A Primitive Solution to the Negation Problem

Derek Shiller

(Penultimate Version)[1]

1 Introduction

Expressivism is a radical theory of mind and language that holds that normative judgments are noncognitive states. Normative assertions don't represent the world as being any way. Rather, they express our feelings. Eschewing the metaphysical extravagances of normative realism comes at a cost. The expressivist must relinquish many explanations provided by representational theories of mind and language; if their theory is to be credible, expressivists are obligated to make up for this explanatory sacrifice. Critics of expressivism have often maintained that this obligation cannot be met. Recently, Nicholas Unwin (1999; 2001) and Mark Schroeder (2008) have alleged that expressivists cannot explain the appearance of inconsistency between certain normative judgments, and they hold that this is deeply problematic. They insist that existing expressivist theories cannot even explain how it is that thinking that bank fraud is wrong is inconsistent with thinking that bank fraud is not wrong.

My response departs from others in the literature (Schroeder, 2008; Horgan and Timmons, 2009; Schwartz and Hom, 2014; Baker and Woods, 2015) in that I think that there needs to be no substantive explanation of the inconsistency of normative judgments. All of the relevant components of normative inconsistency can be treated as primitive without much cost to the expressivist.

In the course of this paper, I will present and defend four claims. First, I will propose that inconsistency between normative judgments can be analyzed into a system of claims

---

[1]The final publication is available at Springer via http://dx.doi.org/10.1007/s10677-015-9670-9

about the behavior of our attitudes and the norms governing them. Second, I will argue that the relevant behavior of our attitudes can be explained, as much as it needs to be, by a functional interpretation of mental states. Third, I will propose that expressivists can regard the relevant norms as primitive. Fourth, I will argue that considerations of the parsimony of a psychological theory's normative consequences shouldn't figure into the (expressivist's) evaluation of that theory. Since the core of expressivism is a psychological theory, the relevance of its normative implications is limited.

The plan is as follows. I will start by presenting the problem – the 'Negation Problem' – as it is conceived of by Unwin and Schroeder. The problem is complex, and connected with a range of other issues facing expressivists. Next, I will try to explicate the notion of inconsistency that is at the heart of the problem, and I will argue that we can decompose inconsistency into descriptive and normative components. After providing my account of the nature of attitudinal inconsistency, I will investigate how expressivists might explain the inconsistency of normative judgments. The problem with normative judgments is especially acute given the variety of complex forms that they can take. I will argue that the account that I have given can be extended to accommodate this variety without a problem.

The arguments that I offer will ultimately rely on thorough-going expressivism. Thorough-going expressivists think that all judgments about norms are noncognitive states. There is nothing special about moral norms. According to the view, sentences such as "one should not believe both a proposition and its negation" do not express beliefs.

Since my response depends on thorough-going expressivism, the realist critic may feel that the expressivist has not provided an adequate explanation because she has not accounted for the genuine normative relations between our attitudes. These criticisms may stick to merely moral expressivists. Given that normativity forms part of what is said to require explanation, there is the possibility of disagreement between realists and thorough-going expressivists over what actually needs to be accounted for. So be it. If expressivists can answer the challenge as that challenge is best conceived in light of their own doctrine, they will have warded off any reasons to doubt it.

2 The Negation Problem

Expressivists think that the meanings of normative words are tied to their use in expressing certain attitudes. Unlike ordinary referential words, normative words do not contribute to the representational content of the sentences in which they occur. Instead, they contribute to the kinds of uses to which those sentences can be put. The occurrence of the predicate 'wrong' in (1) ensures that the sentence can be used to express an attitude that is linked with motivation.

   (1) Bank fraud is wrong.

According to expressivism, this sentence does not express a belief. Let us say that the attitudes, whatever they are, that we express with sentences involving normative language are *normative judgments*. For present purposes, we can say that expressivists think that straightforward, predicative normative judgments, such as the judgment expressed by (1), are a species of *disapproval* rather than belief. According to expressivism, it is part of the meaning of (1) that sincere assertions of it express a kind of disapproval of bank fraud.

   Normative words are not limited to simple predicative contexts as in (1). They are often embedded within other semantic constructions, such as in (2).

   (2) Bank fraud is not wrong.

Semantic consistency demands that 'wrong' in (2) makes the same semantic contribution as 'wrong' in (1). Expressivists cannot say that it contributes to the sentence's representational content. Instead, they should say that it helps determine the kind of attitude that the sentence can be used to express. If (1) is not used to express belief, it is implausible that (2) should be. However, (2) is plainly not used to express an attitude of disapproval, either. On the contrary, the attitude we express with (2) is (in some sense) the opposite of disapproval. Again (for simplicity) we will say that (2) is used to express an attitude of *tolerance*.

One problem facing expressivists is to explain how it is that the meanings of complex sentences like (2) are determined by their parts. Though this is most certainly an important problem, I will leave it aside and just focus on a separate question. It is clear that there is a special relation between the attitudes that sentences (1) and (2) are used to express: they appear to be inconsistent with each other. The question I will focus on is how to make sense of this fact.

