Expressivism and Normative Metaphysics*

Billy Dunaway
University of Oxford
Draft of 11th August 2014

1 Preliminaries: Lewisian eliteness and metaphysical commitments

Begin with David Lewis’s notion of a perfectly natural property, or (to avoid terminological confusion at a later point) a perfectly elite, joint-cutting property. To a rough first approximation, these are the properties that are distinguished from other less elite properties by the theoretical roles that they play. Any two things which share a perfectly elite property will thereby objectively resemble each other to some degree; the perfectly elite properties provide a supervenience base on which all else supervenes, and the presence of perfectly elite properties distinguishes genuine laws from gerrymandered generalizations.

Because eliteness plays these roles, differences over perfect eliteness are not metaphysically inconsequential. Take two theories which agree entirely over what exists—say, both theories hold that properties A, B, and C exist. In Quinean jargon, these theories agree over ontology. But suppose one theory holds that each of the properties A, B, and C are perfectly elite, while the other theory holds that only A and B are. These theories differ, intuitively, over an important point of metaphysics, as the former takes on additional commitment by calling C perfectly elite.¹

The picture described above is a familiar and contentious one from the metaphysics literature; my aim is not to defend it here. Instead I wish to explore the consequences of such a picture for the metaphysical implications of “quasi-realist” expressivism about the normative, of the kind developed by Allan Gibbard.

¹Special thanks to Allan Gibbard, David Manley, Peter Railton, Mark Schroeder, Alex Silk, and an audience at the University of Leeds for helpful comments and/or discussion of earlier versions of this paper.

²Here is a very crude model of this kind of metaphysical difference. Suppose theorists 1 and 2 both agree that there is a single mental state—pain, say—and two physical states—call them A and B. The theorists are in agreement that pain is always co-instantiated among worlds where A is instantiated, and is necessarily not co-instantiated with physical state B. But theorist 1 holds that pain reduces to physical state A in a way that makes pain non-fundamental while theorist 2 holds a proto-Cartesian view on which pain is among the elite, fundamental properties in the world. There is an intuitive metaphysical difference between these views owing to their differences over the eliteness of pain. (Note the difference cannot be captured by claiming that pain does not exist according to theorist 1 but not theorist 2—see Dunaway (MS) for discussion.)
and Simon Blackburn. The central thesis I will explore in this is that quasi-realist expressivism differs from standard non-naturalistic realism about the normative in exactly the way the two views described in the previous paragraph differ from each other. Both theories hold that there are normative truths, that there are normative facts and properties, and that these exist mind-independently, etc. In this respect they hold the same commitments regarding what exists. But they differ over the Lewisian eliteness of normative properties: while the non-naturalist view entails that a halo of eliteness is drawn around the property obligation, the quasi-realist view entails that there is no halo there.

To explain the significance of this thesis, I first discuss in §2 the quasi-realist program in more detail. I then outline in §§3-4 the premises needed for an argument that these two views do in fact differ over the eliteness of the normative. §5 then draws some connections between this way of viewing quasi-realism and other uses for eliteness that have appeared in the literature.

2 Quasi-realism

2.1 A very brief primer on quasi-realism

The quasi-realist program, as I will understand it, proceeds roughly as follows. First, it begins with the familiar expressivist theory of normative language, according to which the meaning of normative sentences is explained by the state of mind those sentences express. Let us follow Gibbard and say that, according to the expressivist, for any atomic normative sentence \( S \), the meaning of \( S \) is explained by the planning-state (or plan, for short) it expresses. Thus on the expressivist view the meaning of sentences like

\[
(1) \quad \text{One ought not to tell lies}
\]

is explained by the fact that \((1)\) expresses a plan that rules out telling lies in a variety of conceivable circumstances.

So far this is a theory of what sentences mean. There is, at this point, the possibility of claiming that while an expressivist can accept the sentence \((1)\), she cannot accept that one ought not to tell lies, since the sentence is only acceptable for the expressivist on something other than its English meaning. Since I do not wish to critique the basic expressivist view on these grounds, I will spot her a premise which guarantees that, since the sentence \((1)\) is acceptable on her view, then that one ought not to tell lies is also consistent with her view (and similarly for other transitions between the formal and material mode). I will call this the

---

2The canonical discussions of the quasi-realist view can be found in Blackburn (1980) and Gibbard (2003: Ch. 9).

3On Gibbard’s official theory, a planning-state is a state of mind that is consistent with some hyperplans—fully decided plans for what to do in every conceivable situation—and rule out others. I will ignore these details of Gibbard’s view for the sake of exposition in what follows.
disquotational principle and will assume it throughout.\footnote{The relevant principle will of course need to be restricted to not permit disquotation of sentences containing indexical or context-sensitive vocabulary, though I will not try to formulate the principle in detail here. See also \cite{Cuneo} for a critical evaluation of uses of disquotation in expressivist theorizing.}

I will also assume here that a theory of atomic normative sentences along these lines can be extended to explain the meaning of complex normative sentences and the normative aspects of “mixed” sentences that contain descriptive and normative elements.\footnote{Cf. the “Frege-Geach” problem as discussed in detail in \cite{Schroeder} and elsewhere.} What is important for the quasi-realist project is the following extension of the basic expressivist approach to normative sentences. Complex sentences about normative truth, facts, belief and so on are plausibly equivalent to normative sentences.\footnote{It is a good question what ‘equivalent’ means here. Plausibly it is not simply necessarily co-extensive since the expressivist needs to be able to claim that, if one can accept the basic normative sentences, it follows that one can also accept their equivalents. I will leave the sense of ‘equivalent’ at issue unspecified at this point, but note that in the specific examples where the expressivist puts these equivalence claims to use, it is plausible that there is a stronger relation than mere necessary co-extensiveness present. For more discussion, see \S\ref{sec:nodifference}.} To take a well-worn example, the quasi-realist might claim that a sentence about normative truth such as (2) is equivalent to the basic normative sentence (1):

\begin{equation}
(2) \quad \text{It is true that one ought not to tell lies.}
\end{equation}

Given that these sentences are equivalent and the disquotational principle, the expressivist can then accept the biconditional (3):

\begin{equation}
(3) \quad \text{It is true that one ought not to tell lies iff one ought not to tell lies.}
\end{equation}

Thus expressivists with certain normative commitments can infer from this that there are normative truths, which include the truth that one ought not to tell lies.

A similar strategy might be employed to show that it is consistent with the Expressivist view that there are some normative facts and properties, and that these facts and properties are mind-independent.

\subsection*{2.2 The \textit{No Difference} conclusion—what is the problem?}

Suppose the expressivist’s basic explanatory resources also extend to allegedly metaphysically substantive claims about truth, facthood, etc. in the way outlined above. \cite{Dreier} summarizes an apparent looming consequence of the success of this program: “those of us who feel confident that there is some difference between the two meta-ethical camps should be concerned that we don’t know how to say what that difference is.”\footnote{\cite{Dreier} (2004: 31)} Gibbard echoes a similar conclusion: “Quasi-realism, in short, elaborates expressivism so that it might turn out to coincide with a form of non-naturalism.”\footnote{\cite{Gibbard} (2013: 233)}
The naive formulation of the conclusion suggested here is No Difference:

No Difference There is no difference between quasi-realist expressivism and non-naturalism.

