Mark Schroeder
University of Southern California
July 1, 2009

how not to avoid wishful thinking

In 2002, Cian Dorr offered a new challenge to metaethical noncognitivism: the wishful thinking problem. Based on considerations from epistemology, it is quite distinct from the usual problems associated with the title 'Frege-Geach' which assail noncognitivists in the philosophy of language, and in contrast to the Frege-Geach Problem, it poses a challenge for noncognitivist views in *ethics* which does not arise for similar views about, for example, probability judgments, epistemic modals, or conditionals. But after an immediate round of attempted solutions from sympathists (James Lenman) and critics (David Enoch) of noncognitivism alike, the ripples in the pond have somewhat quieted.

The aim of this paper is to critically assess the state of play with respect to the wishful thinking problem by putting Enoch's and Lenman's attempted solutions into context with one another, and placing them in a space of possible solutions. The morals are (1) that Enoch's solution is very unpromising, (2) that Lenman's solution takes on extraordinary commitments that from many points of view are problematic in themselves, and (3) that doing better than Lenman's solution may require noncognitivists to develop better tools – in particular, acquiring a better understanding of how to think about concepts like evidence and justification within a noncognitivist framework.

1.1   the target: noncognitivism

We may start by getting straight on a little bit of terminology. Dorr uses the term 'noncognitivism' specifically for the class of views shared paradigmatically by Simon Blackburn ([1984], [1993], [1998]) and Allan Gibbard ([1990], [2003]), which might more properly be called 'noncognitivist expressivism'. As a historical matter, the term 'noncognitivism' began life as a name for metaethical theories according to which moral sentences lack truth values, and was used to describe theories like those of Ayer [1936], Carnap [1935], Stevenson [1937], and Hare [1952]. But by the early 1990's, work by Simon Blackburn in particular made it clear that there could be a view that in many ways resembled these earlier 'noncognitivist' theories, but held that moral sentences *do* have truth-values.

Rather than abandoning the label 'noncognitivism', Michael Smith ([1994a], [1994b]) and others showed how it could be reclaimed as a name for views like Blackburn's, if it was reinterpreted to mean that moral judgments are not beliefs, but rather are desire-like attitudes. The meaning of 'noncognitivism' thus became tied to a view shared by Blackburn and Gibbard, but which was not part of the picture of Ayer, Carnap, Stevenson, or Hare, whose views were formulated primarily about language, rather than about thought. It turns out that we can raise variants of the wishful thinking problem for many of these other views, but I'll follow Dorr in this article in restricting attention to contemporary versions of noncognitivist expressivism (henceforth, following Dorr, 'noncognitivism'), emphasizing that it is important to appreciate the differences from these earlier views and to consider separate cases separately.

In any case, we may understand contemporary noncognitivist expressivism, for purposes of the wishful thinking problem, as primarily a thesis about moral thoughts: that for any moral sentence, 'P', to think that P is to be in a desire-like state of mind. This state of mind may be an 'attitude' (Blackburn [1998]) or a state of 'norm-acceptance' (Gibbard [1990]) or a 'planning state' (Gibbard [2003]) or an 'ought-belief' (Horgan and Timmons [2006]) or a 'preference' (Dreier [2006]) or a state of 'being for' (Schroeder [2008]); the important thing is that this state *is* not and moreover *does not involve* any ordinary belief about an ordinary matter of fact. This characterization is important; as Michael Ridge [2007] has emphasized, any 'hybrid' expressivist theory which maintains that where 'P' is a moral sentence, to think that P is to be in a complex state of mind which involves both a desire-like attitude and an ordinary belief about an ordinary matter of fact has the resources to escape the wishful thinking problem. The problem is thus directed solely at 'pure' noncognitivist theories.

I.2   the wishful thinking problem

The idea behind the wishful thinking problem is that intuitively, we ordinarily think that forming beliefs about what the world is like only on the basis of your desires about how you would like things to be is a kind of irrationality – it is 'wishful thinking'. But if expressivism is right, then it should sometimes be rational to form beliefs about what the world is like, on the basis, essentially, of desires. You do this whenever you accept the conclusion of a *moral-descriptive modus ponens* argument on the basis of accepting its premises.

For example, consider the following moral-descriptive *modus ponens* argument, borrowed from Dorr's original article:

P1	If lying is wrong, the souls of liars will be punished in the afterlife.
P2	Lying is wrong.
C	The souls of liars will be punished in the afterlife.