As a first pass, realists can point to inconsistent representational contents in order to explain attitudinal inconsistency. The mental states expressed by these sentences are inconsistent because they represent the world as being in inconsistent ways. Expressivists do not think that normative judgments have any special normative representational content. Consequently, they don't think that the attitudes are inconsistent in the sense of having inconsistent representations. I will follow Baker and Woods (2015) in using the term 'discordance' to refer to the special sort of not-necessarily-representational inconsistency that attitudes like (1) and (2) have.[2]

Expressivists face the question: why are the mental states that we express with these sentences discordant? Nicholas Unwin (1999; 2001) and Mark Schroeder (2008) have observed that expressivists find it difficult to provide substantive explanations of this discordance. For the realist, sentences (1) and (2) express propositions that represent the world as being different ways. For the expressivist, they do not. If they are discordant, it is not because of the inconsistency of the objects of the attitudes, but because of the attitudes themselves. Unwin and Schroeder suggest that by giving up on the thought that normative language contributes to representational content, expressivists give up anything that might be a help in explaining attitudinal discordance.

Unwin writes:

> It needs to be a logical law (or, at the very least, a 'quasi-logical law' in some sense or other) that we cannot properly hold both [disapproval] and [toleration]. However if [disapproval] and [toleration] are merely understood to denote ex-

---

[2]I take it that Gibbard's term 'disagreement' refers to roughly the same phenomenon.

pressions of two different kinds of attitudes (and that is all) then, as we have just observed, this fact will be lost from sight. It becomes at best a mere brute fact that the attitudes conflict with each other, with no internal complexity that could explain why. (Unwin, 1999, 342)

Schroeder echoes:

It should be very clear that [on the present proposal] we have left completely unexplained and apparently inexplicable why 'murder is wrong' and 'murder is not wrong' are inconsistent... [tolerance and disapproval] can't simply be interdefined, which leads to the conclusion that they are distinct and unanalyzable attitudes. But if they are, then why on earth is it inconsistent to hold them toward the same thing?... Assuming that disapproval and tolerance of murder are inconsistent is taking for granted everything that expressivists need to explain. (Schroeder, 2008, 47-48)

Instead of explaining the discordance of tolerance and disapproval, expressivists such as Blackburn (1985), Gibbard (1992), and Timmons and Horgan (2006) have just helped themselves to it. They have done away with the resources necessary for providing an adequate explanation. They hold that 'not' makes no contribution that reflects anything inside the content of the attitude, it simply allows for a different attitude to be expressed. Since the contribution isn't structural, the structure cannot be used to explain the discordance. And if the structure of the content can't be used to explain this discordance, it is hard to see what else could be.

In any case, it has been recognized that there is a deeper sense of discordance between the attitudes expressible by (1) and (2) than the mere inconsistency of representational content accounts for.[3] Not all attitudes that have inconsistent contents are discordant in this deeper sense. I may be both uncertain whether Prague is the capital of Hungary, and uncertain whether Prague is not the capital of Hungary, without discordance. I can fear missing an important talk that I'm to give, and also fear getting there and having to give it. So while

---

[3]See, for instance, Schroeder (2008).

representational inconsistency may help explain the discordance of attitudinal contents, it is insufficient by itself. Discordance goes deeper than representational inconsistency, and the fact that representationally inconsistent beliefs are discordant in this deeper sense requires explanation.

I do not think that expressivists have much trouble in explaining the relations of discordance between the attitudes expressed by (1) and (2). Given plausible views about the nature of discordance, the existence of attitudes fitting the profile of tolerance and disapproval isn't far-fetched. Ultimately, I suspect that the discordance of tolerance and disapproval (perhaps along with the discordance of belief) is primitive, and not open to substantive explanation.

3   What is Discordance?

Before we try to explain why some normative judgments are discordant, we must get a handle on what discordance amounts to. I propose that discordance between two attitudes is equivalent to a (perhaps very large) set of simple claims that are either purely about how the attitudes actually interact with each other and how they should relate to each other.[4] These behavioral relations and governing norms *constitute* attitudinal discordance.

This equivalent set of claims is broad and complex, and I will not try to do it justice. Formulating the simple claims without the language of discordance would be quite laborious. In this section, I hope instead to give a few examples of the kinds of simple claims that belong to the descriptive and normative sides.

The claims on the descriptive side include (among other things) claims about our inferential dispositions, the phenomenology of reasoning, and the roles that such attitudes have in discourse. For instance, we tend not to have discordant attitudes simultaneously. When we have discordant attitudes, we feel a strong tension that pushes us into deliberation, or else weakens our commitment to one or the other of the attitudes. The tension often has a certain formal flavor: anyone who has ever been moved by a deductive argument can call it to mind. We feel like we are making a sound inference when we move from "predatory

---

[4]The exact set of claims is also likely to be vague.

lending is wrong" and "if predatory lending is wrong, then strategic defaulting is wrong" to "strategic defaulting is wrong". Let us say that any pair of attitudes that interact in the right way are *descriptively discordant*.

