices. This idea is, of course, closely associated with anti-paternalist norms in liberal societies. Second, it means honoring the person's perspective—their wishes, what matters to them, what they care about, and so on. The latter shows up, for exam-

58   Darwall, "Two Kinds of Respect."

ple, in what is required to treat someone of another religion with respect or (a bit more broadly) in respecting claims of conscience, whether religious or secular.[59]

As we have seen, formal views of autonomy are commonly taken to be in the business of specifying conditions for authenticity. Suppose they succeed at this. Then it seems they have the ingredients for vindicating the second manifestation of respect, i.e., the one having to do with respect for conscience. Say you must decide whether to give a blood transfusion to an unconscious Jehovah's Witness to keep her alive. You know for a fact that she would not want to be given the blood transfusion, even though her life depends on it. It is not obvious what you ought, all things considered, to do. Still, whatever the right thing to do is, it seems there are powerful reasons of respect speaking in favor of honoring her (counterfactual) wishes not to receive the transfusion. Contrast this with a case where you know the religious commitments are superficial or have been inculcated in a suspect way. Perhaps the person has only been flirting with the Jehovah's Witness community for a couple of weeks or she has been drugged, manipulated, brainwashed, or coerced into having the commitments she does. In this case, presumably the weight that should be given to respecting the person's wishes is less, if any should be given to them at all. The difference between these cases is not in the content of the patient's request—that is the same. Instead, it is to be found in something like the position of the request vis-à-vis the person's authentic self.[60] Insofar as formal views are equipped to give us a story about authenticity, then, they are in a position to give us a story about this crucial dimension of respect for persons: honoring (i.e., giving some weight to) their point of view.

But it is not clear that formal views have the ingredients for vindicating the first manifestation of respect, i.e., the one having to do with the strong anti-paternalist presumption. As I argued above, authenticity plausibly suffices for attributability-responsibility, but this is not the right kind of responsibility to make sense of strong anti-paternalist practices. The strong liberal anti-paternalist presumption would seem best justified by the assumption that persons are more than merely attributability responsible for their choices. There are often good reasons *in favor* of paternalism, even when paternalism is all-things-considered wrong. In particular, people's welfare matters greatly and any balanced assessment of the ethics of paternalism must recognize this side of the balance sheet. How could a person's foolish choices merit protection? How could the kind of freedom that would allow people to make potentially ruinous life choices be justified? The mere fact that a choice is authentically an agent's own would not

59   Maclure and Taylor, *Secularism and Freedom of Conscience.*
60   Enoch, "Hypothetical Consent and the Value(s) of Autonomy."

seem to suffice to give other persons presumptively decisive reasons to desist from paternalistically interfering with the choice. A more robust form of responsibility seems to be required to make sense of that. This more robust form of responsibility would be secured by the ability to appreciate and appropriately respond to normative features relevant to choice.

To be clear, my point is about the *presuppositions* behind our *general* stance. I am not suggesting that considerations about normative capacity always feature, or always ought to feature, or that they are the only or always the most important considerations in every particular case in which the anti-paternalist presumption holds up. Nor am I claiming that there is not *some* anti-paternalist support from mere attributability-responsibility: there are, as I argued in the last paragraph, *pro tanto* reasons to honor people's points of view in self-regarding matters and these plausibly contribute to the case against paternalism. Rather, I am claiming that the kind of robust anti-paternalist norms characteristic of liberal social morality would be difficult to justify unless people were responsible (in the sense of being on the hook) for their choices. And this is difficult to make sense of in the absence of relevant normative capacities. Normative capacity is the ability to register and appropriately respond to normative features relevant to choice. Thus, if normative capacity is not required for autonomous choice, as formal accounts must maintain, this means one's choice about x can be autonomous, independent of any sensitivity to the features in virtue of which x is choice-worthy. But this is surely puzzling. For how can one be robustly and meaningfully responsible for choosing x in the absence of capacities for tracking what is relevant to the question of whether one should choose x? The requirement that we honor people's perspective is not strong enough for quite general and robust anti-paternalist norms. Perhaps such considerations weigh in here and there in particular cases, but they cannot plausibly be thought to support a regime of vigorous anti-paternalist protections. Assuming that paternalism is (at least in part) presumptively wrong because it conflicts with people's sphere of "sovereignty," and assuming that autonomy-as-a-capacity is the ground of autonomy-as-a-right, we need to ask what view of autonomy capacities would be required to justify a robust sphere of self-sovereignty. Views of autonomy that include a normative-capacity condition would seem much better equipped than views that do not, to vindicate a robust sphere of self-sovereignty.

It is possible, of course, that people's *actual* normative capacities are often quite fragile. In that case, the idea that persons are normative agents, capable of tracking and responding to the normative features bearing on their choice, may be something of an idealization. One might maintain, as I suggested earlier, that the anti-paternalist norms are partly proleptic or educative, cultivating

the thing they appear to presuppose. Or one might maintain that, although it is something of an idealization, people are responsible often enough and, given the other contributing reasons against paternalism and perhaps secondary reasons against too closely tracking facts about normative competence, it is an acceptable idealization. My point is not to defend anti-paternalist norms but to clarify what we are plausibly committed to in accepting them. If someone thinks adults are *not* normatively competent *most* of the time, then it seems to me they should in principle be prepared to accept a much more invasive regime of paternalism than we tend to think appropriate in liberal social orders (even if in practice such a regime would be too difficult or too expensive or too unpopular or too abusive, etc.). For if adults are really not normatively competent most of the time, they will in this respect be a lot like children, and it will be difficult to see what principled objection would remain to treating them like children, except that there might be a variety of practical obstacles to doing so. If one thinks there are strong principled objections against paternalism—that the objections to it are not just incidental or technical—there is substantial pressure to also acknowledge that persons are, or can be, responsible agents. Even if our self-conception as responsible agents is slightly idealized, as long as it does not radically betray the facts about us, we can make sense of strong principled objections to paternalism. This self-conception of responsible agency—perhaps a mix of fact and ideal—is better captured by normative-capacity accounts than by formal accounts of personal autonomy.

In sum, autonomy is a recognizable value in liberal social orders, which prize self-direction and are committed to protecting a significant sphere for individual choice. At the heart of this social vision is the idea of persons as dignified choosers who must chart their own course through life.[61] This idea marks out two normative roles for the idea of autonomy. One is an agency ideal. All else equal, autonomy is a desirable agency characteristic. Another is a principle protecting the exercise of autonomous agency. The choices of an autonomous agent call for respect. These are familiar ideas. And my argument in this section has been that, on the face of it, the normative role played by our concept of autonomy fits much more naturally with normative-capacity accounts of personal autonomy than formal accounts.

It is only fair to acknowledge that the situation has seemed to many philosophers to be exactly the reverse.[62] Indeed, tension with perceived liberal commitments is a central source of resistance to normative-capacity accounts. For example, John Christman is a well-known defender of a formal view of autonomy, and

---

61   G. Dworkin, *The Theory and Practice of Autonomy*; Raz, *The Morality of Freedom*.

62   Thanks to an anonymous reviewer from *JESP* for pressing me on this.

he motivates the view in large part because of its coherence with what he takes to be the best interpretation of liberalism.[63] Normative-capacity accounts, argues Christman, are in tension with liberalism. They seem, among other things, to suggest a "sliding-scale" picture of anti-paternalist protections tailored to match the degree of decision-making competence and to invite state-sponsored perfectionist programs aimed at getting people's choices to align with the true and the good. These concerns are serious. While I cannot vindicate the compatibility of normative-capacity accounts with liberalism, I wish to indicate at least briefly why I think such accounts cannot be too quickly dismissed as illiberal.