It is intuitively possible for someone (let's follow Dorr in calling him 'Edgar') to rationally come to accept the conclusion of this argument, for the very first time, on the basis of reasoning from these premises. For example, Edgar might start by accepting P1 and at the time have no other evidence for C, and then later come to accept P2 – at which point he may rationally go on to infer C on the basis of this argument.

All of this is very intuitive. But it presents expressivists with a dilemma. Either all of this is right, and Edgar really *can* rationally come to accept C on the basis of P2, having started out only by accepting P1, or it is not right, and Edgar cannot rationally come to accept C on this basis. If it is *not* right, then that is its own problem, because on the face of it, this is a completely rational inference. So on the first fork of the dilemma, the expressivist fails to explain the rationality of what is intuitively a perfectly rational inference. But if it *is* right, then by expressivism's lights, it is rational for Edgar to form an ordinary descriptive belief about the world – for after all, to accept C is to have an ordinary descriptive belief about what will happen to the souls of liars in the afterlife – on the basis of a desire-like attitude – for after all, according to expressivists, to accept P2 is simply to have a desire-like attitude. But in that case, it looks like wishful thinking. So on the second horn of the dilemma, the expressivist is committed to allowing that wishful thinking is sometimes rational.[1]

As noted earlier, the problem as stated is a problem specifically for noncognitivist expressivism, rather than for earlier related views, such as those of Ayer, Carnap, Stevenson, or Hare. It is contemporary noncognitivist expressivists who think that coming to accept P2 involves coming to want something – to have a certain desire-like attitude. And the charge, on the second fork of the dilemma, is that it is *wishful thinking* to come to accept a conclusion about the world on the basis of a change in what you want. So the charge is one that applies specifically to these contemporary views.

But there *is* a related problem for these other related views. Suppose, for example, that coming to accept P2 is a matter of issuing a special sort of command or prescription, or of trying to create a special sort of influence, as other kinds of views hold. Now, someone who changes her mind about what the world is like only in order better to fit with the commands she is issuing, or only in order better to fit with the influence that she is trying to create, is not engaged in something that we would ordinarily call 'wishful thinking', precisely. But it does not appear to be any more rational of her to do so. Moreover, *any* view on

---

[1] Note that neither Dorr nor either of his commentators, Enoch [2003] or Lenman [2003], characterizes the problem as a dilemma; all assume that the second fork is obviously to be avoided.

which coming to accept P2 is different from coming to have any new belief or other cognitive state would appear to have this same general property – how could that make it rational to draw a conclusion about how things are? So it looks like the wishful thinking problem generalizes to a problem for a family of related views. Nevertheless, I will continue to set these other views aside, just to fix the issues.

1.3   the shape of the problem

It is important to appreciate the difference between the wishful thinking problem, which invites expressivists to explain the rationality of inferring C on the basis of P1 and P2, and the problem of explaining the *inference-licensing property* of valid arguments.[2]   An argument has the inference-licensing property just in case someone who accepts its premises is rationally *committed* to going on to accepting its conclusion. Famously, it doesn't follow from this that it is actually rational for him to go on and accept the conclusion; it could be that the only rational course would be for him to stop accepting one of the premises. For example, if Edgar has much better evidence against C than he has for P1 or P2, then the rational thing is not to accept C, but to give up on P1 or P2. Or alternatively, if the only reason Edgar accepts P1 is that he is confident that P2 is false (compare: 'if the moon is made of green cheese, then I'm a billy goat'), the rational response to coming to accept P2 is to stop accepting P1 – not to accept C, even in the absence of other evidence against C.

So the inference-licensing property applies to *every* case of a *modus ponens* argument – *whenever* you accept the premises, you are rationally committed to accepting the conclusion. But only in some cases is it rational for you to discharge this commitment by going on to accept the conclusion. In other cases, the only rational way of dealing with the commitment is to give up on one of the premises.

It has been a part of the traditional Frege-Geach Problem to explain the inference-licensing property, which applies to each and every case, because one of the desiderata of the Frege-Geach Problem is to explain the *validity* of moral arguments, and noncognitivists have proposed to turn the usual order of explanation on its head, and to explain validity in terms of the inference-licensing property, rather than following the usual strategy of expecting the inference-licensing property to be explained by validity. But the wishful thinking problem concerns only the rationality of actually going on to accept the conclusion, which applies only in some cases. So the wishful thinking problem is not a problem about logic or about validity; it is a problem in epistemology – about justification.