As it stands, No Difference is not obviously a very plausible or interesting thesis in light of the foregoing. First, it is not plausible because the range of candidate differences is left unconstrained: No Difference, strictly speaking is false if we can find any difference between the supposed quasi-realist and non-naturalist views. And all parties to the debate will grant that there are some differences between the views, if we are sufficiently generous with what would count as a ‘difference’. Thus Gibbard says:

Questions of what we ought to do are questions of what to do, questions we pose in deliberation—and this explains the phenomena to which descriptivists appeal. Indeed, I argue that a form of non-naturalism is correct in a way, as far as it goes—but that it is incomplete.

Here Gibbard is referencing his distinctively expressivist explanation of the meaning of normative language—viz., the explanation of sentences like (1) by reference to the planning state expressed. Since non-naturalists do not typically accept such claims as the explanation of the meaning of normative language, there there is clearly some difference between their views, namely a difference in whether the non-naturalist and expressivist are willing to accept a certain semantic explanation of (1).

But, as stated, such a difference isn’t terribly interesting either. (I am not claiming that the expressivist’s explanation is not related to an interesting difference between the views—indeed I will claim in subsequent sections that the explanatory strategy does give rise to a very important difference. My only point here is that the fact of the expressivist’s acceptance of a particular kind of explanation of meaning is not itself an interesting difference.) The fact that one theory has the property of being believed by Gibbard while another lacks it should strike us as a not-very-interesting way of distinguishing the views. Presumably Dreier is not worried that it is impossible to distinguish the two views by reference to Gibbard’s mental states.

There are a number of ways to supplement No Difference to make it a better characterization of the threatening conclusion that allegedly follows from the quasi-realist’s use of deflationary theses to accept claims about normative truth, fact, etc. I will not attempt to catalogue all of the options here, and instead will simply state a version of the No Difference thesis that has a strong claim to being both plausible in light of the quasi-realist’s use of deflationary measures, and interesting if true. This is the No Metaphysical Difference thesis, which is as follows:

---

\[^9\] Gibbard (2003: 20), see also Dreier (2004: 36).
No Metaphysical Difference There is no difference between the metaphysical commitments of quasi-realist expressivism and non-naturalism.

A metaphysical commitment, as I will understand it here, is a consequence of a view for some metaphysical category—one especially prominent traditional candidate is a view’s implications concerning existence. But as I have indicated in §1, one might extend the list of interesting metaphysical categories beyond traditional ontological questions, and include Lewisian eliteness as well. Thus understood, it will be an interesting and substantive thesis if No Metaphysical Difference were false.

2.3 Theoretical roles: truth, facthood, and eliteness

Before turning to substantive arguments against No Metaphysical Difference, it will be first helpful to make some general observations about why the possibility that it might be true arises in the first place.

Begin with an unpromising resource for showing that No Metaphysical Difference is false, namely truth. Plausibly, if it were the case that non-naturalism entailed that there are normative truths while quasi-realist expressivism denied this, then we would have a substantial metaphysical difference between the views; there would, in particular, be a difference over the existence of normative truth. But the quasi-realist can plausibly accept that there are normative truths, for reasons described above. Here is a conjecture about why this is the case, and consequently why the prospects for showing the existence of a metaphysical difference over truth are not good: the theoretical role truth plays is extremely thin. We know that a candidate meaning for ‘true’ must be one that makes instances the following schema valid, given suitable restrictions:

\[
\begin{align*}
\text{Truth-Role} & \\
\text{If } S, \text{ then it is true that } S, \\
\text{If it is true that } S, \text{ then } S.
\end{align*}
\]

Henceforth I will say that any property satisfies Truth-Role is just one with the following feature: were it to be the referent of ‘true’, all appropriate instances of Truth-Role would be valid.

One significant constraint on the quasi-realist program is that it must not only provide some interpretation of sentences like [2] on which it is acceptable given the expressivist’s commitments, it must in addition provide an interpretation of the sentence which is plausibly what the sentence means in English. The disquotational principle which we spotted the expressivist earlier only applies to English sentences; there would be no path to material mode claims about truth if it could be shown that her claims about what sentences containing ‘true’ mean were incompatible with the meaning of ‘true’ in English. There would in this case be no problem in the vicinity of No Metaphysical Difference.
(To take an extreme example: suppose the quasi-realist accepted the sentence \(2\) but only because she held that it means what ‘grass is green’ means, and she can accept that grass is green. One could in this case show quite easily that she doesn’t accept the existence of normative truth at all, since on her view there is no property that satisfies Truth-Role. \(2\) would take on whatever semantic properties ‘grass is green’ has, and these do not include a satisfier of Truth-Role.\(^{10}\)

But there are straightforward reasons to think the expressivist can not only accept sentences containing ‘true’; she can in addition accept that there is a property which satisfies Truth-Role. She accepts basic normative sentences, via her distinctive expressivist explanation of their meaning. And there is no obvious reason why, given this explanation, she cannot in addition accept that there is a property which satisfies Truth-Role.

To deny at this point that the expressivist genuinely accepts that there are normative truths—to be contrasted with accepting some sentences containing the world ‘true’ on an alternative interpretation—the truth-enthusiast would need to point to some role over and above Truth-Role that is such that (i) it is plausible that the English predicate ‘true’ refers to a property that plays this role, and (ii) expressivism cannot consistently be extended to accept that there is a property that plays the role. Additional roles for truth are bound to be controversial. And even if there is a more robust role in the offing, the quasi-realist may succeed in extending her theory to show that she can accept that there is a property which plays the more robust role.

As a particularly vivid illustration of this point, it is helpful to consider how some writers have tried to emphasize the inadequacy of deflationary truth, whose essence is exhausted by Truth-Role. It is common, for instance, to attempt a distinction between deflationary truth and genuine truth by resort to the all caps term TRUTH, or to add modifiers like ‘real’ or ‘robust’ to the term.\(^{11}\) But it is quite clear that, in the absence of an interpretation of what it takes to satisfy the enhanced truth predicate, the expressivist can consistently maintain that not only are normative sentences true but are also TRUE, really true, robustly true, etc. One might insist that this isn’t so, but (in order to avoid a debate that devolves into table-thumping and reliance on subtle verbal inflections to mark important metaphysical distinctions) this would require pointing to an additional theoretical role that TRUTH or one of its cousins plays over and above Truth-Role. And it would require an argument that the expressivist can’t consistently accept that the satisfier of the role attaches to (for example) that one ought not to tell lies. The prospects for defeating the No Metaphysical Difference thesis with truth as a central resource seem dim, precisely because it is hard to see how Truth-Role can be supplemented in a way that meets these demands.

\(^{10}\)See Dunaway (2010) for application of this broad strategy to some quasi-realist approaches to ‘believes’.

\(^{11}\)See Finch (2011) and Finlay (2012) for some examples.
This dialectic repeats itself with other candidates for driving a metaphysical wedge between non-naturalism and quasi-realism. Facts and properties likewise seem to likewise be promising candidates for the metaphysical differences between non-naturalism and quasi-realism, but as before there are deflationary roles for these entities, as there is apparently nothing which prevents the expressivist from thinking that there are things which satisfy the following schemas:

**Fact-Role**

- If \( S \), then it is a fact that \( S \),
- If it is a fact that \( S \), then \( S \).

**Property-Role**

- If \( x \) is \( F \), then \( x \) exemplifies \( F \)-ness,
- If \( x \) exemplifies \( F \)-ness, then \( x \) is \( F \).**\(^{12}\)**

And as before, we might hold that the English terms ‘fact’ and ‘property’ refer to things which play much richer theoretical roles. The quasi-realist might in turn either reject that the theoretical role in question is one that the referent of the English terms ‘fact’ and ‘property’ require their referents to play, or show it is consistent with quasi-realism that there are things which play the enhanced role. If we wish to avoid table-thumping in our attack on *No Metaphysical Difference*, we need to find plausible claims about facts or properties that satisfy these dual considerations.