As I have already suggested, the idea of responsible agency seems a crucial bulwark in any principled anti-paternalist case, and normative-capacity accounts seem better positioned than formal accounts to interpret what this sort of responsible agency comes to. Moreover, normative-capacity accounts can, and I think should, accept the idea of negative autonomy rights. Most of us think (unjustified) paternalism *wrongs* people. We operate with the assumption that people have a *right* to decide for themselves in certain matters and that paternalism constitutes a usurpation of their *rightful authority* to do so.[64] A negative autonomy right protects a person (within suitable limits) in the making of bad self-regarding choices. This is just another way of saying that competent adults have powerful claims against others not to be interfered with in self-regarding matters. Since the right attaches to the capacity rather than its exercise, the right need not be thought of as conditional on making good choices.

One might, however, worry that even with negative autonomy rights in place, fidelity to the underlying normative structure would push normative-capacity accounts toward three unpalatable conclusions: (i) significant scope-restrictions on who enjoys negative autonomy rights, (ii) variations in autonomy levels and therefore different autonomy rights for different competent individuals, and (iii) an invitation to make minute discriminations among persons concerning their normative competence. But this is not necessarily so.

First, we should distinguish between scalar and threshold assessments of autonomy. On a plausible view, negative autonomy rights attach to threshold normative competence. It is a further question where to set the threshold, but there is no reason to suppose normative-capacity accounts are committed to setting it particularly high.[65] The lower the threshold is set, the less revisionary the account will be vis-à-vis standard liberal practice.

---

63   Christman, "Constructing the Inner Citadel," "Liberalism and Individual Positive Freedom,"
     "Procedural Autonomy and Liberal Legitimacy," and *The Politics of Persons*.

64   Darwall, "The Value of Autonomy and Autonomy of the Will," 267–68.

65   Cf. Griffin, *On Human Rights*, 156; Kauppinen, "The Social Dimension of Autonomy," 297.

Second, variation above the threshold does not necessarily yield differential allocation of rights. To be sure, there is an important question about how to resist this conclusion. But the problem is more general and is familiar from discussions of equality. Egalitarians are committed to ignoring variation above some threshold, treating persons as equals even when they exhibit morally relevant properties to different degrees.[66] Hence, the problem is no worse for normative-capacity theorists of personal autonomy than it is for egalitarians in general.

Third, it is open to normative-capacity theorists to say that above some threshold of competence, treating persons equally requires what Ian Carter calls "opacity respect," that is, in such a way as to not make fine-grained distinctions about their normative competence.[67] If so, then there would be moral reasons to desist from too closely tracking or using information about normative competence above the threshold, at least by certain agencies and within specified contexts (e.g., the state in relation to its citizens).

Fourth, a normative-capacity account is consistent with thin procedurally defined operative standards in different domains. For example, one might think that a normative-capacity account would demand a stingy approach to medical consent, e.g., in who gets deemed "capacitous." But this is not obvious. There may be good secondary practical and moral reasons for the existing standards, whether or not they adequately track normative competence. A variety of considerations—evaluative disagreement, proneness to error, liability to abuse, practical serviceability, and so on—speak in favor of thin procedural-looking proxy measures for normative competence that may, in practice, be overinclusive from the point of view of genuine normative competence. Since the pressures of crafting realistic and well-justified policy may license and even require a departure from attempting to use genuine normative competence as criterial for the determination of negative autonomy rights in various setting, we must be cautious about attributing to normative-capacity views pro-paternalist or illiberal policy implications in practice.

In short, normative-capacity accounts need not be wildly revisionary vis-à-vis widely accepted liberal views about equality, rights, and respect. But do they commit us to perfectionist politics? The answer, I think, is that they do not. It is one question what autonomy is; it is a further question how autonomy is to be promoted—and by whom. Even if there are quite general agent-neutral reasons to promote anyone's autonomy, there may be good secondary reasons for insisting that it is not everyone's business to promote everyone else's autonomy, and in

66 Cf. Waldron's discussion of equality in terms of the idea (originally from Rawls) of a "range property" (*One Another's Equals*, 84–127).

67 Carter, "Respect and the Basis of Equality."

particular, there may be special reasons to insist that states not be in the business of promoting autonomy. Whether the state may promote its citizens' autonomy is an important question, but it is orthogonal to the nature of autonomy. To see this, notice that it arises whether one adopts a normative-capacity view or a formal view. Suppose, for example, that autonomy is, as Gerald Dworkin maintains, "a second-order capacity of persons to reflect critically upon their first-order preferences, desires, wishes, and so forth and the capacity to accept or attempt to change these in light of higher-order preferences and values."[68] This type of view, too, might be combined with either perfectionist or anti-perfectionist political commitments. There is nothing about formal views of autonomy that commits those who hold such views to say to states or other agencies in authority, "Hands off on promoting this property!" A formal theorist might welcome state intervention in promoting autonomy, e.g., by promoting critical reflection. Conversely, there is nothing about normative-capacity accounts that commits those who hold such views to say, "This property may (or should) be promoted by the state." Whether one has an inviting posture to state intervention is orthogonal to which view one takes about the nature of autonomy. The debate between liberal perfectionists and liberal anti-perfectionists—interesting and important though it is—should not drive our theorizing about personal autonomy.

Suppose, however, that normative-capacity accounts do invite perfectionism in politics. Would this really be damning news? I think that is far from obvious. Some philosophers take it as virtually axiomatic for an account of autonomy of a liberal bent that it must respect neutrality and safeguard anti-perfectionism in politics.[69] But the question of how best to interpret the requirement of state neutrality is notoriously complex and controversial. Proponents of formal theories all too often simply take for granted that liberalism favors their view. Yet liberalism is a broad camp. There are sensible forms of perfectionist liberalism that have as good a claim as Rawlsian justificatory liberalism to being *bona fide* versions of liberalism.[70] To suggest that *all* substantive accounts of autonomy are illiberal will not work: normative-capacity accounts, as I have suggested, need not have radically illiberal implications—and they do a good job of interpreting the picture of responsible agency that seems to be presupposed by our liberal

---

68 G. Dworkin, *The Theory and Practice of Autonomy*, 20.

69 Cf. Christman, *The Politics of Persons*; G. Dworkin, *The Theory and Practice of Autonomy*; Westlund, "Rethinking Relational Autonomy."

70 E.g., Green, *Lectures on the Principles of Political Obligation and Other Writings*; Hurka, *Perfectionism*; Mill, *On Liberty*; Raz, *The Morality of Freedom*; Sher, *Beyond Neutrality*; and Wall, *Liberalism, Perfectionism, and Restraint*. On Mill as a perfectionist liberal, see Brink, *Mill's Progressive Principles*.

anti-paternalist practice. There is work to be done interpreting liberalism. At the very least, I think those who leverage anti-perfectionist arguments in favor of formal accounts of autonomy have more work to do in showing why we should antecedently favor non-perfectionist over perfectionist liberalism. And even if they make this case convincingly, it does not, so far as I can see, follow that autonomy is best interpreted in formal terms. For as I suggested in the last paragraph, what autonomy is and who gets to promote it are separate questions.

I conclude that normative-capacity accounts need not necessarily conflict with liberal commitments. Much more, of course, would need to be said to allay fears that normative-capacity accounts commit us to unattractive views of politics. What I have tried to show here, at least in brief outline, is that the conflict between liberalism and at least one variant of a substantive view of autonomy may not be as sharp or deep as sometimes supposed.

5. CONCLUSION

I have argued that formal views of autonomy face serious challenges. In particular, I have argued that they do not give us the right building blocks to make sense of the kind of responsibility that is plausibly at stake in autonomy, that they posit a fact/value asymmetry that creates an explanatory burden, and that they supply rather meager resources for helping us make sense of autonomy's normative role. Moreover, I have suggested that normative-capacity accounts need not be on a collision course with liberal commitments. None of this is decisive. Perhaps formal accounts can be developed in such a way that they meet the *prima facie* challenges I have sketched. And perhaps normative-capacity accounts are, after all, in significant tension with liberal commitments—or face other insurmountable problems. My aim in this paper has not been to give a full-scale assessment of competing theories, but to offer some reasons for rethinking the status quo. The choice between formal and normative-capacity accounts of personal autonomy is deep and consequential. It represents a quite fundamental fork in the road that any theorist of personal autonomy must face. Many philosophers have bounded down the formal path, thinking it would take them in the right direction. But if the arguments in this paper are along the right lines, it may be time to revisit the fork in the road and consider going the other way.[71]

University of Graz
jonathan.knutzen@gmail.com

REFERENCES

Arneson, Richard J. "Mill versus Paternalism." *Ethics* 90, no. 4 (July 1980): 470–89.