The claims on the normative side primarily serve to validate the interactions between discordant attitudes. The norms under consideration are norms of rationality. For instance, we should not hold both "bank fraud is wrong" and "bank fraud is not wrong" at the same time. If we do, we should feel some pressure to change our minds about one or the other. Many of the claims that constitute descriptive discordance are not mere matters of contingent psychology – they are the correct properties for our attitudes to have. Let us say that any attitudes that are supposed to interact in the right way are *normatively discordant*.

I claim that there is nothing more to attitudinal discordance than descriptive and normative discordance. At the very least, until the critics specify what else is involved in discordance, the expressivist can solve the problem if she satisfactorily accounts for both the descriptive and normative discordance of disapproval and tolerance. She will have provided two different ways in which the attitudes clash where it was suggested that no kind of clash was possible.

One of the primary lessons to be drawn from van Roojen's work on expressivism and the Frege-Geach problem is that there are multiple ways in which attitudes may clash, and not every sort of clash involves the same sort of discordance that is displayed in (1) and (2). Van Roojen argues that "an expanded notion of inconsistency applicable to the contents of attitudes cannot easily be defined in terms of norms governing the rationality or irrationality of accepting certain attitudes. For the rational connections among attitudes do not neatly mirror the logical connections between contents" (1996, 312-13). For instance, Moore's paradox demonstrates a kind of clash: the belief that bank fraud is difficult and the indexical belief that I don't believe that bank fraud is difficult clash in a different sense from (1) and (2). One difference is that the clash doesn't persist in the context of suppositions. I can consistently suppose that bank fraud is difficult and suppose that I mistakenly believe that it is not. Van Roojen suggests that the relevant sort of clash involved in judgments and their negations can't be explained in terms of facts about rationality, but we haven't explored

the space of possible norms sufficiently to justify that conclusion. I don't have a story to tell about the differences between these two varieties of discordance, but I am committed to thinking that whatever they are, they can be spelled out in terms of either behavioral differences or normative differences.

4   The Discordance of Tolerance and Disapproval

The challenge before the expressivist is to explain why mental states like those expressed by "bank fraud is wrong" and "bank fraud is not wrong" are discordant. Realists think that these attitudes are beliefs. If the Negation Problem is truly problematic, it must be because expressivists both need to explain the discordance of disapproval and tolerance and they cannot. Whereas the discordance of beliefs is deeply tied to inconsistency of representational content, the discordance between tolerance and disapproval cannot be. I will argue that expressivists do not need to provide a substantive account either of the descriptive or of the normative discordance of the attitudes. The reasons why they do not need a substantive account are a bit different for descriptive and normative discordance. For this reason, I will employ a divide-and-conquer strategy by offering a separate treatment of descriptive and normative discordance.

4.1   Accounting for Descriptive Discordance

Descriptive discordance is a fact about psychology; there is surely some neurological explanation for the discordance of our discordant attitudes that eschews talk of propositions and representations, so descriptive discordance doesn't obviously require representational contents. We may not understand this explanation until we have a much better grasp on how our brains function. Even if we never actually understand it, we can be sure that there is an explanation.

I do not think that Unwin and Schroeder intended to raise doubts about the possibility of engineering noncognitive states that are descriptively discordant. Given that the interactions that characterize descriptive discordance are not themselves sufficient to ensure that the

states have representational content, this should be possible. Not all aspects of cognition have a philosophical explanation. Some attitude clashes may simply be a product of the way that our brains are wired – they may have no deeper philosophical explanation than the existence and behavior of headaches. The fact that we cannot occupy each of two attitudes simultaneously may just turn out to be a matter of the details of their implementation.

It might be regarded as a cost if expressivism were forced to postulate hybrid states that share something with beliefs (discordance) and something with disapproval (non-representationality and a role in motivation). It would certainly complicate our theories of human cognition. Nevertheless, it is hard to assess the cost of this complexity from the armchair. How bad complexity is for a psychological theory depends in part on whether we should expect human cognition to be complex in that way. If complexity is to be expected as a natural result of the evolutionary pressures that shaped our brains, then it is not especially costly for a psychological theory to be complex.

Normative judgments are clearly very closely related to beliefs. It is plausible that the same sorts of mechanisms involved in producing beliefs might have evolved, without drastic modification, to produce nonrepresentational motivational judgments. If evolutionary pressures could easily adapt one kind of mental state to a new and useful purpose, then the cost of positing both the original state and its adapted form is not too high. It is especially plausible that normative attitudes arose in such a way. They clearly play an important role regulating social interactions for our benefit, and does not seem vital to that role that the attitudes have any particular representational content (Gibbard, 1992). If we evolved to have certain judgments about morality because they were useful to us in stabilizing social interactions (Street, 2006; Joyce, 2007), it is not obvious that we should expect them to be representational.

Neurological facts can only explain the discordance of token attitudes. The particular implementations of our attitudes might explain why it is that *our* attitudes of tolerance and disapproval behave in the way characteristic of discordant attitudes, but it leaves open the possibility that the attitudes of other creatures do not. The discordance of tolerance and disapproval is not parochial. It holds for all creatures whatsoever. For this reason, there is

bound to be another, more general, explanation.