The point of emphasizing this aspect the dialectic surrounding quasi-realism is that some theoretical resources will plausibly be associated with richer theoretical roles than others. These resources will thereby potentially be more promising for showing that that *No Metaphysical Difference* is false. There are two reasons why this might be so.

First a notion with a rich theoretical role will require additional interpretive work on the part of the expressivist. Given a theoretical tie between \( t \) and \( t^* \), the expressivist will not only need to provide an adequate interpretation of \( t \), she will also need to provide an interpretation of \( t^* \) which preserves the connections between \( t \) and \( t^* \). And given further additional theoretical connections for \( t \), the constraints on an adequate interpretation will ramify. The quasi-realist’s success with incorporating talk using the theoretically thin notions ‘true’, ‘fact’ and the like may not extend to more theoretically rich notions.

But second even if the theoretical connections encoded in the relevant role can be accommodated on the expressivist’s interpretation, it doesn’t follow that the resulting quasi-realist view will agree with non-naturalism over matters metaphysical. This is because in general it doesn’t follow from the fact that two views accept the same sentences that they agree—they might accept the relevant

---

\(^{12}\)As with Truth-Role, these will need to be restricted to avoid paradox.
sentences by attaching different meanings to them.\footnote{As a simple illustration consider two people who accept the claim ‘Sally is at the bank’. The linguistic similarity would not indicate any agreement over how the world is if one person means ‘financial institution’ by her use of ‘bank’ and the other means ‘side of a river’.} Call such agreement merely linguistic. In the case of ‘true’, which has a thin role, it is difficult to show that someone who has provided an interpretation of ‘true’ on which there is a satisfier of TRUTH-ROLE nevertheless succeeds only at earning a merely linguistic agreement with English speakers who use ‘true’. By contrast, a notion that bears rich theoretical connections is a more promising candidates for a diagnosis of merely linguistic agreement. If the theoretical role for $t$ encodes connections to other familiar subject-matters, this affords us an antecedent grasp on what $t$ is. To the extent that this antecedent grasp of the role is firm, we will have a plausible methodology that does not resort to mere table-thumping for settling whether a view which provides an interpretation for $t$ (including an interpretation for the sentences capturing $t$’s theoretical connections) succeeds only at procuring merely linguistic agreement.

I will elaborate on this latter point more in §5. The limited point here is just that, in order to make progress on No Metaphysical Difference, we will do better a metaphysical notion that has a thicker theoretical role than those of truth, fact, or property. Below I outline a metaphysically substantive notion that plausibly plays a richer theoretical role and is therefore a more promising resource for attacking No Metaphysical Difference.

2.4 The eliteness-role

Lewisian naturalness, or ‘eliteness’ as I am calling it here, is a notion with a rich theoretical role. Theorists have that the perfectly elite properties play a number of important connections to other domains; I mention a few below; this list is by no means intended to be exhaustive.\footnote{For a full catalogue, see Dorr and Hawthorne (2013). I should also add here that I am focusing on the property of perfect eliteness here; there may also be a related gradable notion of relative eliteness (expressed by ‘is more elite than’). One might plausible go in for connections between relative eliteness and degrees of ease of reference: the more elite a property is, the easier it is to refer to it. (And similarly for lawhood and confirmation as well.) For simplicity, I will not discuss relative eliteness any more here.}

**REFERENCE** The perfectly elite properties are easy to refer to.

**LAWS** The perfectly elite properties feature in genuine laws.

**CONFIRMATION** Claims including perfectly natural properties are easy to confirm.

Qualifications and clarifications of each of these claims are in order, although I will not spend too much time on them here. The basic idea behind Reference is that, while we can refer to all kinds of properties, facts about reference are
determined by the twin considerations of fit with use of our terms and eliteness. The perfectly elite properties need to attain a substantially lower fit with use to qualify as referents than other not-very-elite properties. Laws claims that what distinguishes the claim ‘electrons repel each other’ from other true but gerrymandered universal generalizations is that only the former contains an elite property. And Confirmation claims that what makes the laws of physics confirmable on the basis of limited observation, while the generalization that all emeralds are grue is unconfirmable on a similar basis, is that only the former contains perfectly elite properties.

It will be convenient to speak of specific properties having the features specified in Reference, Laws, and Confirmation. In what follows, I will adopt the convention of speaking of satisfaction of the relevant roles; strictly speaking, however, each of these claims expresses a general proposition and is not the kind of thing that can be satisfied. But there are analogous schema for each claim, which are as follows:

Reference$_{\Phi}$  \( \Phi \) is easy to refer to.

Laws$_{\Phi}$  \( \Phi \) is a candidate to feature in genuine laws.

Confirmation$_{\Phi}$  Claims including \( \Phi \) are easy to confirm.

Here, \( \Phi \) is a schematic variable which takes property-names as substitution instances. Hence in what follows I will be asking whether obligation (as our placeholder example of a basic normative property) satisfies Reference$_{\Phi}$; this amounts to the question of whether obligation is easy to refer to. Similar points apply to Laws$_{\Phi}$, Confirmation$_{\Phi}$, and any other structural features of eliteness.

Finally there is a general feature of each of these roles which goes beyond the mere claim that perfect eliteness has the structural features specified by Reference, Laws, and Confirmation. Consider Reference—not only are perfectly elite properties like mass easy to refer to, what explains why mass is easy to refer to is that mass is perfectly elite. (Recall that our rough gloss on “easy to refer to” is just that the property doesn’t need to fit with use to a very high degree in order to be the referent of a term. So in full form the claim here is that the perfect eliteness of mass explains why we can refer to mass without using ‘mass’ in a way that makes for a high degree of fit.) Similarly for Laws and Confirmation: not only are they true, but the perfect eliteness of certain properties also explains why they are true. That is:

Explanation If a property \( p \) is perfectly elite, then the perfect eliteness of \( p \) explains why \( p \) satisfies Reference$_{\Phi}$, Laws$_{\Phi}$, Confirmation$_{\Phi}$, etc.

The claim that there is a notion which has the structural features specified by Reference, Laws, Confirmation, and Explanation is a well-known one from Lewis and his followers, and my aim in what follows is not to defend it but to explore the consequences for No Metaphysical Difference. The forgoing considerations suggest that the notion is a rich one, bearing a wide range of explanatory connections to the linguistic, metaphysical, and epistemic domains. It will, then, be correspondingly more difficult for an expressivist account of the meaning of normative sentences to be turned into a view on which normative properties have all of the structural features of a perfectly elite property. Thus the notion of eliteness is a promising metaphysical category to explore when evaluating No Metaphysical Difference.

3 Eliteness and non-naturalism

The aim of this section and the next is to explore an argument that promises to establish that No Metaphysical Difference is false because non-naturalism and quasi-realism differ over the eliteness of obligation. The argument is necessarily somewhat framework-dependent, as we can’t even formulate what this claim would come to in eliteness-free frameworks. Nevertheless, once the choice of framework is granted, it would appear that given some fairly simple premises we can conclude that any non-naturalist view will hold that obligation is perfectly elite.

Before proceeding, a few caveats are in order.

First caveat: I do not take the arguments below to show that, as a textual matter, non-naturalists have said, or even privately believed, that obligation is perfectly elite. Most discussions of non-naturalism are not conducted in an ideology that distinguishes between elite and non-elite properties, and some proponents of the view would reject the distinction as I am conceiving of it. None of these points will matter to for the issues I will discuss below; the question at hand is what the commitments of non-naturalism and related views are, not what commentators say these commitments are.