Arpaly, Nomy. "Responsibility, Applied Ethics, and Complex Autonomy Theories." In Taylor, *Personal Autonomy*, 162–81.

———. *Unprincipled Virtue: An Inquiry Into Moral Agency*. Oxford: Oxford University Press, 2003.

Benn, S. I. "Freedom, Autonomy and the Concept of a Person." *Proceedings of the Aristotelian Society* 76 (1975–76): 109–30.

Benson, Paul. "Feminist Intuitions and the Normative Substance of Autonomy." In Taylor, *Personal Autonomy*, 124–42.

———. "Feminist Second Thoughts about Free Agency." *Hypatia* 5, no. 3 (September 1990): 47–64.

———. "Free Agency and Self-Worth." *Journal of Philosophy* 91, no. 12 (December 1994): 650–68.

———. "Freedom and Value." *Journal of Philosophy* 84, no. 9 (September 1987): 465–86.

Berofsky, Bernard. *Liberation from Self: A Theory of Personal Autonomy*. Cambridge: Cambridge University Press, 1995.

Betts, Alexander, and Paul Collier. *Refuge: Rethinking Refugee Policy in a Changing World*. Oxford: Oxford University Press, 2017.

Bratman, Michael E. "Intention, Practical Rationality, and Self-Governance." *Ethics* 119, no. 3 (April 2009): 411–43.

Brink, David O. *Mill's Progressive Principles*. Oxford: Oxford University Press, 2013.

———. *Perfectionism and the Common Good: Themes in the Philosophy of T. H. Green*. Oxford: Clarendon Press, 2003.

Brink, David O., and Dana K. Nelkin. "Fairness and the Architecture of Responsibility." In *Oxford Studies in Agency and Responsibility*, vol. 1, edited by David Shoemaker, 284–314. Oxford: Oxford University Press, 2013.

Buss, Sarah. "Autonomous Action: Self-Determination in the Passive Mode." *Ethics* 122, no. 4 (July 2012): 647–91.

Carter, Ian. "Respect and the Basis of Equality." *Ethics* 121, no. 3 (April 2011): 538–71.

Christman, John. "Autonomy in Moral and Political Philosophy." *Stanford Encyclopedia of Philosophy* (Spring 2018). https://plato.stanford.edu/archives/spr2018/entries/autonomy-moral/.

———. "Autonomy and Personal History." *Canadian Journal of Philosophy* 21, no. 1 (March 1991): 1–24.

———. "Constructing the Inner Citadel: Recent Work on the Concept of Autonomy." *Ethics* 99, no. 1 (October 1988): 109–24.

———. "Liberalism and Individual Positive Freedom." *Ethics* 101, no. 2 (January 1991): 343–59.

——— . *The Politics of Persons: Individual Autonomy and Socio-Historical Selves.* Cambridge: Cambridge University Press, 2009.

———. "Procedural Autonomy and Liberal Legitimacy." In Taylor, *Personal Autonomy*, 277–98.

Christman, John, and Joel Anderson. "Introduction." In *Autonomy and the Challenges to Liberalism: New Essays*, 1–26. Cambridge: Cambridge University Press, 2005.

Darwall, Stephen. "Two Kinds of Respect." *Ethics* 88, no. 1 (October 1977): 36–49.

———. "The Value of Autonomy and Autonomy of the Will." *Ethics* 116, no. 2 (January 2006), 263–84.

Dworkin, Gerald. *The Theory and Practice of Autonomy*. Cambridge: Cambridge University Press, 1988.

Dworkin, Ronald. *Life's Dominion: An Argument about Abortion, Euthanasia, and Individual Freedom.* New York: Vintage Press, 1993.

Ekstrom, Laura Waddell. "Autonomy and Personal Integration." In Taylor, *Personal Autonomy*, 143–61.

Enoch, David. "Hypothetical Consent and the Value(s) of Autonomy." *Ethics* 128, no. 1 (October 2017): 6–36.

Feinberg, Joel. *The Moral Limits of the Criminal Law*. Vol. 3, *Harm to Self.* Oxford: Oxford University Press, 1986.

Fischer, John Martin, and Mark Ravizza. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press, 1998.

Frankfurt, Harry G. "Autonomy, Necessity, and Love." In *Necessity, Volition, and Love*, 129–41. Cambridge: Cambridge University Press, 1999.

———. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy* 68, no. 1 (January 1971): 829–39.

Friedman, Marilyn. *Autonomy, Gender, Politics*. Oxford: Oxford University Press, 2003.

Green, T. H. *Lectures on the Principles of Political Obligation and Other Writings*, edited by Paul Harris and John Morrow. Cambridge: Cambridge University Press, 1986.

Griffin, James. *On Human Rights*. Oxford: Oxford University Press, 2008.

Groll, Daniel. "Paternalism, Respect, and the Will." *Ethics* 122, no. 4 (July 2012): 692–720.

Hurka, Thomas. *Perfectionism*. Oxford: Oxford University Press, 1993.

———. "Why Value Autonomy?" *Social Theory and Practice* 13, no. 3 (Fall 1987): 361–82.

Kauppinen, Antti. "The Social Dimension of Autonomy." In *Axel Honneth: Critical Essays*, edited by Danielle Petherbridge, 255–302. Leiden, Netherlands: Brill, 2011.

Killmister, Suzy. "Autonomy and False Beliefs." *Philosophical Studies* 164, no. 2 (June 2013), 513–31.

———. *Taking the Measure of Autonomy: A Four-Dimensional Theory of Self-Governance*. New York: Routledge, 2018.

Mackenzie, Catriona. "Three Dimensions of Autonomy: A Relational Analysis." In *Autonomy, Oppression, and Gender*, edited by Andrea Veltman and Mark Piper, 15–41. Oxford: Oxford University Press, 2014.

Mackenzie, Catriona, and Natalie Stoljar. "Introduction: Autonomy Refigured." In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, 3–34. Oxford: Oxford University Press, 2000.

Maclure, Jocelyn, and Charles Taylor. *Secularism and Freedom of Conscience*. Cambridge, MA: Harvard University Press, 2011.

McDowell, John. "Autonomy and Its Burdens." *Harvard Review of Philosophy* 17, no. 1 (2010): 4–15.

McKenna, Michael. "The Relationship Between Autonomous and Morally Responsible Agency." In Taylor, *Personal Autonomy*, 205–34.

Mele, Alfred. *Autonomous Agents: From Self Control to Autonomy*. Oxford: Oxford University Press, 1995.

Meyers, Diana Tietjens. *Being Yourself: Essays on Identity, Action, and Social Life*. Lanham, MA: Rowman and Littlefield, 2004.

———. "Decentralizing Autonomy: Five Faces of Selfhood." In *Autonomy and the Challenges to Liberalism: New Essays*, 27–55. Cambridge: Cambridge University Press, 2005.

Mill, John Stuart. *On Liberty.* Edited by George Kateb and David Bromwich. New Haven: Yale University Press, 2003.

Nelkin, Dana Kay. *Making Sense of Freedom and Responsibility*. Oxford: Oxford University Press, 2011.

Nussbaum, Martha C. *Creating Capabilities: The Human Development Approach*. Cambridge, MA: Harvard University Press, 2011.

Oshana, Marina A. L. "Personal Autonomy and Society." *Journal of Social Philosophy* 29, no. 1 (March 1998): 81–102.