The general explanation of why attitudes of tolerance and disapproval are descriptively discordant is straightforward. The reason why the attitudes always play a certain psychological role is because *playing that role is part of what it is to be those attitudes*. I do not mean to suggest that we can find some essential core functional purpose of the attitudes – that they are planning states or wishes for how the world might be – that will explain discordance. My claim is more radical: we can just build it into our concept of what it is to be attitudes of tolerance and disapproval that they are descriptively discordant. They would then be descriptively discordant for no further reason. Surely, the concept *non-representational motivational attitudes that behave in accordance with descriptive discordance* is coherent. My proposal is that we should regard our concepts *tolerance* and *disapproval* as requiring descriptive discordance in this way.

Some relations between kinds of attitudes will be primitive. It is widely recognized that our beliefs have a generalizable relation to desires and intentions: we are disposed to intend to do those actions that we believe will best satisfy our desires. Someone might wonder why it is that these three attitude kinds, belief, desire, and intention, which lack internal structure to explain these roles, nevertheless play them. The reason is that it is part of what it is to be a system of beliefs, desires, and intentions involves exhibiting such relations. If particular mental states did not relate to each other in those ways, those particular mental states would not be beliefs, desires, and intentions. In exactly the same way, if attitudes did not clash in the descriptive sense, they would not be attitudes of tolerance and disapproval with the same content. It is simply characteristic of tolerance and disapproval that particular mental states must be descriptively discordant with each other in order for them to count as attitudes of tolerance and disapproval.

4.2   Accounting for Normative Discordance

I have so far argued that expressivists can offer a trivializing account of the descriptive discordance of tolerance and disapproval. Part of what it is for a mental state to be an attitude of tolerance is that it is descriptively discordant with certain other attitudes of

disapproval. Now I'll turn to normative discordance. If we are to account for the normative discordance of discordant attitudes, we must explain why the attitudes are subject to the norms that they are. We might try to offer the same explanation of normative discordance that I offered for descriptive discordance. That explanation would rest on the thought that *tolerance* and *disapproval* are normatively thick concepts, and that part of what it takes for attitudes to be attitudes of tolerance and disapproval is to be governed by the relevant norms. The fact that two attitudes are normatively discordant is not posterior to the fact that they are attitudes of disapproval and tolerance with their respective contents – the fact that they are those attitudes with those contents lies partially in the fact that they are normatively discordant.

This explanation can't amount to a full response to the problem by itself. It is possible to provide an explanation for why mental states behave in the ways that they must in order to qualify as attitudes of tolerance and disapproval. How it is that such states come to exist is a question of engineering. If we were to simply say that, in order for two attitudes to be attitudes of tolerance and disapproval, they must be normatively discordant, we haven't yet explained how it is that actual mental states get to be normatively discordant (and hence get to be attitudes of tolerance and disapproval).

Instead, I'll instead suggest that *expressivists should not be committed to providing any explanation of the norms governing tolerance and disapproval.* Allan Gibbard (2003) has already proposed that we take attitudinal discordance as primitive, and on my analysis, this includes taking the norms as primitive. Jamie Dreier (2006) proposes that the problem with doing so is that "if we have no explanation for why some states can be disagreed with and others cannot, then for all we know it could be that the correct explanation is not amenable to expressivism." (716) This worry only makes sense if we interpret Gibbard as proposing that we take discordance as a mere theoretical primitive, rather than a genuine metaphysical or normative primitive. In the latter case, there will be no explanation at all for why some states are discordant and others are not. I propose that the attitudes which are normatively discordant are normatively discordant for no further reason. Most normative theories bottom out somewhere; I suggest that expressivists include the norms governing

normative discordance at the bottom level.

On most plausible normative theories, there will be a more than a few distinct primitive norms. Some primitive norms will concern which actions are right. Some will concern what is fundamentally valuable. Some will concern what we should believe on what evidence. Some may concern which kinds of desires we should have. Some will concern how we should act, given our evidence and our desires. And some, I propose, will concern the proper relation between attitudes of tolerance and disapproval. Given the number of primitive norms that are necessary in order to accommodate all of the different normative judgments we're committed to, the cost of postulating a new primitive norm to account for the discordance of tolerance and disapproval can't be too high. Nevertheless, it may be thought that expressivism should regard novel primitives as a theoretical cost. In the remainder of the paper I will try to argue that assessing this cost is delicate. Even some egregious violations of parsimony won't warrant giving up on the core claim of expressivism.

5   The Infinite Hierarchy of Attitudes

The Negation Problem is particularly worrisome because it generalizes. In this section, I will start by addressing how to handle the generalized problem. I will suggest that the explanation of descriptive discordance can be extended to handle the full problem. The explanation of normative discordance can be extended as well, but extending the account of the normative discordance will require us to posit a huge number of primitive norms. While this may seem problematic, I will argue there are good reasons for thinking that normative parsimony isn't relevant to the assessment of the core claim of expressivism – normative considerations don't bear on how we should think about scientific theories, and expressivism is primarily a theory about psychology. Then I will explain how it is these considerations relate back to parsimony: expressivists should think there are at least two different kinds of parsimony, and our intuitions suggest thinking that the kind of parsimony at issue isn't relevant to psychological theory choice. Therefore, the fact that expressivism requires accepting a large number of primitive norms, if it does, isn't a reason not to accept

12

the core claim of the theory.