---

[2] See, for example, Schroeder [2009], and [forthcoming], especially chapter 6.

Another way of seeing that the wishful thinking problem is distinct from the Frege-Geach Problem – a point which Dorr himself highlights – is that the Frege-Geach Problem arises in full force for expressivist theories in any domain – including theories about probability judgments, epistemic modals, or indicative conditionals. For example, an expressivist about probability judgments might hold that to think that the probability of P is 60% is to have a credence of 60% in P, an expressivist about epistemic modals might hold that to think that Jack might be in Seattle is to have a positive credence that Jack is in Seattle, and an expressivist about indicative conditionals might hold that to have a confidence of n that if you ask, she'll say 'yes', is to have a conditional credence of n in the proposition that she'll say 'yes', conditional on the proposition that you ask.

All of these theories face the traditional Frege-Geach Problem, and need to explain how the sentences of which they seek to provide a special account can combine in complex sentences with the right semantic properties – including validating the right arguments. The Frege-Geach Problem is a *general* problem for expressivist theories. But none of these theories face the wishful thinking problem or any analogue of it, for there is nothing problematic about the idea that a subject could come to be justified in forming an ordinary descriptive belief about a matter of fact, on the basis of having a credence of 60% in P, on the basis of having a positive credence that Jack is in Seattle, or on the basis of having full credence in the proposition that she'll say 'yes', conditional on the proposition that you ask. Forming beliefs on the basis of other cognitive attitudes – such as levels of credence or conditional credence – is not intuitively problematic in the way that wishful thinking is, so there is no second fork to the dilemma.

In short, in light of these considerations it should be clear that the wishful thinking problem is not just a part of the Frege-Geach Problem, wrapped up in new trappings; it is a distinct problem that arises for noncognitivist expressivism within epistemology, and is neither a problem about logic nor a general problem for expressivism.

1.4   mapping out strategies for a solution

Because, as we noted in the last section, the wishful thinking problem focuses on a problem about justification that does not arise in the case of every valid argument or every case in which a subject entertains some valid argument whose premises she accepts, I'm going to introduce a distinction between the cases in which it *is* intuitively rational for Edgar to go on to accept C – which I'll unimaginatively call the *target included* cases – and the cases in which it is *not* intuitively rational for Edgar to go on to accept C –

5

which I'll call the *target excluded* cases. A 'case' is just a situation in which a subject entertains some valid argument whose premises she accepts.

Using this terminology, the first horn of the dilemma is that not all cases are target excluded cases. At least some cases are target included cases. We can then think of the dilemma as arising separately for each of the target included cases. Either – on the first horn – our expressivist theory denies that it is really rational for Edgar to accept the conclusion in that case, or – on the second horn – our expressivist theory claims that it *is* rational for Edgar to accept the conclusion in that case, in which case Dorr argues, first, that the conclusion is adopted only because of a change in desire-like attitudes, and second, that it consequently counts as a case of wishful thinking.

It is helpful to think of the dilemma as arising separately for each case, because one way of responding to the problem is to make further assumptions about Edgar's case, and to try to use those assumptions in order to explain why Edgar's rationality in coming to accept C is not just a case of wishful thinking, because *in that case*, Edgar has independent, ordinary descriptive, evidence for C, so that it is supported on the basis of Edgar's other, ordinary descriptive beliefs, and hence not only supported on the basis of P2, his desire-like attitude. In evaluating each of the proposals that this might be the case, the important thing for us to be keeping track of, is not so much whether any such explanation *works*, as whether some such explanation works for *every* target included case. All that it takes for there to be a problem is that there are *some* target included cases for which no such explanation is possible.

On the other hand, as David Enoch has emphasized, a solution to the problem along these lines would not need to offer a one-size-fits-all solution which needs to apply to each and every case – it could be that different solutions go for different sorts of case, but each target included case is adequately covered by *some* such solution. So what we should be looking for, in evaluating these solutions is whether they *jointly cover* the target included cases. The closer they come to jointly covering the target included cases, the less unintuitive residue will remain. But if any target included cases are left over, then the wishful thinking problem will not have been completely discharged.