———. *Personal Autonomy in Society*. New York: Routledge, 2006.

Raz, Joseph. *The Morality of Freedom*. Oxford: Oxford University Press, 1986.

Rosen, Gideon. "Culpability and Ignorance." *Proceedings of the Aristotelian Society* 103, no. 1 (2003): 61–84.

Savulescu, Julian. "Rational Non-interventional Paternalism: Why Doctors Ought to Make Judgments of What Is Best for Their Patients." *Journal of Medical Ethics* 21, no. 6 (1995): 327–31.

Sayre-McCord, Geoffrey, and Michael Smith. "Desires . . . and Beliefs . . . of One's Own." In *Rational and Social Agency: The Philosophy of Michael Bratman*, edited by Manuel Vargas and Gideon Yaffe, 129–51. Oxford: Oxford University Press, 2014.

Sen, Amartya. *Development as Freedom.* Oxford: Oxford University Press, 1999.

Sher, George. *Beyond Neutrality: Perfectionism and Politics*. Cambridge: Cambridge University Press, 1997.

Stoljar, Natalie. "Autonomy and the Feminist Intuition." In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edited by Catriona Mackenzie and Natalie Stoljar, 94–111. Oxford: Oxford University Press, 2000.

———. "Feminist Perspectives on Autonomy." *Stanford Encyclopedia of Philosophy* (Winter 2018). https://plato.stanford.edu/archives/win2018/entries/feminism-autonomy/.

Taylor, James Stacey, ed. *Personal Autonomy: New Essays on Personal Autonomy and Its Role in Contemporary Moral Philosophy*. Cambridge: Cambridge University Press, 2005.

Terlazzo, Rosa. "Conceptualizing Adaptive Preferences Respectfully: An Indirectly Substantive Account." *Journal of Political Philosophy* 24, no. 2 (June 2016): 206–26.

Vargas, Manuel. *Building Better Beings: A Theory of Moral Responsibility*. Oxford: Oxford University Press, 2013.

Waldron, Jeremy. *One Another's Equals: The Basis of Human Equality*. Cambridge, MA: Harvard University Press, 2017.

Wall, Steven. *Liberalism, Perfectionism, and Restraint*. Cambridge: Cambridge University Press, 1998.

Watson, Gary. "Two Faces of Responsibility." *Philosophical Topics* 24, no. 2 (Fall 1996): 227–48.

Westlund, Andrea C. "Rethinking Relational Autonomy." *Hypatia* 24, no. 4 (Fall 2009): 26–49.

Wolf, Susan. *Freedom within Reason*. Oxford: Oxford University Press, 1994.

WHAT NORMATIVITY CANNOT BE

Matthew S. Bedke

OVER THE COURSE of many years Derek Parfit argued that reducing the normative to the natural is "conceptually excluded." His most recent incarnation of this argument is the *Normativity Objection*, where he moves from the conceptual impossibility of identifying rivers with sonnets, or heat with a shade of blue, to the conclusion that all normative-natural property identities are conceptually impossible.[1] This objection to reductive naturalism has been subject to heated debate. On the one hand, many philosophers inclined toward nonreductive normative realism are sympathetic to this view or something very similar. David Enoch, for example, thinks normative facts are "just too different from naturalist, not-obviously-normative facts" to be reducible or identical to them.[2] On the other hand, plenty of philosophers remain unconvinced, such as Patrick Fleming, who has recently concluded that the Normativity Objection has "no argumentative force against reductionism."[3]

Here, I hope to provide some guidance to the perplexed. Section 1 briefly rehearses the main problems with the argument as Parfit articulates it. Section 2 considers and criticizes a recent attempt to improve the argument by Nathan Howard and Nicholas Laskowski.[4] And sections 3 and 4 suggest and critically evaluate an improved argument. As we shall see, my suggestion relies on the highly controversial claim that normative cognition is *transparent* in the following sense: normative concepts reveal the nature of the properties they are about. I think this is the best way forward for those who wish to conceptually exclude normative-natural reductions.

---

1   Parfit, *On What Matters*, 2:324–27 and 3:72–84.

2   Enoch, *Taking Morality Seriously*, 80, and see 104–9; see also FitzPatrick, "Robust Ethical Realism, Non-Naturalism, and Normativity" and "Skepticism about Naturalizing Normativity"; Dancy, "Nonnaturalism"; Scanlon, *Being Realistic about Reasons*, 46.

3   Fleming, "The Normativity Objection to Normative Reduction," 421; see also Copp, "Normativity and Reasons"; Van Roojen, Review of *On What Matters*, vol. 3.

4   Howard and Laskowski, "The World Is Not Enough."

1. PARFIT'S NORMATIVITY OBJECTION

The version of naturalist realism criticized by Parfit targets property identities, such as the identification of rightness with maximizing happiness, or the (normative) reason relation with something about promoting desire satisfaction. One could argue against these particular identities, of course, but the Normativity Objection is not aimed at the merits of any particular proposal or set of proposals. It is aimed at the reductive gambit *tout court*. Moreover, it does not merely say that no normative property *is* the same as any natural property. It says that no normative property *could be* the same as any natural property. Last, it adds that the modality here is conceptual—normative-natural identities are conceptually excluded.

The main support for this contention comes from other cases where identities seem to be conceptually excluded. Parfit maintains that, just as rivers could not be sonnets, and heat could not have turned out to be a shade of blue, no normative property or fact could be some natural property or fact.[5]

The argument is similar to G. E. Moore's infamous open question argument. Arguably, the "open feel" that accompanies questions like "x is [natural predicate], but is x [normative predicate]?" reflect conceptual gaps. Application of a natural predicate just does not conceptually entail application of a normative predicate. Hence the open feel. Moore then thought that this conceptual gap entailed a metaphysical gap—that the property ascribed by the natural predicate was not the same as the property ascribed by the normative predicate. And that is the Achilles heel of the argument. At most the open feels reveal that our concepts alone do not fix any normative-natural identities. Granted, they do not *entail* identities, but that is a far cry from *excluding* them.

Parfit's Normativity Objection succumbs to a similar criticism. There might be a conceptual gap between application of natural predicates and application of any normative predicate, even an unbridgeable one. However, all this shows is that concepts alone *fail to fix* normative-natural identities. It does not show that concepts alone *rule out* all normative-natural identities.

Parfit's analogies do not show otherwise. We might grant that rivers could not be sonnets, or heat could not be a certain shade of blue, as a conceptual matter. And we might grant that *some* normative-natural property identities are conceptually excluded—as Parfit says, justice could not be the number 4 (for that would be a category mistake). But how are we to rule out *all* such identities? Maybe justice could not be the number 4, but why could it not be the same as an

5   Parfit, *On What Matters*, 2:324–27 and 3:72–84.

extensionally equivalent natural property?[6] This, in brief, is why the Normativity Objection fails to convince.

It is tempting to conclude that there is no good argument against normative-natural identities based on conceptual gaps. But the long history of meta-ethicists advancing such arguments, from Moore through Parfit and Enoch, suggests a more charitable reaction—that we have yet to uncover the best version of this argument. I will consider and criticize one attempt to improve the argument in section 2 before offering my own improvement in section 3.