5.1   The Problem

The expressivist's basic account suggests that normative judgments are attitudes of disapproval. We had to add that "bank fraud is not wrong" expresses a different attitude once we recognized that the negated sentence doesn't itself express an attitude of disapproval. We said that it expressed tolerance instead. We will need to repeat this for more complicated varieties of normative judgments. Take the judgment that "bank fraud is either wrong or bank fraud is not wrong". While to say "bank fraud is wrong" is to express some attitude of disapproval and "bank fraud is not wrong" is (perhaps) to express an attitude of tolerance, the disjunction expresses neither a straightforward attitude of disapproval nor tolerance.

It appears as if we'll need to postulate a novel kind of attitude for each kind of judgment with a different logical profile. Expressivists have responded by describing the relations that hold between all of the attitudes that they postulate. Thus, we can say what kind of attitude "bank fraud is either wrong or bankers have it coming" expresses by saying how that attitude relates to the attitudes expressed by "bank fraud is either wrong" and "bankers have it coming". There is something systematic about this. Each complex structure that a normative judgment might take will correspond with a different functional characterization, and these functional characterizations are easy to predict given the structures of the sentence that we use to express them.

Despite being easy to predict and characterize, expressivists are still forced to postulate a range of special attitudes distinct from disapproval in order to capture the full logical complexity that our normative judgments display. We can't resolve the Negation Problem by postulating one new kind of attitude. We must postulate an infinite hierarchy of attitudes along with primitives describing the relations between them. This is the Hierarchy Problem.

5.2   The Hierarchy and Descriptive Discordance

If there is no problem characterizing the functional profile of each of the postulated attitudes in the hierarchy, as Timmons and Horgan (2006) and others have adequately done, then we

can repeat the same approach to account for the discordance between all members of the normative attitude hierarchy. What it is to be a set of judgments to be the judgments that "bank fraud is either wrong or bankers have it coming", "bank fraud is wrong", and "bankers have it coming" is, in part, for them to relate in some particular way.

I see two possible problems with extending the approach to descriptive inconsistency to all of the attitudes in the hierarchy. First, there is a potential problem of engineering: we are actually capable of having a tremendous number of attitudes in this hierarchy – would that be plausible if they really were separate kinds of attitudes? How does our brain give rise to such a huge variety of attitudes? Second, there is the question of how we could possibly have come to have the capacity for so many different attitudes, even if we do.[5] Evolution could not have selected for all of them individually. If there is not some systematic core, we should only expect to be able to have a small number of them.

Neither of these two problems should worry us too much, because they both bear equally on the infinite hierarchy of beliefs that we are capable of having.

Beliefs form a hierarchy just as much as moral attitudes. We are able to form beliefs in a vast number of different propositions and each kind of belief will have a slightly different functional profile. This capacity shouldn't be taken for granted. Jerry Fodor and Zenon Pylyshyn (1988) famously argued that we need to posit a language of thought in order to account for it. According to their proposal, our attitudes are built up by combining basic concepts that collectively determine the content of the attitudes in which they figure in something like the way that the meaning of individual words determine the meanings of the sentences in which they occur. However, it remains controversial whether there is a language of thought and presently there is no consensus on how our brains generate the functional hierarchy of beliefs. This is important because without knowing how our brains manage to produce the hierarchy of beliefs, we do not know whether we will need a separate explanation for our capacity to adopt the attitudes in the hierarchy of normative judgments.

It is hard to see why it should be easier to account for our capacity to have the whole hierarchy of beliefs. The existence of a highly structured set of abstract of objects outside

---

[5]Thanks to an anonymous reviewer for pointing out this problem.

of our brain does not explain how our brain manages to have beliefs that take them as objects. It is deeply plausible that what it is to be propositional attitudes is to be a kind of attitudes whose functional profiles happen to be fruitfully demarcated by the propositions they are said to be about.[6] Content is an epiphenomenon of functional role. What makes a belief count as a belief in a particular proposition is that there is a natural mapping from broad functional profiles to propositions that takes that belief's functional profile to that proposition. Propositions do not themselves do anything to explain our productive and systematic capacity for beliefs.

If this is the right story about how beliefs get to be about propositions, then the main difference between the normative judgment hierarchy and the hierarchy of beliefs is that the latter can be fruitfully demarcated by an association of particular states with a certain kind of abstract object. It isn't anything that makes an obvious difference in plausibility from the perspective of neural engineering. There is nothing obviously problematic about postulating a well-ordered hierarchy of distinct attitudes that have systematic functional roles that don't align neatly with any abstract objects.