Second caveat: many will find the claim that obligation is perfectly elite to be implausible or perhaps incoherent on the grounds that the central notion, and its most salient structural features, were introduced by Lewis as primarily applying to physical properties. Thus they will conceive of Reference as constraining naturalistic reference, Laws as a claim about physical laws, and Confirmation as an explanation of scientific confirmation. There are, however, compelling reasons to reject these as constraints on perfect eliteness. These claims might be true, but they cannot be built into the structural features of eliteness at the outset. The reason is that there is a plausible case to be made that structural features which are restricted in application to physical properties only would not be features of a theoretically interesting property. An artificial constraint on application only to properties in the physical domain would, on other words,
threaten to make the satisfier an unnatural or highly disjunctive kind.\textsuperscript{16} Hence it is reasonable to assume that the structural features of perfect eliteness do not preclude its application to non-natural normative properties, even if in the end the only instantiators of perfect eliteness are physical.

Reference is the structural feature I will be giving most attention to here, so it will be helpful to say make a few additional remarks about what the claim amounts to. It is, at a first pass, a claim about how eliteness interacts with other reference-determining features to provide a referent for a term. The basic idea is that broad facts about how a community uses a term are not sufficient to pin down a sufficiently determinate referent for a term; multiple properties will “fit” with the community’s use to roughly similar degrees.\textsuperscript{16} Reference is the claim that reference is not thereby wildly indeterminate because use alone does not determine reference. Eliteness is also a component—reference, very roughly, is determined by maximization of fit and eliteness. An elite property need not fit best with a community’s use of a term in order for the property to count as the referent; some non-elite properties might fit just as well or better, but the eliteness of one moderately good fit candidate referent can in principle override the greater degree of fit of other less elite candidates.

To help gain fluency with this idea, consider an eliteness-based solution to a puzzle in Kripke (1982). (This is the famous ‘Kripkenstein’ puzzle about the reference of the term ‘+’.) The puzzle in basic form points out that it is very natural to think that the referent of ‘+’ in the mouths of a community much like ours is plus, but that there are many other candidate referents that fit the community’s of ‘+’ equally well. One example is the quus function: it fits the use of the community’s use of ‘+’ just as well as plus, as (on Kripke’s description) it is a function which outputs the sum of any input of a pair numbers under 57; otherwise it outputs 5. Thus for a community that has no dispositions to answer questions involving numbers above 57, just as many of the utterances of the community in question come out as true on an assignment of quus as the referent of their ‘+’ as on an assignment of plus as the referent. Use-based considerations then will not alone secure the conclusion that ‘+’ as used by such a community refers to plus. Those who are attracted to Reference will be tempted to say: even if plus and quus fit with use equally well, the former is plausibly highly elite while the latter is not. According to the Reference-based solution, plus is for this reason determinately the referent of ‘+’.\textsuperscript{18}

3.1 Non-naturalism: explanatory desiderata

A typical non-naturalist view will accept the following claims:

\textsuperscript{16}Plausibly, moreover, Explanation would not be true of perfect naturalness if it is gerrymandered in this way.
\textsuperscript{17}See §5 for more detailed discussion of this issue.
\textsuperscript{18}I discuss variants on the Kripkenstein case more in §8.
No Indeterminacy  There is no widespread indeterminacy in what ‘ought’ and other normative terms refer to.

Disagreement  Communities who use normative terms slightly differently nevertheless succeed in (determinately) referring to the same property, and hence disagree with each other.\(^{19}\)

Importantly a purely use-based conception of the reference of normative terms fails to explain No Indeterminacy and Disagreement. There are communities for whom the facts about the use of the term ‘ought’ are limited—picturesquely, they have only managed to apply ‘ought’ to a finite range of cases, call this \(R\). There exist cases outside of \(R\), and multiple properties which fit with the community’s application of ‘ought’ within \(R\) but which diverge in extension outside of \(R\) will all fit with the community’s use equally well. Hence, given purely use-based constraints on reference, the it will be indeterminate what the referent of ‘ought’ is, contra No Indeterminacy.

For similar reasons fit with use alone will not explain Disagreement: take for example two communities use ‘ought’ slightly differently (say, one persistently applies it to happiness-maximizing violations of autonomy, while the other consciously and persistently refrains from making such applications). Since distinct properties best fit the use of ‘ought’ by these communities, a purely fit-based approach to reference will predict that these communities refer to distinct properties, contra Disagreement.\(^{20}\)

Analogous claims will be plausible given other more sophisticated accounts of reference available to the non-naturalist. For instance, a “conceptual role” approach as in [Wedgwood (2007)] claims (very roughly) that the referent of ‘ought’ is the property which explains why the inferences required for competence with ‘ought’ are correct. (Wedgwood is clear that the inferences in question will be those of practical reasoning in addition to those from traditional theoretical reasoning. In the case of ‘ought’, the inference in question is one from the state expressed by \(⌜\text{it ought to be that } a \phi s \text{ in } c⌝\) to a plan to \(\phi\) if \(a\) in circumstance \(c\).) The basic point is the same as before: there will be many properties such that if ‘ought’ refers to that property, it is correct to make the corresponding inference. (Thus it is correct to plan to \(\phi\) if \(a\) in \(c\) if one is correctly in the state expressed by \(⌜\text{it ought to be that } a \phi s \text{ in } c⌝\) and ‘ought’ refers to obligation. But it is also correct

\(^{19}\)Cf. [Horgan and Timmons (1992)].
\(^{20}\)Such cases obviously bear some resemblance to the “Moral Twin Earth” cases from [Horgan and Timmons (1992)] and elsewhere. Horgan and Timmons clearly take these cases to be especially problematic for naturalistic moral realism, since they take themselves to have an argument that the naturalistic versions are committed to giving the wrong result in these cases. (For interpretation and criticism, see [Dunaway and McPherson (MS)].) Even though they do not single out moral realism in its non-naturalist incarnations for criticism on similar grounds, a non-naturalist will still need an explanation for the same disagreement data. Moreover I am making no assumptions about the extent to which genuine disagreement extends—I am only assuming that some communities manage to use their normative terms differently yet still refer to the same property.
to plan in this way if ‘ought’ refers to the property of either being an action one is obligated to do or something that is entertained by an amoeba.\footnote{\textsuperscript{21}}

3.2 The exhaustive conception of eliteness

How does non-naturalism explain commitment to No Indeterminacy and Disagreement? Certainly one option available to the non-naturalist on these issues is to hold that obligation is highly elite. This would be an interesting in relation to the No Metaphysical Difference thesis, which is plausibly false if non-naturalism is consistent with an eliteness-based explanation of No Indeterminacy and Disagreement while expressivism is not.

Here however I will consider an argument for a stronger and potentially more interesting thesis: that it follows from the basic features of the non-naturalist view that obligation is perfectly elite. At the center of this argument is the that satisfying components of the eliteness-role like Reference\(_\Phi\) pretty much exhausts what it is to be a perfectly elite property. I will call this the exhaustive conception of eliteness.\footnote{\textsuperscript{22}} On the exhaustive conception, what it is to be elite just is to be easy to refer to, to feature in generalizations that are eligible for lawhood, to be a constituent of claims that are easy to confirm, etc. This is only a rough characterization because Explanation is a constraint on eliteness as well: it must be that the perfect eliteness of a property is what explains why it is easy to refer to, an so on. The exhaustive conception isn’t an essential thesis for the argument that follows, but it will make the argument much simpler. (Even if someone denies it, one could still give a abductive argument that follows the outline of the argument below.)