2. THE WORLD IS NOT ENOUGH

In a recent bid to improve the argument, Howard and Laskowski propose that we reformulate it as a dilemma.[7] On the first horn, the reductive naturalist embraces a theory of meaning for normative terms in tension with the view that some normative-natural property identities are excludable *a priori* and conceptually.[8] That is problematic because some identities do seem conceptually excluded. Alongside "justice is not the number 4" consider "rightness is not a yellow rose" and "rightness is not a rocket." These appear to be conceptual truths, or at least *a priori* knowable truths.[9] Yet many semantic and metasemantic theories popular among naturalists have a hard time explaining how. Howard and Laskowski point out that Kripke-style reference by baptism plus the Millian view that meaning is exhausted by reference fail to explain how truths like these could be conceptual or *a priori* knowable.[10] After all, for such a view the cognitive sig-

---

6 Cf. Copp, "Normativity and Reasons," 46–47; Railton, "Two Sides of the Meta-Ethical Mountain?" 54–58; and Streumer, *Unbelievable Errors*, ch. 2. In what follows, property talk is to be robustly construed. There is a weaker claim that normative properties minimally construed (i.e., without metaphysical commitment) are not identical to natural properties minimally construed. But that claim is much less interesting, for it is too close to the uncontroversial claim that normative concepts are not to be identified with natural concepts (without regard to whether they ascribe the same robust properties). That said, Parfit might be happy with the weaker claim. As I read the exchange between Parfit and Railton in volume 3 of *On What Matters*, it appears that Parfit is not only willing to accept the weaker claim, but also willing to grant that, in terms of the robust metaphysics, normative properties and facts can be identified with certain natural properties and facts. These issues are vexed by some puzzling claims and distinctions Parfit draws concerning properties in pleonastic, description-fitting, and necessary co-extensionality senses.

7 Howard and Laskowski, "The World Is Not Enough."

8 Howard and Laskowski speak of analytic truths, but here I make no distinction between analytic truths and conceptual truths.

9 Cf. Copp, "Normativity and Reasons," 46–49.

10 Howard and Laskowski, "The World Is Not Enough."

nificance of a normative concept is not going to be part of its meaning; it will not help to fix reference, and it does not otherwise provide resources for purely conceptual or otherwise *a priori* access to what is excluded from the eligible referents of the normative term.

Now, one could quibble with the alleged conceptual or *a priori* status of these claims. I will not. So I take it that the lesson of the first horn is this: normative naturalists need to have a semantic (or metasemantic) theory whereby *some* normative-natural property identities are conceptually excluded, even if we cannot rule out *all* such identities. This is a nice point. At the very least, we need a semantics (or metasemantics) whereby we can conceptually exclude category mistakes, as this seems to be what is amiss with identifying rightness with roses or rockets.

On the second horn of the dilemma, reductive naturalists embrace a semantic (or metasemantic) theory that is consistent with conceptually excluding *some* normative-natural identities. Howard and Laskowski illustrate one such possibility: neo-descriptivism. For example, it could be that "rightness" is associated with a description such as *the actions of the action-type that a maximally informed observer would desire to perform*, which would constrain the causal chains relevant to reference fixation. In turn, such a description would also make the truth of the sentence "rightness is not a yellow rose" knowable *a priori*, in part because it is plausibly *a priori* that a yellow rose is not an action type.[11]

Once reductive naturalists go down this road they can grant that some *but not all* normative-natural identities are conceptually excluded. That sounds like the right thing to say. So what is the remaining problem with this horn of the dilemma?

Howard and Laskowski think the problem appears when we turn our attention to fundamental normative principles. They ask us to suppose that the Principle of Utility is a fundamental normative principle, and then they ask this question: Is this principle true because of (or in virtue of) some natural fact? *A priori*, it seems not. Consider:

2a. That it's Monday does not partly explain why we ought morally to do what would maximize the balance of pleasure over pain.
2b. The Principle of Utility is true not even partly in virtue of the fact that it's Monday.
2c. The Principle of Utility is not even partly grounded in the fact that it's Monday.

---

11 Howard and Laskowski, "The World Is Not Enough."

2d. The Principle of Utility is not even true partly in virtue of the fact that we're in Chicago.

2e. The Principle of Utility is not even partly grounded in the fact that we're in Chicago.[12]

Howard and Laskowski argue that these are conceptual truths knowable *a priori* (assuming that the Principle of Utility is a true moral principle). Granted, these are just examples, and grounding morality in facts about days of the week or cities is really implausible. But, like Moore's open question argument, we are supposed to see that *any* natural fact, or at least any spatiotemporal fact, can be plugged into these grounding claims to yield the same result: grounding in such facts is conceptually excluded.

Next, they borrow some ideas from Kit Fine to argue that moral/normative principles are *unworldly*.[13] For Fine, to be an unworldly truth is to be a truth that is true *regardless* of how things are at any world. This is contrasted with truths that are made true by how things are at a world or at some worlds. Unworldly truths are also to be distinguished from necessary truths. Consider "2 + 2 = 4" and "Socrates exists or Socrates does not exist." Both are true at every world, but for Fine the first sentence is true *regardless* of how things are at any world—for each world, nothing there makes the mathematical claim true—whereas the second sentence is true *whatever* the circumstances—each and every world makes the disjunction true, some by making the first disjunct true, others by making the second disjunct true.

Tapping into this theoretical apparatus, the idea is that the true, fundamental moral principles are unworldly. Howard and Laskowski conclude:

> Moral principles have unworldliness as part of their meaning exactly in the way that "seven is prime" does. Put another way, sentences expressing truths about moral principles have *unworldliness* as part of their content. This is why, on a brand of non-reductivism that takes Fine as inspiration, 2a–e are knowable analytically or on the basis of our competence with normative concepts alone. This is also why the impossibility of at least one highly intuitive version of naturalism is knowable on the basis of our competence with normative concepts alone.[14]

This argument goes by pretty quickly. But it is clear that the conclusion is that grounding fundamental normative principles in the natural (or at least the spa-

---

12 Howard and Laskowski, "The World Is Not Enough."

13 Fine, *Modality and Tense*.

14 Howard and Laskowski, "The World Is Not Enough."

tiotemporal) is conceptually excluded. And they take this to vindicate the spirit of Parfit's Normativity Objection.

I am not convinced. First, a minor concern: 2a–2e help to show that the normative principles are not grounded in facts about Monday or Chicago. As already mentioned, these particular grounding claims are not very plausible. Of course, we can fix the claims so that they feature more plausible natural grounds, like facts about pain and pleasure. It might be harder to argue that these grounding claims will also be conceptually excluded, but let me grant conceptual exclusion for the sake of argument.

A more significant concern is that 2a–2e and the surrounding argument only address one way of grounding normative principles in the natural, namely, grounding fundamental moral principles *directly* in particular natural facts. They do not address *indirect* grounding in natural facts. Let me clarify this distinction in the following way. One grounding option—the one Howard and Laskowski seem to have in mind—has it that fundamental normative principles are metaphysically prior to particular normative facts, where the principles help to ground particular normative facts. On this picture, what *makes* an executioner's action wrong, for example, is the Principle of Utility in combination with certain natural facts about the act of execution. If we then ask if the Principle of Utility is itself grounded, it seems plausible to say, no, it is not grounded and *a fortiori* it is not directly grounded in natural facts.

An alternative grounding option has it that particular normative facts are metaphysically prior to the fundamental normative principles, where the principles are simply systematizing and explicitly stating the patterns we find in the particular. On this picture, if the particular normative facts are to be grounded, they are to be grounded in natural facts, so that the executioner's action is wrong, for example, fully because of certain natural features it has. And what would then make the Principle of Utility true? The fact that this execution is wrong, that denying Sally the right to vote is wrong, that donating to an ineffective charity is wrong, and all the other particular normative facts about wrongness that are best systematized (allegedly) by the Principle of Utility. On this view, normative principles are ultimately grounded in the natural, albeit indirectly, by being grounded in particular normative facts that are themselves more directly grounded in the natural.