So how did the hierarchy evolve, if we could only ever had need for a finite number of attitudes in that hierarchy? We cannot say for sure, but here is one reasonable account. There is something about our brains that generates our capacity to have a tremendous variety of beliefs. That capacity was selected for because we face a complex and incredibly varied environment that we must adapt to. At some point along the way, the capacity to make straightforward normative judgments also became useful to our ancestors. As part of the process of duplication and modification that is prevalent in evolution and an important part of the evolution of the brain (Striedter, 2005), the same mechanisms involved in belief generation were co-opted for the purpose of producing normative judgments. Since those mechanisms already allowed for a hierarchy of beliefs, they also naturally also produced a hierarchy of normative attitudes. Such a capacity may have been useful, or it may have just come along as an evolutionary spandrel.

Is it very bad to posit two infinite hierarchies of attitudes when one could posit one?

---

[6]See, for instance, Brian Loar 1981.

I don't think so, especially if the same cognitive mechanisms employed in generating one could have been adapted to generate the other. It should not be too surprising if we have two different sorts of similar attitudinal hierarchies, one of which can be fruitfully associated with propositions and one of which cannot.

Consequently, the approach that I took to explain descriptive discordance generalizes.

5.3 The Hierarchy Problem and Normative Discordance

The Hierarchy Problem does not particularly complicate the explanation of descriptive discordance. How does our explanation of normative discordance fare? I suggested that noncognitivists can posit the relevant norms as primitives. Now, however, it looks like that doing so will require an infinite number of primitive norms.

There are two avenues for dealing with this concerning development. One is to meet the problem by giving a general explanation of discordance that will apply simultaneously to each pair of discordant complex normative judgments. In order to do this, it is necessary to attribute some structure to these attitudes and to find some way of making generalizations over the hierarchy, so that only a handful of normative primitives are needed.

The second is to deny that there is a problem in the first place with postulating a large number, even an infinitely large number, of primitive norms. I think both avenues are promising. The first avenue may be more satisfying – nevertheless I shall only consider the second, because it provides a deeper resolution of the issue. While we may be able to find a parsimonious version of expressivism, looking for such a theory would suggest that parsimony should bother us in the first place. It shouldn't.

Before proceeding in the next few sections to make my case that we should not be too bothered by the need to postulate an infinite hierarchy of norms, let me address the worry that any metaethical theory that required an infinite number of primitive norms would make it a inexplicable mystery how we were able to recognize them.

The fact that the infinite primitive norms cannot all be reduced to a single norm doesn't mean that they can't be systematically generated. We can provide a systematic account of what kinds of normative judgments there are, what kinds of functional roles they have,

and what normative relations hold between them and others. For any normative judgment, for instance, there is another judgment that is its *negation*, that is discordant with it both descriptively and normatively. For any two normative judgments, there is another normative judgment that is their *conjunction*, that is discordant with the negations of those judgments. The hierarchy is organized in such a way that it can be systematically generated.[7] Thus, the fact that an infinite hierarchy of primitive norms exists wouldn't make our capacity to recognize the primitive norms miraculous. If there is a problem with the norms, it is independent of our ability to recognize them.

5.3.1 Scientific Theories and Normative Evidence

Expressivists might get away with a few additional primitives, but it would look bad if they had to posit an infinite hierarchy of primitive norms. It isn't nearly as bad as it looks. Over this and the next few sections, I will argue that the kind of parsimony that expressivism lacks isn't relevant to our assessments on whether the core claim of that theory is correct.

Expressivism is primarily about the psychology of judgments about normativity. The core claim of expressivism is that normative judgments behave in the manner of disapproval rather than in the manner of beliefs and that as such, they lack distinctively normative representational contents. We are all pretheoretically committed to thinking that "insurance fraud is wrong" and "insurance fraud is not wrong" express discordant attitudes. And expressivist doctrine tells us that "insurance fraud is wrong" and "insurance fraud is not wrong" express attitudes that lack the behavioral roles necessary to count as beliefs. So expressivists are committed to thinking that attitudes other than beliefs can be discordant, but I consider this an implication of their view, rather than a proper part of it.

My skepticism of the significance of normative parsimony is related to the phenomenon that Alex Barber (2013) has called "science's immunity to moral refutation":

> The phenomenon of moral immunity is this: while the scientific finding (plus
> background assumptions and logic) may count as evidence against the moral

---

[7]How can it be systematically generated while still being primitive? The norms themselves are primitive, in that they can't be explained in terms of other norms, but they fit into a pattern that allows us to give a concise summary of what norms there are.

assumption, it seems deeply misguided to regard the moral assumption (plus this same background) as evidence against the scientific finding. (Barber, 2013, 635)

Here is Barber's thought applied to an example. Many people strongly believe that it is deeply wrong to torture animals. It is, let's suppose, only deeply wrong to torture animals if they are conscious. We shouldn't let these moral stances affect what we believe about whether animals are conscious. That is a scientific question. The deep feeling that it is morally wrong to torture animals may justify the judgment that it is wrong to torture animals, but it does not justify the belief (in the absence of other evidence) that animals are conscious. If a researcher claims to have found strong evidence that dogs lack the cognitive structures necessary to be conscious, we should not allow the fact that persistent feeling that it is wrong to vivisect dogs to provide any counter-evidence. In Barber's words, "moral claims as such cannot figure as evidence within the scientific domain." (636)

Barber recognizes that this poses a puzzle for the realist. If we are licensed to believe in certain moral claims, and those moral truths imply certain scientific claims, why can't we treat the moral claims as evidence for the scientific claims? Barber thinks that it is especially easy to explain the scientific immunity if one is an expressivist. For the expressivist, the epistemic gap parallels a difference in the metaphysical substantivity of the relevant facts.