To get a sufficiently precise characterization, we can begin by taking the Ramsey sentence obtainable a characterization of eliteness which stays that the eliteness of a property explains why it satisfies Reference\(_\Phi\), explains why it satisfies Laws\(_\Phi\), and so on. These are the components of the eliteness-role. A property plays the eliteness-role simpliciter, then, just in case it has a feature which explains why it satisfies all of the components. As a bit of shorthand, properties satisfy Eliteness-Role:

\[
\text{Eliteness-Role } \exists x : \Phi \text{ satisfies } \text{Reference}_\Phi \text{ because } \Phi \text{ has } x ; \Phi \text{ satisfies } \text{Laws}_\Phi \text{ because } \Phi \text{ has } x ; \Phi \text{ satisfies } \text{Confirmation}_\Phi \text{ because } \Phi \text{ has } x \ldots
\]

On the exhaustive conception of perfect eliteness, there is nothing more to being a perfectly elite than playing the eliteness-role; any property that satisfies Eliteness-Role is perfectly elite. There is, in other words, nothing more to

\footnote{A similar point applies to a platitudes-based approach as in \textit{Jackson and Pettit (1995)}.}

\footnote{Compare: there are two conceptions of what it is to be an electron. On one, attracting protons and repelling other electrons is just what it is to be an electron; anything which plays this role is an electron. On the second, electronhood might have played the mass-role, the charge-role, etc. The first conception is akin to a denial that properties have “quiddities”; there is nothing to what it is to be a certain property over and above playing a certain role. The approach to eliteness that I am adopting here is akin to the quiddity-denying approach to electronhood.}
being perfectly elite than simply having some property that explains why you are easy to refer to, a candidate for constituency in laws, feature in easy to confirm generalizations, and so on. We can encode this claim in **Eliteness Def.**:

**Eliteness Def.**  \( \Phi \) has the property \( x \) and \( x = \) perfect eliteness iff \( \Phi \) satisfies 

1. **Reference**\( _{\Phi} \) because \( \Phi \) has \( x \); \( \Phi \) satisfies **Laws**\( _{\Phi} \) because \( \Phi \) has \( x \); \( f \) satisfies 
2. **Confirmation**\( _{\Phi} \) because \( \Phi \) has \( x \) . . .

**Eliteness Def.** provides us with a workable criteria for settling whether obligation is perfectly elite on the non-naturalist view. Obligation must satisfy the **Reference** component by (i) being easy to refer to because of some property it has, and (ii) being such that the same property also explains why obligation plays the other components of the eliteness role. I will take each of these points in turn.

### 3.3 Non-naturalism, reduction, and ease of reference

The first question is whether obligation has a property that explains why it is easy to refer to (I will assume from here on that it must somehow be easy to refer to; this is needed for an adequate explanation of No **Indeterminacy** and **Disagreement.**) The alternative is that the ease of reference to obligation is explained by its relation to other, distinct properties. Take a simple case: if being a Marge is being a thing that either has mass or charge, Margehood will be easy to refer to. One needs only to introduce it via its canonical definition ‘having mass or charge’. But it isn’t a property of Margehood that explains why it is easy to refer to; the ease of referring to mass and charge separately does all of the explanatory work. So the first step is to rule out that the non-naturalist can explain the ease of referring to obligation via the ease of referring to other, distinct properties.

It is beyond the scope of this paper to give a full characterization of the non-naturalist view, but I take it that a central aspect to the view is its **anti-reductionism:** at the very least non-naturalist is incompatible with the reducibility of the normative. I will label this thesis **ANTI-REDUCTIONISM:**

**ANTI-REDUCTIONISM** If non-naturalism is true, then obligation does not reduce, and hence all reductive explanations of the features of obligation are false.

I also won’t say much about what reduction is here, but the key claim for our purposes is that it is a dependence relation, and hence is explanatory. A reduction

---

23 I assume here that the property which Margehood does have—the property of having constituent properties that are easy to refer to—is too gerrymandered to be a genuinely explanatory property; only elite properties can enter into explanations.

24 Of course even non-naturalists can accept intra-normative reductions, and hence might reduce obligation to some other normative property, or vice versa. **ANTI-REDUCTIONISM** as I am understanding it here presupposes that obligation is the basic normative property and so won’t be on the reduced end of any intra-normative reductions. Analogous theses will be available to those who have different views about normative priority.
As to Bs, I assume, requires that all of the features of the As be explainable in B-theoretic terms.\textsuperscript{25}

The first part of the argument that obligation satisfies ELITENESS Def. and hence is perfectly elite is that the ease of referring to obligation must be explained by a feature of obligation itself, since the non-naturalist is committed to ANTI-REDUCTIONISM. The non-naturalist is as a result committed to the falsity of any view which attempts to explain the ease of referring to obligation in terms of the ease of referring to other properties that are distinct from it. While the ease of referring Margehood is explainable in non-Marge-theoretic terms, the ease of referring to obligation isn’t explainable in non-obligation-theoretic terms: it must be a feature of obligation that, at the most basic level, explains why the property is easy to refer to.

3.4 Parsimony and consistent roles

The second part of the argument is to establish that it is the same feature which explains why obligation satisfies the other components of the eliteness-role as well. The alternative is that, while obligation has some property which explains why it plays each of the individual components of the eliteness-role, there is no one property that it has which explains why it plays all of them. On this alternative, obligation satisfies not ELITENESS-ROLE but rather the following distinct schema:

\textbf{Disjunctive-Role} \ \exists x, y, z \ldots : (x \neq y) \lor (y \neq z) \lor \ldots \text{ and } \Phi \text{ satisfies Reference}_\Phi \text{ because } \Phi \text{ has } x; \Phi \text{ satisfies Laws}_\Phi \text{ because } \Phi \text{ has } y; \Phi \text{ satisfies Confirmation}_\Phi \text{ because } \Phi \text{ has } z \ldots

That is: there is some feature of obligation x that explains why it satisfies Reference$_\Phi$; there is some feature of obligation y that explains why it satisfies Laws$_\Phi$, there is some feature of obligation z that explains why it satisfies Confirmation$_\Phi$, and so on; and in addition, not all of these explanatory features are identical. Hence if obligation satisfies Disjunctive-ROLE, it doesn’t satisfy ELITENESS-ROLE.

There is a strong case to be made that the non-naturalist should reject that obligation satisfies Disjunctive-ROLE. First there are parsimony considerations: given that there are some unreduced features of obligation which explain why obligation satisfies Reference$_\Phi$, Laws$_\Phi$, Confirmation$_\Phi$, etc., the non-naturalist should, all things considered, hold that it is the same property which is explanatory each case. She would have a substantially less parsimonious theory if she were to hold that the properties in question were distinct.

Of course there are plenty of cases where a lack of parsimony is called for: sometimes a theory cannot be explanatorily adequate without introducing additional explananda. (The non-naturalist should be familiar with this need; her commitment to unreducible normative properties is motivated by the view that the

\textsuperscript{25}See Barnes (2012) for more on dependence and related metaphysical notions.
more parsimonious naturalist view is explanatorily inadequate.) In this case the properties are distinct would be motivated if the components of the eliteness-role could not be jointly satisfied by a single property. For instance: if reference and confirmation were structurally related to each other in ways which required that some properties which are easy to refer to are not properties that feature in easy to confirm claims, then it would perhaps be warranted to think that obligation satisfies Disjunctive-Role instead.