The examples and arguments of Howard and Laskowski only address the direct grounding of fundamental normative principles in natural facts. So even if they successfully argue that the principles are not so grounded, and this is *a priori* knowable, such a result is consistent with the possibility that the principles are *indirectly* grounded in the natural, and I take it that they would count

this as a form of naturalism. Moreover, when we do consider whether particular normative facts are grounded in natural facts, like whether the wrongness of the executioner's action is grounded in natural facts about the action, even fully grounded, the negation of such claims does not seem to be a conceptual truth and knowable *a priori*. Such grounding claims are contested in the literature and it would be surprising if those on one side of the debates were making a conceptual mistake. In addition, the particular-normative-facts-first option is consistent with *a priori* knowledge of fundamental normative principles, like the Principle of Utility. For it could be that particular normative facts are knowable *a priori* from these natural facts (though probably not *qua* conceptual truths), and systematizing principles are knowable *a priori* from there (again, probably not *qua* conceptual truths). In fact, I think this epistemology is faithful to our normative practices, which are usually case-based and casuistic, not inferential from first principles. But the important point for now is that Howard and Laskowski have not ruled out *a priori* all the ways in which the normative can be grounded in the natural.

These concerns all assume that grounding in the natural suffices for naturalism. But this itself is highly questionable. Distinguish two kinds of reductive projects. On one project, *A* is reduced to *B* just in case *A* is grounded in *B*. Note that, on this notion of reduction, one could "reduce" *A* to *B* even if *A* is a distinct existence from *B*, for grounding could be a real (explanatory) relation between distinct existences. Another reductive project is to come up with type-type property identities (or fact identities). On this second notion of reduction, if *A* is reduced to *B* (*A* is a property-type identical to *B*), *A* is not a distinct existence from *B*. I think naturalists worthy of the name should aim for the second sort of reduction. Why? Because even if the first project succeeds and normative properties are fully grounded in the natural, normative properties could still be distinct existences from natural properties, and what is more, they could have a *nature* that can only be articulated in normative terms. This possibility seems to me more clearly on the nonnatural side of the naturalism/nonnaturalism divide. After all, Mackie was not assuaged of his metaphysical concerns over nonnaturalism after admitting that there might be some sort of *because* relation between the moral properties and their natural subvenient base.

Note that Howard and Laskowski themselves characterize reduction as follows: "the metaphysical *nature* of morality—and of normativity, more generally—is … fully explicable in nonnormative terms."[15] I like this focus on the metaphysical *nature* of normativity, but it is not clear how grounding one property (or fact) in another distinct property (or fact) directly addresses the question of the

---

15 Howard and Laskowski, "The World Is Not Enough," emphasis added.

nature of the grounded property (or fact). To address that, Howard and Laskowski need to defend a controversial theory of grounding according to which the grounded cannot enjoy a nature/essence not enjoyed by its grounds.

This is not the place to provide an argument against those controversial theories of grounding. But it is worth noting that property or fact identities would *uncontroversially* address the nature of normativity and its naturalizability. An argument against natural to normative grounding, on the other hand, only questionably addresses the nature of normativity and its naturalizability.

Last, even if grounding the normative in the natural suffices for naturalism, arguments against such grounding do not suffice to rule out naturalism. For even if the normative is not grounded in the natural, the existence of normative-natural property identities—the most straightforward form of naturalism—is a live option. In a way, this is easy to see. For suppose that the property of being wrong is identical to the property of failing to maximize happiness, and suppose the fact that x is wrong and the fact that x fails to maximize happiness are the same fact. It is fairly uncontroversial that grounding is irreflexive, so this one fact would not be grounded in itself. We can even make it a conceptual truth that the fact that x is wrong is not grounded in the fact that x fails to maximize happiness. We just need two ways of conceiving of this fact that build in the lack of grounding. So here we have *a priori* knowledge that this wrongness fact is not grounded in a natural fact, combined with a normative-natural fact identity. This shows that those who wish to argue against naturalism should not rest content with an argument that the normative is not grounded in the natural. They must do what Parfit tried to do and rule out property identities.

Let me make a similar point using Fine's language: unworldly normative principles might nevertheless trade in properties that are type identical to worldly (natural) properties.[16] First, from Fine's own discussion, it is not entirely clear which sentences count as unworldly—as true *regardless* of the circumstances. In addition to mathematical truths, Fine talks about applications of *transcendental predicates* (he gives as an example, "Socrates is self-identical"), and certain substance sortals (he gives as an example, "Socrates is a man"). His animating metaphor is what is not under God's control as they go about creating a possible world.[17] Whether there is an intelligible grouping of unworldly truths to be found here and whether positing the group does important theoretical work is disputed.[18] But even if we grant for the sake of argument an intelligible and theoretically useful category of unworldy truths, and grant as *a priori*/conceptual

16 Fine, *Modality and Tense*.

17 Fine, *Modality and Tense*, 325.

18 See, e.g., Forbes, "Critical Notice of Kit Fine's *Modality and Tense*."

truths propositions to the effect that fundamental normative principles are not grounded in the natural, it is still a leap to infer the conceptual impossibility of normative-natural property identities.

Insofar as I get my head around unworldly truths, outside of normativity there seem to be many claims that are not grounded in natural facts (as a matter of conceptual necessity, knowable *a priori*), but that leave open interesting property identities. Consider:

1. That water is whatever shares the fundamental explanatory nature as the stuff around here that falls from the skies, fills lakes, etc., *is not even partly grounded in* [fill in natural world fact].[19]
2. That red is a color *is not even partly grounded in* [fill in natural world fact].
3. x is luminous iff x radiates light, and this *is not even partly grounded in* [fill in natural world fact].

The truths on the left-hand side of these grounding claims are not within God's control. It is not up to them to decide whether water is going to be the watery stuff, red is going to be a color, or being luminous is coextensive with radiating light. These seem to be true regardless of how things are at any world. And there are lots of ways to fill in the brackets such that we can know these claims *a priori* (as a matter of concepts?). We seem to be able to infer that these truths are not grounded in *any* natural facts. Nevertheless, 1 leaves open whether water is a substance identical to H_2O, 2 leaves open whether redness is to be property-type identified with some natural property (like a surface reflectance property), and 3 states an extension equivalence similar to the one stated by the Principle of Utility (x is right iff x maximizes net happiness) that clearly leaves open this identity: luminous = radiates light.[20] Generally, the conceptual exclusion of grounding claims like these leaves open genuine property identities.

And so it goes with Howard and Laskowski's argument. Even if our normative concepts rule out the possibility that the fundamental normative principles are grounded in the natural (directly or indirectly), this leaves open certain normative-natural property identities. As I put it earlier, even if grounding the normative in the natural *suffices for* naturalism, arguing against such grounding does

---

19 I do not think the use of "here" in the principle makes its truth worldly. The principle is still true regardless of how things are at each world. What is worldly is the truth that water is H_2O, but the principle is neutral on that question.

20 The property identity is taken from Parfit (*On What Matters*, 3:66). He also maintains that these properties are identical in the "description fitting" sense, though I admit it is not clear to me what "description fitting" means.

not suffice to reject naturalism. To strike at the heart of naturalism we need to address normative-natural property identities.

3. NORMATIVE TRANSPARENCY

Let me suggest the Argument from Normative Transparency.

1. Normative concepts have a distinctive cognitive significance, and they present the properties they ascribe as having a certain nature, one that can only be characterized in normative terms (e.g., in terms of inherent, authoritative guidance).
2. Natural properties are properties whose nature need not be characterized in normative terms.
3. So normative-natural identities, including property identities and fact identities, are conceptually excluded.

Premise 1 has three key ingredients: cognitive significance, mode of presentation, and nature revealing. First, it makes a claim about the cognitive significance of normative concepts. I take this to include their inferential and motivational roles as well as their distinctive modes of presentation, including the phenomenal qualities of occurent normative thoughts. And I take it that this cognitive significance is distinctive—it helps us identify a normative thought as a normative thought. This should be fairly uncontroversial, for it leaves open whether or not the concept is in the business of ascribing properties, and if it is in that business, it leaves unsettled what kind of property is ascribed.