This epistemic gap shows up in Cian Dorr's (2002) Wishful Thinking Problem. Dorr suggests that if expressivism were true, expressivists would be committed to thinking that many seemingly reasonable arguments are irrational. For any argument with moral premises is an argument from our feelings, and:

> It is often rational to modify your views about one part of the world so that they cohere with your views about the rest of the world. It is irrational to modify your views about the world so that they cohere with your desires and feelings. That's wishful thinking! (Dorr, 2002, 99)

Like Dorr, I agree that expressivists should avoid wishful thinking. Unlike Dorr, I don't see this as a problem for the expressivist.[8] In fact, I think that it is an advantage, because the

---

[8]I think that James Lenman (2003) has the right response to it. This response involves denying that expressivists are committed to wishful thinking while agreeing that this would be problematic.

epistemic gap that Dorr relies upon helps to explain science's immunity to moral refutation and it provides the expressivist with a justification for discounting considerations of normative parsimony.

My response to the Hierarchy Problem depends on the thought that we shouldn't let the parsimony of normative commitments dictate what we believe about scientific theories. I will need to extend Barber's proposal in two different ways. First, Barber's proposal concerns specifically *moral* norms, while the norms of normative discordance are norms of *rationality*, so we need to extend the idea to include the immunity of scientific claims from revision on the basis of norms of rationality as well. Second, Barber's proposal concerns the use of normative claims in support of scientific claims – it doesn't explicitly say how to evaluate the use of considerations of normative parsimony. We'll need to extend the idea to include the thought that considerations of normative parsimony should not be desiderata for deciding between scientific theories.

Barber's proposal is suggestive, but it will be far more so if we can figure out why it should be true. In the next few sections, I will suggest that expressivists should think that it is true, and I will an offer an account of the phenomenon that can be used to deal with the expressivists' apparent problem with parsimony. The upshot is that we should not use considerations of normative parsimony to reject the core claim of expressivism.

5.3.2 Metaphysical and Normative Parsimony

Parsimony is a general label for a kind of consideration that occurs in many different places. Expressivists should divide parsimony into at least two main varieties.

First, consider the kind of parsimony employed by detectives, astronomers, and metaphysicians to choose among competing theories. This kind of parsimony revolves around how complex the theory presents the world as being. I call it *metaphysical* parsimony, because the degree of metaphysical parsimony of a theory is determined by the number of distinct representational ontological or ideological commitments that theory has. It is, in effect, a measure of how complex reality would have to be in order for that theory to represent it correctly.

It is important to note that metaphysical parsimony of a theory doesn't depend upon the complexity of the beliefs involved in accepting the theory. Very simple theories with very few representational commitments may be given complex formulations. How exactly we should go about calculating the representational commitments of a theory is a difficult matter, but it is clear that only genuine representational commitments should be included. We should not count illusory or eliminable posits against the parsimony of a theory. If we find that we can, in principle, reconstruct physics without reference to numbers, we should not hold the use of numbers *as a shorthand* against the parsimony of a physical theory.

On the other hand, according to expressivists, the kind of theoretical parsimony counted by ethicists and epistemologists doesn't revolve around how complex the theory makes the world out to be because expressivists think that the complexity of normativity doesn't have any close connection with metaphysical complexity. Expressivists take judgments about normative properties to be disguised attitudes of disapproval. Just as the complex emotions that a scientist's theory might give rise to don't detract from the metaphysical parsimony of the scientist's theory, the normative commitments of expressivism do not detract from the metaphysical parsimony of that theory. Acceptance of a theory that is parsimonious in the normative sense involves having parsimonious attitudes of disapproval. There is no distinct notion of the commitments of a normative theory – there is only the complexity of the states required to accept it. The difference between the two kinds of parsimony is striking.

5.3.3   The Significance of Normative Parsimony

If expressivists are right, purely normative theories don't differ in terms of metaphysical parsimony. Since the metaphysical parsimony of a theory depends on the complexity of its representational commitments, normative parsimony must be something different. If anything, normative parsimony is a matter of the systematicity of the noncognitive attitudes involved in accepting a normative theory. If we are to favor more parsimonious normative theories over less parsimonious ones, it must be because parsimony provides a reason for (or against) disapproval. In particular, we should aim to have parsimonious normative attitudes. Why should we care if our normative attitudes are parsimonious?

Parsimony is used to decide between competing views in first-order ethics. Ethicists try to subsume intuitions about individual cases under unified theories in an attempt to make their theories more parsimonious. Expressivists could repudiate this practice, but it is widely agreed to be preferable if we can develop expressivism in a way that doesn't have a radical effect on our first-order practices. So I will suppose that some kind of parsimony is a legitimate criterion to use in deciding between first-order normative theories. We need an account of why is it that normative claims provide support for other normative claims in order to accommodate basic philosophical practice. Until we are sure that such an explanation won't undermine the prima facie irrationality of wishful thinking, the expressivist cannot ignore the Hierarchy Problem on the grounds that it involves wishful thinking.