Now, in principle several different responses to the wishful thinking problem are possible. The first is to embrace the first fork, biting the bullet and allowing that even though a given case is *intuitively* one in which it is rational for Edgar to accept C, in fact this is really not so. A second response is to embrace the second fork, biting the bullet and agreeing that wishful thinking really is sometimes rational. A third response – which I'll say more about later on – is to agree that it is rational to accept C on the basis of no further evidence than P2 (together with P1), but to deny that this is really *wishful thinking*. But the main sort of response to the problem, offered in both published responses to date, is to try to find a way between the

forks of the dilemma, by arguing that in every target included case, Edgar is in possession of ordinary descriptive evidence for C, which can justify his concluding C without it being a case of wishful thinking.

This last strategy, of course, requires an account of where Edgar's descriptive evidence for C comes from. We can distinguish two possibilities for how this might happen. Since the only thing that changes when Edgar comes to accept P2 is that he comes to have a certain desire-like attitude, the first possibility is that the *fact that Edgar has this attitude* is itself ordinary descriptive evidence that Edgar comes by for C, by coming to accept P2. David Enoch [2003] has tried to exploit this possibility, by arguing that anyone who is *justified* in accepting P1 would also be justified in inferring C from the fact that he has the attitude expressed by P2.

A second possibility is that since Edgar is *justified* in coming to accept P2 (otherwise this wouldn't be a target included case, because it's never justified to draw a conclusion on the basis only of unjustifiedly accepting a premise), he must have some evidence for it. So perhaps it is Edgar's evidence *for P2* which is also evidence for C, and hence which guarantees that Edgar's acceptance of C is, because it is supported by ordinary descriptive evidence, not merely wishful thinking. This second possibility is exploited by James Lenman [2003] in his response to Dorr. In part 2 we'll look at how far these two possibilities can take us, starting with Enoch and the first, and moving on to Lenman and the second.

2.1   enoch on accepting P2

The main idea of Enoch's proposed solution to the wishful thinking problem is that in every target included case, when Edgar comes to accept P2, he comes to have available an independent, purely descriptive, argument for C, which can justify him in accepting C without any wishful thinking. The new premise which Enoch holds becomes available to Edgar when he comes to accept P2 is P2*:

P2*   I disapprove of lying.[3]

And so to get a descriptive argument for C, Edgar must also have available the additional premise P1*:

P1*   If I disapprove of lying, the souls of liars will be punished in the afterlife.

So for Enoch's strategy to work, every target included case must be one in which Edgar is justified in accepting P1*.

---

[3] Here I'm following the practice I've used elsewhere of using 'disapprove' as a generic term for that desire-like attitude toward lying, whatever it is, which is what it is to think that lying is wrong, according to the noncognitivist theory under discussion.

His strategy for establishing this is piecemeal; noting that it can be rational for Edgar to conclude C on the basis of P1 and P2 only if he is justified in accepting P1, Enoch proposes to consider the different ways in which Edgar could be justified in accepting P1, and for each, to argue that if *that* is how Edgar is justified in accepting P1, then he would also be justified in accepting P1*.[4] In this way, he proposes to cover all of the target included cases, dividing and conquering. One of the illustrative cases that Enoch considers is the case in which Edgar's justification for P1 comes from inductive evidence for its universal generalization, ∀P1:

> ∀P1 For any action A, if it is wrong to do A, then the souls of those who do A will be punished in the afterlife.

So let's walk through this case and evaluate whether someone with inductive evidence for ∀P1 would also have to have inductive evidence for ∀P1*, as Enoch claims.

> ∀P1* For any action A, if I disapprove of doing A, then the souls of those who do A will be punished in the afterlife.

The reason why Enoch thinks that anyone with inductive evidence for ∀P1 would also have inductive evidence for ∀P1* is straightforward. It is that getting inductive evidence for ∀P1 involves having come across a significant series of cases of actions A, for which he thinks, 'doing A is wrong and the souls of those who do A will be punished in the afterlife', without having come across any actions, B, for which he thinks, 'doing B is wrong and the souls of those who do B will *not* be punished in the afterlife'. But Enoch reasons, every case in which Edgar thinks, 'doing A is wrong and the souls of those who do A will be punished in the afterlife' is one in which he is in a position to recognize that he disapproves of doing A, and hence in a position to think, 'I disapprove of doing A and the souls of those who do A will be punished in the afterlife', thereby collecting inductive evidence for ∀P1*.

This is clearly a very clever idea. But we should be suspicious of it. For one thing, the very same sort of reasoning would seem to predict that anyone is justified in accepting the following thesis:

> **hubris** For any action A, doing A is wrong just in case I disapprove of doing A.