Second, premise 1 says that normative concepts *present* their properties in a certain way. The concept of being a reason presents the property under the guise of inherent, authoritative guidance (or favoring), for example. This serves to further characterize at least some aspects of the cognitive significance of normative terms. It is not a full characterization of that significance, which likely includes certain inferential roles and perhaps links to motivation or intention. But it does characterize a rather salient and important part of cognitive significance—the mode of presentation as of inherent, authoritative guidance. Again, this should not be too controversial, for so far we are just focused on cognition. We have not yet said that there is inherent, authoritative guidance *in the world*, built into the property being ascribed. And even if this mode of presentation does lay down a condition on the properties being ascribed, being a property of "inherent, authoritative guidance" needs to be interpreted. It could turn out that what it is to be a property of inherent, authoritative guidance is to be a certain sort of natural

property, like promoting desire satisfaction, or being the basis of advice for an ideal advisor. So far, nothing excludes normative-natural property identities.

Hence premise 1 includes the third, stronger idea that normative concepts present the properties they ascribe as having a certain *nature*, one that can only be characterized in normative terms, such as in terms of inherent, authoritative guidance.[21] When we add this element, the mode of presentation as of inherent, authoritative guidance is no longer confined to cognition, as it were. It is representing something worldly—or, as I like to put it, it is acting as a transparent window onto something worldly—as having a nature only describable in the very terms we use to describe that distinctive mode of presentation in cognition—in terms of inherent, authoritative guidance. This is certainly controversial, and I discuss some of my reservations in section 4. Before I do, let me finish unpacking the argument.

Premise 2 then says that natural properties have a nature that need not be characterized in normative terms. The key idea here is similar to Jean Hampton's claim that normative authority violates the strictures of science because science rejects explanations that invoke Aristotelian final causes.[22] Hampton thinks that Aristotelian final causes posit that (a) certain places, states of affairs, or motions are fitting or right for certain objects, (b) the objects are able to detect this fittingness, and (c) the objects can respond to this fittingness. For Hampton, varieties of nonnaturalism countenance this type of explanation for some actions, whereas naturalist realism would appeal to only material, formal, or efficient causes.

To argue against normative naturalism, I do not think we need to talk about different sorts of explanations and we need not claim that relations of fittingness help to explain some actions. All we need is the idea that there are concepts that are about worldly properties whose nature can only be characterized in normative terms. That alone is inconsistent with naturalism, regardless of whether these properties also help explain action. For natural properties have no such nature. That is part of the point of distinguishing the natural from the nonnatural. Just as a divine being would not be a natural entity, any property whose nature is only describable in normative terms would not be a natural property.

In saying this, nonnaturalists need not have a full-blown theory of the natu-

21 Note that I do not offer the much less plausible claim that normative properties can only be ascribed with normative thought and language (well criticized by Eklund, *Choosing Normative Concepts*, 77). Nonnaturalists should grant that nonnormative concepts might ascribe normative properties (see Eklund's "thgir" example), but they should deny that such ascriptions reveal the nature of the properties ascribed. They should maintain that only normative thought and language can do that.

22 Hampton, *The Authority of Reason*, 111–14.

ral. Maybe the natural is all that can be completely characterized with structural, functional, or causal language. Any proposal here is bound to be controversial. Fortunately, all we need to note is one negative feature of the natural—natural properties do not have a nature that can only be characterized in normative terms. All sides to the debate should agree with that limited claim.

This combined with premise 1 yields Parfit's desired conclusion: normative-natural identities are excluded by normative concepts (plus a negative claim about the natural).[23]

4. EVALUATION OF THE ARGUMENT

I think the Argument from Normative Transparency is the best way—perhaps the only plausible way—to argue that normative-natural reductions are conceptually excluded. I submit it as an improvement on Parfit's Normativity Objection and it also might be the best way to develop other objections in this ballpark, like Enoch's "just too different" intuition, and Scanlon's worry that normative-natural identities destroy normativity altogether.[24] Moreover, I can see why some would find it plausible and even persuasive.

That said, I am not convinced that premise 1 is true. We should grant that normative concepts have a distinctive cognitive significance and even a distinctive mode of presentation aptly characterized in terms of inherent, authoritative guidance. It is much harder to see how this cognitive significance speaks to the nature of the properties allegedly ascribed. Most concepts do not do this. Most are not transparent windows onto the natures of their worldly contents. If normative concepts are, this cries out for argument and explanation. I think the best strategy for nonnaturalists is to draw our attention to the special mode of presentation enjoyed by normative concepts, including the distinctive phenomenal quality of occurent normative thoughts, and to argue that this mode of presentation does purport to represent the nature of their worldly referents. It could be that this presentation as of inherent, authoritative guidance *in cognition* uniquely captures the natures of some *worldly* properties or facts—properties or facts of inherent, authoritative guidance in-the-world, as it were.

There are examples outside of normativity where similar issues arise. Focus-

---

23 We could replace premise 2 with the stronger premise that our *concept of a natural property* precludes it from having a nature only describable in normative terms. Then from the concept of the normative and the concept of the natural alone we exclude normative naturalism. I will not weigh in on the merits of this stronger premise, and I will stick with premise 2 as stated.

24 Enoch, *Taking Morality Seriously*, 80, 104–9; Scanlon, *Being Realistic about Reasons*, 46.

ing on phenomenal and presentational qualities, some have seriously considered whether visual experience reveals the nature of the colors.[25] Using Mark Johnston's example, the idea is that the nature of canary yellow is revealed by experiences of canary yellow. It is there, laid bare before the mind when one experiences canary yellow. Putting a similar thesis in terms of concepts, Philip Goff has argued that

> in having a direct phenomenal concept, the token conscious state being attended to is directly presented to the concept user, in such a way that … the complete nature of the type to which it belongs is apparent to the concept user.[26]

And again:

> [direct] phenomenal concepts reveal the complete nature of the conscious states they refer to … *we know what pain is* through feeling pain.[27]

I think this is where nonnaturalists should look to develop the argument that normative-natural property identities are conceptually excluded. The key idea is that normative concepts (and particularly their distinctive mode of presentation) are transparent in that they reveal the nature of the properties they are about. Here are some proposals: we *know what the normative reason relation is* by having thoughts about normative reasons; we *know what wrongness is* by having thoughts involving the concept WRONG; etc. To be sure, such thoughts would not transparently reveal whether these properties are *instantiated* in any given case, but the suggestion is just that they reveal the nature of the property type in question. This is no doubt worthy of exploration, but there are difficulties ahead. Let me mention some of my reservations.

First, phenomenal colors might provide the *best case* of worldly natures that are transparent to cognition. Still, the transparency claim about color cognition and the colors is hotly contested. The most plausible version of it would say that what is transparent is the nature of a mental state, or some property of a mental state. But this is implausible in the case of normativity. The normative properties posited by nonnaturalists are not meant to be some subset of phenomenal properties, and more generally they are not mental states or properties thereof. Wrongness, for example, is meant to be a property of actions, which is something instantiated outside of one's mental life. Once you locate it in the world,

---

25 Campbell, "A Simple View of Colour," 178; Johnston, "How to Speak of the Colors," 138; Yablo, "Singling Out Properties," 480–90.

26 Goff, *Consciousness and Fundamental Reality*, 107.

27 Goff, *Consciousness and Fundamental Reality*, 124–25.

however, any analogy with phenomenal transparency is certainly weaker. That said, not all hope is lost. There is room to maintain that normative concepts reveal the nature of properties that are not properties of the mind, as some have considered whether phenomenal concepts reveal the nature of extra-mental properties in the world.[28]

However, and this is my second point, once we have this transparency thesis clearly formulated, we are faced with the epistemic problem of how we know whether it is true. In the case of normativity, one possibility is that being nature-revealing is itself somehow manifest in normative cognition, or otherwise evident enough, and in need of no further argument.[29] If that is right, one simply needs to carefully attend to certain aspects of normative cognition to establish both that it is about worldly properties and that it is transparently about the natures of those properties. Unfortunately, many philosophers have attentively explored their own normative cognition without finding it manifest that it reveals the nature of some worldly properties. Elizabeth Anscombe, for example, ridiculed what she called the "mesmeric force" that attends the emphatic ought.[30] Far from revealing the nature of normative properties in the world, she thought it was a holdover from a defunct conceptual scheme. Others have carefully considered various aspects of normative cognition and have come to the conclusion that such cognition is about perfectly natural properties, or that it is a projection of our sentiments. After paying careful attention to first-personal ought judgments, for example, Terry Horgan and Mark Timmons have concluded, "It is not introspectively accessible whether or not direct moral experiences carry ontological objective purport," where to have ontological objective purport is to "purport to be about some in-the-world moral properties."[31] If we cannot tell whether they carry objective ontological purport, we certainly cannot tell whether they purport to reveal the very nature of their worldly referents.