In summary, expressivists are committed to distinguishing metaphysical from normative parsimony and to thinking that normative parsimony is relevant for normative theory choice. They should not be committed to thinking that normative parsimony is relevant for psychological theory choice. Given that normative and metaphysical parsimony are so different, we can't infer that that normative parsimony is relevant to psychological theory choice from the fact that metaphysical parsimony is. We must actually look to our intuitions about particular cases. In order to accommodate the intuitions shared by Barber, Dorr, and Lenman while respecting the role of considerations of parsimony in ethics, I propose that normative parsimony is a consideration that applies to the choice of normative theories. Normative considerations don't provide evidence for scientific theories.

This distinction between considerations of normative and metaphysical parsimony give us grounds for thinking that expressivists shouldn't be worried about the normative parsimony of their theories, no matter how bad it gets. Such parsimony isn't entirely irrelevant, because it matters when it comes to evaluating normative theories, but it is irrelevant to what psychological views we should accept. Since expressivism is at its core a psychological theory and not a normative theory, we shouldn't take normative parsimony into account.

6   Conclusion

I have offered sketches of trivializing explanations for the normative and descriptive discordance of tolerance and disapproval. I made four primary claims. First, attitudinal discordance is a combination of descriptive and normative discordance. Second, descriptive discordance can be accounted for with a functional characterization of mental states. Third, expressivists should not be too opposed to thinking that the normative component of the discordance of tolerance and disapproval is primitive. Fourth, considerations of normative parsimony shouldn't figure into selecting a psychological theory. I left out many details concerning the precise form of the explanation, but there does not appear to be any principled barrier to filling them out. Consequently, I conclude that the Negation Problem is not a problem for expressivists.

The arguments that I have given have relied on thorough-going expressivism and consequently I have not done anything to help expressivists provide explanations of any realistically interpreted normative discordance. Expressivists, of course, think that this is unnecessary. Realists should think that the account that I have offered fails to get to the heart of the matter. In this way, I think that Unwin and Schroeder are right to think that from their perspective, expressivists do face a significant problem. Their challenge lacks dialectical force; since expressivists are blind to the phenomenon that realists think needs to be explained, there is no problem from her perspective.

We have a stalemate. For realists, there may yet be a Negation Problem – the expressivist has not even tried to explain the *genuine* norms that govern tolerance and disapproval. Realist metaethical views suggest that there are norms that need to be explained. For thorough-going expressivists, however, the absence of such explanations is not a problem for expressivism.

Acknowledgements

References

Derek Baker and Jack Woods. How Expressivists Can and Should Explain Inconsistency. *Ethics*, 125(2):391–424, 2015.

Alex Barber. Science's Immunity to Moral Refutation. *Australasian Journal of Philosophy*, 91(4):633–653, 2013.

Simon Blackburn. *Spreading the Word*. Oxford: Clarendon Press, 1985.

Cian Dorr. Non-Cognitivism and Wishful Thinking. *Noûs*, 36(1):97–103, 2002.

Jamie Dreier. Disagreeing (About) What to Do: Negation and Completeness in Gibbard's Norm-Expressivism. *Philosophy and Phenomenological Research*, 72(3):714–721, 2006.

Jerry Fodor and Zenon Pylyshyn. Connectionism and Cognitive Architecture. *Cognition*, 28(1):3–71, 1988.

Allan Gibbard. *Wise Whoices, Apt Feelings*. Cambridge: Harvard University Press, 1992.

Allan Gibbard. *Thinking How to Live*. Cambridge: Harvard University Press, 2003.

Terry Horgan and Mark Timmons. Cognitivist Expressivism. In Terry Horgan and Mark Timmons, editors, *Metaethics after Moore*. Oxford: Clarendon Press, 2006.

Terry Horgan and Mark Timmons. Expressivism and Contrary-Forming Negation. *Philosophical Issues*, 19(1):92–112, 2009.

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

James Lenman. Noncognitivism and Wishfulness. *Ethical Theory and Moral Practice*, 6(3): 265–274, 2003.

Brian Loar. *Mind and Meaning*. Cambridge: Cambridge University Press, 1981.

Mark Schroeder. *Being For: Evaluating the Semantic Program of Expressivism*. Oxford: Clarendon Press, 2008.

Jeremy Schwartz and Christopher Hom. Why the Negation Problem Is Not a Problem for Expressivism. *Noûs*, 48(2):824–845, 2014.

Sharon Street. A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies*, 127(1):109–166, 2006.

George Striedter. *Principles of Brain Evolution*. Sunderland: Sinauer Associates, 2005.

Nicholas Unwin. Quasi-Realism, Negation and the Frege-Geach Problem. *The Philosophical Quarterly*, 49(196):337–352, 1999.

Nicholas Unwin. Norms and Negation. *The Philosophical Quarterly*, 51(202):60–75, 2001.

Mark van Roojen. Expressivism and Irrationality. *The Philosophical Review*, 105(3):311–335, 1996.