---

[4] This gloss isn't completely accurate; things are slightly more complicated in the case in which Edgar is justified in accepting P1 only on the basis of testimony. In that case, the non-moral conditional which Enoch's account requires Edgar to be justified in accepting has the form, 'If Jack thinks that it is wrong to lie, then the souls of liers will be punished in the afterlife.' I believe that Enoch's treatment of this case is also flawed, but it isn't necessary to go into it for our purposes, here.

But surely expressivists should not accept a 'friendly suggestion' which leads to this hubristic prediction.[5]

In fact it turns out not to be hard to see that Enoch's reasoning is problematic in at least a couple of places. First, whenever Edgar thinks that doing A is wrong without thinking that he disapproves of doing A, he will collect inductive evidence for $\forall$PI without collecting inductive evidence for $\forall$PI*. Second, it is possible – even rationally possible – for Edgar to be wrong about what he disapproves of. Suppose, for example, that Edgar does not disapprove of viewing pornography, but that consultation with his trusted psychotherapist has led him to believe that he does. And suppose, moreover, that Edgar further thinks that the souls of pornography-viewers are not, as it turns out, punished in the afterlife. In that case, Edgar will accept 'I disapprove of viewing pornography and the souls of pornography-viewers will not be punished in the afterlife' and hence be in possession of conclusive counterevidence for $\forall$PI*, without being in possession of any counterevidence for $\forall$PI. So he could be inductively justified in accepting $\forall$PI without being inductively justified in accepting $\forall$PI*.

Enoch's reasoning also fails in a third way. Even if Edgar justifiedly thinks that he disapproves of something in precisely all and only the cases in which he does justifiedly disapprove of it, things can still go wrong. And that is because the appropriateness of inductive inferences depends on the suitability of the predicates that are being applied – on their *projectability*. Observation of a series of eagles to determine whether they fly will lead to a successful generalization to the effect that eagles fly. But observation of the same series of birds to see whether birds fly will not lead to a successful generalization. You might observe many birds which do fly, and generalize that all do, but the class of birds is heterogeneous with respect to locomotion in a way that the class of eagles is not. Enoch's reasoning requires the inductive evidence to work equally well when Edgar generalizes on what he thinks, as when he generalizes on what is wrong. But there does not seem to be any *a priori* reason to think that this is so. In fact, Edgar may explicitly think that it is not.

So in conclusion, it doesn't look like Enoch's account could apply to *all* cases in which Edgar is inductively justified in accepting PI. Of course, it might, for all that, apply to *some* such cases. But it doesn't look like it will succeed at covering the full range of target included cases, and the foregoing discussion makes it look like a poor candidate to cover even some of the most central and common sorts of cases.

---

[5] See Schroeder [unpublished] for further discussion.

2.2   the structure of lenman's solution

On the face of it, it shouldn't be too surprising that Enoch's strategy ran into trouble. For it doesn't even make use of the full set of resources that ought to be available for explaining how Edgar is justified in accepting C. As Enoch points out, it is rational for Edgar to accept C on the basis of P1 and P2 only if he is *justified* in accepting P1 and P2. But Enoch's explanations only appeal to the assumption that Enoch is justified in accepting *P1* — no work is done by the assumption that Edgar is justified in accepting P2 — only by the assumption that Edgar *does* in fact accept P2. This both overgeneralizes the explanation of Edgar's justification for accepting C to cases in which intuitively Edgar should *not* be justified in accepting C (because he is not justified in accepting P2), and leaves a whole set of possible resources for solving the problem unutilized.

So a different strategy hopes to explain why Edgar always has ordinary descriptive evidence for C in target included cases, by trying to show that the evidence Edgar has for P1 and P2 must itself be descriptive evidence for C. I think this is the right way to understand the strategy taken by James Lenman [2003], in his reply to Dorr, and when described in this way, it is easy to see why this is a more promising strategy than Enoch's, for the reasons just articulated.