Perhaps Anscombe, Horgan and Timmons, and others are just wrong about what we can glean from careful attention to the normative mode of presentation. Perhaps.[32] More promising, I think, is to push a non-introspective argument for

---

28   Cf. Chalmers, "The Content and Epistemology of Phenomenal Belief," and his discussion of Edenic content.

29   Cf. Goff, *Consciousness and Fundamental Reality*, 108–9.

30   Anscombe, "Modern Moral Philosophy," 8.

31   Horgan and Timmons, "What Does Moral Phenomenology Tell Us about Moral Objectivity?"

32   I note that, though Anscombe and Horgan and Timmons do not think their normative cognition is a window onto *robust*, worldly normative reality that is *actually instantiated*, one might grant that it is a window into possible normative reality that is not instantiated, or maybe a minimal normative reality. See note 6 above on properties minimally construed.

the transparency of normative concepts. Perhaps the transparency thesis is part of the best explanation in answer to a question that has received too little attention, namely: Why do normative concepts have the mode of presentation that they have (that of authoritative guidance) rather than some other mode of presentation or perhaps none at all? Of course, it could be that modes of presentation are accidental features and inexplicable. But if one wants an explanation, one answer is that they have this mode of presentation because they are transparent windows onto the nature of their subject matter, and their subject matter has a nature as of authoritative guidance. (Similarly, one could say that the phenomenal concept of yellow has its distinctive mode of presentation because it is a window onto a property whose nature is reflected in the concept itself.)

This is an interesting proposal. And it is hard to see how naturalists can offer a better explanation.[33] If the naturalist is to avoid classifying the normative mode of presentation as inexplicable, the best move is to lean on certain metaphors that have cropped up in the literature, like the metaphors of projecting, gilding, staining, or coloring. There, the hope is that *somehow* our conative attitudes help to explain why normative cognition has the authoritative mode of presentation. But a metaphor is no substitute for an explanation. Ideally, we would be given a mechanism that shows just how some conative attitude combines with a concept so as to imbue that concept with the mode of presentation as of authoritative guidance. Without this extra step, the nonnaturalist suggestion that the normative mode of presentation is as it is because there is a worldly subject matter whose nature is reflected in normative concepts might just be the better explanation.[34]

5. CONCLUSION

This is not the place to fully prosecute the case. The Argument from Normative Transparency has been suggestive and exploratory. But it does strike me as an improvement on similar arguments. Parfit's own attempts to conceptually exclude normative-natural property identities fail in much the same way as Moore's open question argument fails, and relying on a "just too different" intuition seems too thin. Howard and Laskowski try to exclude a certain kind of naturalism—that of grounding fundamental normative principles in the natural—but they do not exclude indirect grounding, and they do not rule out property identities (the

---

33 But see Bedke, "Naturalism and Normative Cognition." For another attempt to explain away the "just too different" intuition, see Copp, "Just Too Different."

34 Again, for an alternative explanation consistent with a naturalist metaphysics, see Bedke, "Naturalism and Normative Cognition."

holy grail of naturalism). Whether the Argument from Normative Transparency will succeed where these others fall short is yet to be fully adjudicated. It has analogues in the philosophy of mind, but there are important disanalogies between normative and phenomenal concepts, and an epistemic defense needs to be worked out. But it offers a promising way forward for those who wish to argue that normative-natural reductions are conceptually excluded.[35]

University of British Columbia
matt.bedke@ubc.ca

REFERENCES

Anscombe, G. E. M. "Modern Moral Philosophy." *Philosophy* 33, no. 124 (January 1958): 1–19.
Bedke, Matthew S. "Naturalism and Normative Cognition." *Philosophical Studies* (forthcoming). Published ahead of print, February 4, 2020. https://doi .org/10.1007/s11098-020-01425-y.
Campbell, John. "A Simple View of Colour." In *Reality, Representation and Projection*, edited by John Haldane and Crispin Wright, 257–68. Oxford: Oxford University Press, 1993.
Chalmers, David J. "The Content and Epistemology of Phenomenal Belief." In *Consciousness: New Philosophical Perspectives*, edited by Quentin Smith and Aleksandar Jokic, 220–72. Oxford: Oxford University Press, 2003.
Copp, David. "Just Too Different: Normative Properties and Natural Properties." *Philosophical Studies* 177, no. 1 (January 2020): 263–86.
———. "Normativity and Reasons: Five Arguments from Parfit against Normative Naturalism." In *Ethical Naturalism: Current Debates*, edited by Susana Nuccitelli and Gary Seay, 24–57. Cambridge: Cambridge University Press, 2012.
Dancy, Jonathan. "Nonnaturalism." In *The Oxford Handbook of Ethical Theory*, edited by David Copp, 122–45. Oxford: Oxford University Press, 2006.
Enoch, David. *Taking Morality Seriously: A Defense of Robust Realism*. New York: Oxford University Press, 2011.
Eklund, Matti. *Choosing Normative Concepts*. Oxford: Oxford University Press, 2017.

Fine, Kit. *Modality and Tense: Philosophical Papers*. Oxford: Oxford University Press, 2005.

FitzPatrick, William. "Robust Ethical Realism, Non-Naturalism, and Normativity." In *Oxford Studies in Metaethics*, vol. 3, edited by Russ Shafer-Landau, 159–205. Oxford: Oxford University Press, 2008.

———. "Skepticism about Naturalizing Normativity: In Defense of Ethical Nonnaturalism." *Res Philosophica* 91, no. 4 (October 2014): 559–88.

Fleming, Patrick. "The Normativity Objection to Normative Reduction." *Acta Analytica* 30, no. 4 (December 2015): 419–27.

Forbes, Graeme. "Critical Notice of Kit Fine's *Modality and Tense: Philosophical Papers*." *Philosophical Review* 117, no. 2 (April 2008): 275–87.

Goff, Philip. *Consciousness and Fundamental Reality*. Oxford: Oxford University Press, 2017.

Hampton, Jean E. *The Authority of Reason*. Cambridge: Cambridge University Press, 1998.

Horgan, Terry, and Mark Timmons. "What Does Moral Phenomenology Tell Us about Moral Objectivity?" *Social Philosophy and Policy* 25, no. 1 (January 2008): 267–300.

Howard, Nathan Robert, and N. G. Laskowski. "The World Is Not Enough." *Noûs* (forthcoming). Published ahead of print, May 15, 2019. https://doi.org/10.1111/nous.12293.

Johnston, Mark. "How to Speak of the Colors." *Philosophical Studies* 68, no. 3 (December 1992): 221–63.

Parfit, Derek. *On What Matters*. 3 vols. Oxford: Oxford University Press, 2011–17.

Railton, Peter. "Two Sides of the Meta-Ethical Mountain?" In *Does Anything Really Matter?* edited by Peter Singer, 35–60. Oxford: Oxford University Press, 2017.

Scanlon, T. M. *Being Realistic about Reasons*. Oxford: Oxford University Press, 2014.

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

Van Roojen, Mark. Review of *On What Matters,* vol. 3, by Derek Parfit. *Notre Dame Philosophical Reviews*, October 21, 2017. https://ndpr.nd.edu/news/on-what-matters-volume-iii.

Yablo, Stephen. "Singling Out Properties." *Philosophical Perspectives* 9 (1995): 477–502.