It is a large question beyond the scope of this paper whether the components of the eliteness-role are structurally consistent in this way. All I will point out here is that this is a question about the viability of an eliteness-based metaphysics in general, and is not a question specific to the eliteness of obligation in particular. It is not at all obvious that the structural features of an eliteness-like property are all mutually consistent, and the Lewisian eliteness-centric framework may not be viable for this reason. But in light of this, holding that Disjunctive-Role rather than Eliteness-Role is satisfied by obligation, on account of inconsistency among the components of the eliteness-role, amounts to rejecting the entire Lewisian eliteness-centric framework that we introduced in §1.

The argument for the perfect eliteness of obligation on the non-naturalist view leaves a few options for those who wish to reject its conclusion. They can question the exhaustive conception of eliteness and raise challenges to analogous abductive arguments for the same conclusion that do not rely on the premise. They can raise worries about the relevance of Anti-Reductionism to non-naturalist explanations of normative reference, lawhood, confirmation, and the like. And they can claim that there are reasons to hold that the properties of obligation that feature in these explanations satisfy Disjunctive-Role rather than Eliteness-Role. I haven’t given knock-down arguments here that each of these options are unavailable. But the premises in this argument are all fairly plausible, and are yield a conclusion that provides significant leverage on the No Metaphysical Difference thesis.

4 Eliteness and quasi-realism

The hypothesis laid out in §1 was that non-naturalism and quasi-realism differ in their commitments regarding the perfect eliteness of obligation. We have already seen how the structural features of eliteness might be used to argue for the first half of this thesis: obligation is perfectly elite on the non-naturalist view. This section explores a separate argument for the second half: obligation is not perfectly elite on the quasi-realist view. This is because obligation does not satisfy Eliteness-Role on the quasi-realist view.

Playing the eliteness-role, we will recall, requires having a property which explains why one satisfies the various components of the role. An argument that obligation is not a property that plays this role on the quasi-realist view would show that there is no feature it has which explains why it satisfies Reference$_{Ψ}$ or one of the other components. As an explanatory notion, eliteness is subject to
parsimony considerations: we can infer that the quasi-realist rejects that there is
such a feature if her view explains all of the relevant claims without reference
to such a feature. Thus the main concern of this section is two arguments
for the conclusion that the quasi-realist view does in fact accomplish all of the
relevant explanatory work without the need for a property that plays the eliteness-
role. The first is that it is a consequence of the general expressivist explanatory
strategy detailed in §2. The second is that the most (only?) well-developed
expressivist account of reference in Gibbard (2013) explains the relevant claims
without invoking the eliteness of normative properties.

I will look at each of these arguments in turn.

4.1 The quasi-realist explanatory strategy, revisited

Recall that the expressivist starts by offering a theory about the meaning of atomic
normative sentences such as

(1) One ought not to tell lies.

The rough characterization of the expressivist theory we are working with here
holds that the meaning of (1) is explained by the planning-state it expresses. Being
in that planning state is all there is to accepting (1). Given the disquotational
principle, the expressivist’s apparatus explains why she can accept that telling lies
is wrong.

The quasi-realist component of expressivism, as I have presented it, consists in
finding equivalences between the normative claims like (1) that they already have
an explanation for, and metaphysically heavy-duty claims that are prima facie
acceptable only to a realist. Given these equivalences, plus the fact that the first-
order normative claims are acceptable to the expressivist, it follows that the claims
that are prima facie are acceptable only to a realist are also claims an expressivist
can accept. Via the disquotational principle, the expressivist view has the same
metaphysical commitment as the realist.

This is just a quick recap of the quasi-realism program as introduced in §2.
Here it is worth exploring in more detail the central claim in any instance of
this quasi-realist strategy: that for any realist-sounding sentence $E$, it is equivalent
to a normative claim $N$ that is antecedently acceptable to the expressivist. The
equivalence between $N$ and $E$ cannot be just any equivalence relation if the quasi-
realist strategy is to work: it must one which entails that the basic the explanatory
resources for showing ordinary normative claims like $N$ are acceptable are also
sufficient for showing that $E$ is acceptable. That is, the equivalence-relation must
be such that the following holds:

**Explanatory Equivalence**: If $N$ is equivalent to $E$, then whatever explains why
$N$ is acceptable is also sufficient to explain why $E$ is acceptable.

Suppose we denied Explanatory Equivalence and held that the consequent
is false because the resources that explain the acceptability of $E$ are distinct from
those that explain the acceptability of \( N \). For instance, suppose \( R \) is the needed resource (it isn’t needed to explain the acceptability of \( N \) but is needed for \( E \)). Even if the antecedent is true—there is some sense of ‘equivalent’ on which \( N \) is equivalent to \( E \)—the fact that \( N \) is equivalent to \( E \) in this sense would be of no help in carrying out the quasi-realist’s program. Given our supposition that \( R \) is required for \( E \) but not \( N \), the fact that \( N \) is equivalent to \( E \) in some sense would not entail that \( E \) is acceptable if \( N \) is. Someone who denied that \( R \) existed could coherently accept \( N \) but would not be in a position to coherently accept \( E \). The quasi-realist strategy assumes that this isn’t a live option: that once we have normative sentences which are candidate equivalents for realist-sounding sentences, those who accept the relevant normative claims can thereby accept the realist equivalents. The quasi-realist needs these claims to be equivalent in a sense for which EXPLANATORY EQUIVALENCE holds.

Without worry about the details for the moment, let’s assume that the quasi-realist can carry out the project of finding a normative sentence for every one of the realist’s sentences that satisfies EXPLANATORY EQUIVALENCE. In particular, we can assume that the quasi-realist accepts that obligation satisfies \( \text{REFERENCE}_\theta \), that is, that \( \text{REFERENCE}_\theta \) is true:

\[
\text{REFERENCE}_\theta \quad \text{Obligation is easy to refer to.}
\]

In keeping with the quasi-realist program, \( \text{REFERENCE}_\theta \) will be acceptable to the quasi-realist because she has found some normative claim that it is equivalent to—let’s call this claim \( N_\theta \).

The argument that obligation isn’t elite on this view then goes as follows. If \( \text{REFERENCE}_\theta \) is equivalent to \( N_\theta \), then by EXPLANATORY EQUIVALENCE whatever that explains why \( N_\theta \) is acceptable also explains why \( \text{REFERENCE}_\theta \) is acceptable for the quasi-realist. But \( N_\theta \) won’t be explained by the existence of any properties that play the eliteness-role; it is a complex of ordinary normative claims, and all the expressivist needs to explain these are the familiar planning-states. Thus given EXPLANATION EQUIVALENCE the eliteness of some property won’t explain \( \text{REFERENCE}_\theta \) either. Hence there will be no property that plays the eliteness-role on the quasi-realist view; obligation might be easy to refer to, but this will not be in virtue of its superior eliteness.

---

26 Suppose ‘equivalent’ just means ‘necessarily co-extensive’. Showing that two sentences are equivalent in this sense won’t meet the quasi-realist’s needs because necessary co-extensiveness won’t imply EXPLANATORY EQUIVALENCE. As a simple case suppose mathematical platonism and theological realism are true and consider the necessary truths (under these suppositions) expressed by ‘\( 2 + 2 = 4 \)’ and ‘God exists’. These claims are necessarily co-extensive but it doesn’t follow that an explanation for the acceptability of ‘\( 2 + 2 = 4 \)’ will suffice for an explanation of ‘God exists’; someone who denied the existence of God would be in a position to accept the former but could not coherently accept the latter.

27 Given an application of the disquotational principle to allow us to say not only that expressivists can accept \( \text{REFERENCE}_\theta \), but also that they can accept that obligation is easy to refer to, some care is needed with the talk of ‘explanation’ here. Although expressivists like Gibbard will talk of “explaining” normative claims by reference to the planning-states they express, they explicitly deny
The argument of the previous section is somewhat abstract, as we made no assumptions about how a quasi-realist would find an equivalent normative sentence for \( \text{Reference}_0 \); we only assumed that she would be able to find some normative sentence \( N_0 \) to suit her purposes. The abstractness is to some extent a good thing; without making concrete assumptions about how the quasi-realist might go about explaining why obligation satisfies \( \text{Reference}_\Phi \), we have an argument that regardless of the details of implementation, the quasi-realist strategy commits itself to disagreement with the non-naturalist over the eliteness of obligation.

But it can be helpful to consider how the argument goes once the details are in place, so before moving on to concluding remarks I will consider a case study: the discussion of reference, and in particular its relation to \text{Disagreement}, in Gibbard (2013).

Gibbard’s discussion is set in the larger context of an expressivist treatment of ‘means’, which he analyzes as a normative notion. He then applies the explanatory resources to the related notion of reference, which is our primary interest here. As a primer on Gibbard’s treatment, consider his hypothetical subject Quursula, who will be recognizable from our earlier discussion of Kripke (1982). Quursula has no dispositions to use the sign ‘+’ when presented with queries involving ‘+’ and numbers over 50. (She simply answers randomly for such questions, or looks confused and doesn’t answer at all.) For simpler questions involving ‘+’, she reliably adds. But her surrounding community is not exactly like her—while they do add for the same simple questions, they also take a stance on the more complicated questions they respond by applying the \text{quus} function to the relevant pairs, answering ‘5’ for all of them.

Gibbard next introduces two theorists about meaning; these are theorists who make claims about what ‘+’ means in Quursula’s mouth. The first is Tyler, who holds that both the dispositions of an individual and the dispositions of her linguistic community determine what she means. The second is Jerry, who holds that an individual’s dispositions and the relative simplicity of candidate referents determine what she means. Tyler and Jerry hold two incompatible meta-semantic theories, and Gibbard describes their disagreement in the familiar expressivist way. In short, he says that Tyler and Jerry, in accepting different theories about
what Quursula means, adopt different plans for their own linguistic behavior in Quursula’s situation. Tyler plans to quadd with the rest of Quursula’s community in the circumstance where he is presented with a problem containing ‘+’, while Jerry plans for the same circumstance to add—that is, to apply the simplest function that fits Quursula’s dispositions. Since these plans are incompatible—it would be incoherent to plan for one and the same circumstance to both respond with the result of adding and quadding—Tyler and Jerry disagree about what Quursula means. Gibbard summarizes:

An Expressivist could further maintain that the issue is just this: which statements of Quursula’s to accept if one is Quursula. At issue between Jerry and Tyler is at least this: whether, if one is Quursula, to accept or reject the sentence she writes ‘68 + 57 = 125’ […] It is a difference on what to accept if in a hypothetical circumstance, that of being Quursula.28

It is straightforward to turn this into an account of what it is to accept reference-claims about what ‘means PLus’ refers to. If accepting meaning-claims is just to plan to accept particular sentences in certain circumstances, then accepting that ‘means PLus’ refers to a property \( P \), roughly, just is to plan to accept a sentence \( s \) in a circumstance \( c \) just in case \( s \) has \( P \) in \( c \). In Gibbard’s example, there are two candidate referents, \( P_{Jerry} \) and \( P_{Tyler} \), which are as follows:

\[
P_{Jerry} = \text{the property of scoring the best on both metrics of fit with an individual’s linguistic dispositions and simplicity}
\]

\[
P_{Tyler} = \text{the property of being the best fit with the linguistic dispositions of an individual’s community}
\]

The addition function has \( P_{Jerry} \) in Quursula’s situation, while the quaddition function has \( P_{Tyler} \). Jerry, who endorses the individualistic constraints on reference, plans in a circumstance to accept in any circumstance (including Quursula’s) the sentence that has \( P_{Jerry} \); Tyler, who accepts community-wide constraints on reference, plans to accept in any circumstance (including Quursula’s) the sentence that has \( P_{Tyler} \). Jerry and Tyler disagree about reference because these plans are inconsistent: in some circumstances (e.g., Quursula’s) different sentences will have \( P_{Jerry} \) and \( P_{Tyler} \) and so planning with Jerry means thereby not planning with Tyler, and vice versa.

To return to the issues at hand: Reference\( _O \) is just a paraphrase of a conjunction of claims about the reference of another normative term, namely ‘ought’. If obligation is easy to refer to, then for a wide range of possible uses of the normative term ‘ought’, those who use it in these ways will still manage to refer to the same property. Reference\( _O \) is then a claim about what ‘ought’ refers to in

28Gibbard (2013: 46)
the mouths of a range of possible speakers; we can think if it as a large conjunction of claims about what different speakers in a wide range of circumstances use their word ‘ought’ to refer to, as follows:

\[ R_1 \text{ ‘ought’ as used by } s_1 \text{ in } c_1 \text{ refers to } P_O; \]
\[ R_2 \text{ ‘ought’ as used by } s_2 \text{ in } c_2 \text{ refers to } P_O; \]
\[ R_3 \text{ ‘ought’ as used by } s_3 \text{ in } c_3 \text{ refers to } P_O; \]

\[ \vdots \]

Gibbard explains each of \( R_1, R_2, R_3 \ldots \) by showing that it is equivalent to a claim about what one ought to do with the word ‘ought’ in the circumstance of being the possible speaker \( S_n \) in circumstance \( c_n \). In particular, since obligation is easy to refer to, what one ought to do is apply ‘ought’ to those actions which instantiate obligation. And this is a claim Gibbard already has explained: it expresses a plan, for the circumstance of being one of each of the possible speakers, to apply ‘ought’ to the actions which instantiate obligation. So Gibbard plausibly has the resources to accept \( \text{REFERENCE}_O \) and by the disquotational principle, can accept that obligation is easy to refer to.\(^{29}\)

Let’s grant that this gives an outline of how reference-claims are equivalent to familiar normative claims. A claim of the form \( R_1, R_2, R_3 \ldots \) can be explained by the planning-state it expresses, and this explanation has no need of elite normative properties. (Gibbard uses the exact same resources to explain it that he uses to explain \([1]\) and isn’t committed to eliteness claims by virtue of accepting that explanation.) Hence by EXPLANATORY EQUIVALENCE, \( \text{REFERENCE}_O \) is not explained by a property that plays the eliteness-role either. If we spot the quasi-realist everything she needs to get the expressivist theory of normative language off the ground, and to make the quasi-realist approach to metaphysics-heavy realist talk workable, we have thereby spotted her a theory which does all of the explanatory work without the need for Lewisian eliteness.

Thus there are arguments from both the general contours of the quasi-realist program, and the most detailed implementations of it in Gibbard’s work, for the second half of the conclusion promised in §III. Again these arguments might

\(^{29}\)To illustrate this, take a simple and familiar example. Suppose we ought to maximize happiness. Sally, however, is a possible speaker who doesn’t believe this—she thinks there are constraints on when we can maximize happiness and in particular thinks that one is obliged to refrain from doing this when it would involve one’s violating the autonomy of another person; call the distinct property she applies ‘ought’ to \( P_A \). But since obligation is easy to refer to, Sally doesn’t manage to refer to the property \( P_A \); she refers to happiness-maximization in spite of her divergent usage. To say that Sally refers with ‘ought’ to happiness-maximization is just to say that she ought in her circumstance to apply ‘ought’ to the property of happiness-maximization. And this is something Gibbard can clearly accept, since he explains it as an expression of the plan for being in Sally’s circumstance to apply ‘ought’ to all and only actions that maximize happiness. Similar points go for other possible speakers whose use of ‘ought’ does not perfectly track the property it refers to in their mouths.
be resisted, but they are fairly straightforward, only appealing to the **Explanatory Equivalence** principle plus some basic assumptions about the original expressivist explanation of ordinary normative claims. If these arguments are successful, they suggest a response to **No Metaphysical Difference** on which non-naturalism and quasi-realism might agree entirely about what there is, but nonetheless differ over a notion with a rich theoretical role—perfect eliteness.

5 Concluding remarks: Putnam and permuted interpretations

What I have given above is at best an outline for how one might go about using perfect eliteness as a resource for showing that **No Metaphysical Difference** is false. I have however argued that the eliteness-based arguments sketched above are much more promising than traditional resources for attacking the problem, since they employ a metaphysical notion with a rich theoretical role. In closing I will make one final point about this general strategy by connecting it to other uses of the Lewisian eliteness-centric framework in the literature.

One thing I have not done is to ask whether there is some way for the quasi-realist to accept the sentences that are used state Eliteness-Role using the basic quasi-realist strategy. For all I have said here, the expressivist could carry out the latter project; I have not argued that this is not possible. But it would be of little comfort to the defender of **No Metaphysical Difference**.

Consider a related job for eliteness, where it appears in a Lewis-style response to the “permutation argument” in *Putnam* (1981). To a rough first approximation, Putnam argues that on a “realist” view, reference is going to be massively indeterminate—given that reference is roughly determined by “fit” with a community’s use of their sentences, there will be multiple, wildly divergent interpretations each of which fits use equally well. Reference will be indeterminate between the referents assigned by the divergent interpretations on the realist view.

There are two aspects to this rough characterization of the Putnamian argument that can be given a little more detail (I won’t elaborate on his focus on “realism” here). First “fit” is to be cashed out in terms of truth: an interpretation fits with a community’s use of sentences to the extent that the interpretation counts those uses as true. A correct interpretation for Putnam, then, maximizes the degree to which a community’s use of their sentences are true. Second, we can get a handle on the relevant “divergent” interpretations as follows. Begin with what we might call the *standard* interpretation of English, on which ‘cat’ refers to cathood, ‘person’ refers to personshood, etc. A *permuted* interpretation then assigns a referent that is determinately distinct from cathood to ‘cat’, and then makes structural adjustments to the rest of the assignments of referents so as to not score more poorly on fit-based considerations (hence personhood won’t be assigned to ‘person’; ‘cats and persons are both mammals’ will need to come out true to preserve fit).

The response in *Lewis* (1984) to the permutation argument is to hold that
fit alone does not determine reference: for Lewis, the correct interpretation is one that fares best on the twin considerations of fit and eliteness. While the permuted interpretations fare just as well as the standard interpretation on the fit metric, they will score considerably worse on eliteness considerations: a permuted interpretation may start out by assigning something roughly as elite as cathood as the referent of ‘cat’, but it will have to make corresponding adjustments in the assignments it gives to the rest of the language in order to preserve fit, and will as a result assign highly gerrymandered and not-very-elite referents in order to accomplish this. Wildly divergent permuted interpretations are not among the correct interpretations on the Lewisian approach.

Lewis insists that this approach not be confused with another approach which would offer no help at all in light of the permutation argument. The alternative approach is to hold that, among the sentences that must come out true in order for an interpretation to preserve fit are sentences containing ‘elite’. For instance we might imagine a community who, in addition to using sentences like ‘cats have tails’ and ‘there is a cat in my yard’, also says things like ‘cats are very elite’. If all that is required of an interpretation is that the sentence containing ‘elite’ is among the sentences that the interpretation must account for, then so long as a permuted interpretation finds some property to assign to ‘elite’ that preserves fit, it fares no worse on reference-determining considerations.

The Lewisian view is instead that correct interpretations must maximize fit plus eliteness, and not simply provide some fit-preserving interpretation for ‘elite’. The permuted interpretations which succeed at the later task still don’t do very well on eliteness considerations; the property they assign to ‘elite’ is not eliteness but rather some distinct gerrymandered property that has nothing to do with reference. Lewis is not denying that some permuted interpretation will provide fitting referents for ‘eliteness’-talk, but this is beside the point; all his solution requires is that the referents assigned on the permuted interpretation will not be elite.

Is worth emphasizing in the context of the eliteness-based response to No

---

*Lewis puts this point as follows:*

> Suppose that we say it is constraint C that saves the day—a causal constraint, perhaps, or what have you. We offer an account of how constraint C works, a bit of theory in fact. If this bit of theory looks good, it will deserve to be incorporated into total theory. Suppose it is. Then an intended interpretation must make C-theory come true, along with the rest of total theory. But it will still be true, as much as ever, that (almost) any world can satisfy (almost) any theory. Adding C-theory to the rest of total theory doesn’t help. It is still trivially easy for a world to make total theory come true, and in fact to do so in countless ways. And the point is general: it applies to any constraint (or, at least, to any otherwise satisfactory constraint) that might be proposed. […]

To which I reply: C is not to be imposed just by accepting C-theory. That is a misunderstanding of what C is. The constraint is not that an intended interpretation must somehow make our account of C come true. The constraint is that an intended interpretation must conform to C itself. *(Lewis 1984: 225)*
Metaphysical Difference that an analogous position is available. First, a proponent of the solution I have explored here might legitimately claim, with Lewis, that they are not claiming that quasi-realists cannot mimic ‘eliteness’-talk using the standard quasi-realist strategy. Instead they might claim is that, even if quasi-realists can do this, normative properties are not elite on the resulting view. Second, because of the rich and theoretical connections of eliteness, it is an especially promising and non-ad hoc resource to feature in a response of this kind. We can imagine someone making the analogous claim about truth in responding to the quasi-realist: she could say that the quasi-realist has managed to capture ‘truth’-talk but has nonetheless failed to provide a view which is about truth. But if the theoretical role of truth is exhausted by Truth-ROLE, then the debate over this response will only amount to an exchange of flimsy intuitions about whether the quasi-realist’s satisfier really is truth. The rich connections between eliteness and epistemology, language, and science (as captured by Eliteness-ROLE) lend a greater amount of credibility to accusations that someone has managed to accept all of the sentences used to talk about the crucial notion, but not succeeded in actually talking about the crucial notion.

Note finally that once we have a difference over the eliteness of normative properties in play, we can give derivative characterizations of differences over truth, facthood, property-instantiation, etc. as well. If a fundamental truth involves instantiation of a single elite property, then quasi-realists will deny that normative claims are fundamentally true. Similarly for the analogous notion of a fundamental fact. And if we restrict quantification to those properties that are perfectly elite (and not Boolean operations on them), the quasi-realist must deny that there is a property telling lies has. But these differences in the standard metaphysical categories will only arrive once we use eliteness to supplement the notions of truth, facthood, and property, and thereby connect them to the robust theoretical roles that eliteness plays.

References


Terence Cuneo. The Expressivist Slide. MS.


Billy Dunaway. The Metaphysical Conception of Realism. MS.

Billy Dunaway and Tristram McPherson. Reference Magnetism as a Solution to the Normative Twin Earth Problem. MS.