Lenman actually adopts a very strong version of this strategy; he holds that whenever Edgar is justified in accepting both P1 and P2, it is on the basis of beliefs which, *independently of P1 and P2,* can be used to directly argue for C. The clearest example that he gives for how this might work is the following argument:

R1   Derek never contravenes the Decalogue.
R2   All and only contraventions of the Decalogue are wrong.
R3   Therefore: Derek never does anything wrong.
R4   Therefore: if looking at a woman with lustful intent is wrong, then Derek never looks at a woman with lustful intent.
R5   Looking at a woman with lustful intent contravenes the Decalogue.
R6   Therefore: Looking at a woman with lustful intent is wrong.
R7   Therefore: Derek never looks at a woman with lustful intent.[6]

In the example, R4 and R6 constitute a moral-descriptive *modus ponens* argument for R7 (just like Edgar's argument for C). But R4 is justified on the basis of R1 and R2, and R6 is justified on the basis of R2 and R5, and R1 and R5 (which are part of the justification of R4 and R6) constitute an independent, direct, descriptive argument for R7. So given the way that R4 and R6 are justified, coming to accept R7 on the

---

[6] Lenman [2003, 272].

basis of R4 and R6 can't lead Edgar any more astray than his beliefs have already led him – because it is independently supported by R1 and R5.

2.3   how far does this solution go?

Lenman's case shows that there are at least *some* examples of target included cases which needn't involve wishful thinking in any objectionable way. So it solves the problem for at least some cases. But recall that there is still a problem, unless Lenman's solution covers *all* target included cases. This is exactly what Lenman claims his solution can do. He alleges that *all* target included cases are like this case, except perhaps simply a little bit more complicated. What Lenman says is that if Edgar does *not* have background beliefs which guarantee that he is in possession of an independent descriptive argument for C, then

> in *that* case the noncognitivist may readily concede that Edgar, so characterized, is irrational. Such a concession is altogether harmless, as it is *independently* highly plausible - *whether we are noncognitivists or not*. If Dorr insists on considering a case where this disconnection is total, we get irrationality by *anybody's* standards.[7]

What Lenman is saying here is that every case which doesn't meet the condition laid down in his account should be classified as being a target *excluded* case – one in which it is intuitively irrational for Edgar to come to accept the conclusion of the argument, anyway.

This is intuitively quite a surprising claim. There is no reason to suspect that in *non-moral* arguments, someone is rational in accepting their conclusion only if they are in possession of some further, different, argument which would independently justify its conclusion. In fact, that *can't* be the case, because it would lead to a vicious regress. *Some* arguments have to support their conclusions without the help of further arguments, or no arguments would support their conclusions at all. So if Lenman's assumption is true of moral-descriptive *modus ponens* arguments, that would be quite a surprising and restrictive conclusion.

It is important, in order to see what is going on, to distinguish Lenman's thesis that the evidence for P1 and P2 must provide an *independent* justification for C from the weaker and more plausible thesis that the evidence for P1 and P2 must provide a justification for C. This latter thesis is compelling because since C follows from P1 and P2, any evidence sufficient to justify both of them would also be sufficient, *derivatively*, to justify C *by way* of justifying P1 and P2. What Lenman's solution requires, is the stronger thesis that there must be a *direct* argument from the evidence for P1 and P2 to C, as in his Decalogue case. This is what we've seen no reason to think is satisfied in the full range of target included cases.

---

[7] Lenman [2003, 269], italics in original.

2.4   what could lenman be thinking?

What we've seen so far is that Lenman's claim that the assumptions required for his solution are independently plausible, 'whether we are noncognitivists or not', is clearly false. His argument requires the assumption that every time someone is justified in accepting a moral thesis, it is justified partly on the basis of non-moral assumptions which themselves suffice to justify any further non-moral conclusions which can be drawn from that moral thesis. Lenman claims that this is 'independently highly plausible – whether we are noncognitivists or not', but so far we've seen no reason why ordinary cognitivists must accept this assumption – indeed, it is easy to think of cognitivist views which are committed to denying it. For example, any intuitionist theory is committed to the view that at least some moral theses are directly justified – and hence not justified on the basis of any non-moral assumptions.

It is worth trying to sort out, however, what Lenman must have been thinking, and his Decalogue example is instructive, here. In the Decalogue example, it is not the non-moral premises alone which justify the premises of the moral-descriptive *modus ponens* argument, but rather the non-moral premises *in combination with* the moral premise, R2, which says that all and only contraventions of the Decalogue are wrong. So there is at least one picture of the structure of moral epistemology which has the feature that if we accept it, then we can adopt Lenman's solution. According to this picture of moral epistemology, anyone who is justified in accepting any moral thesis whatsoever is so justified on the basis of her acceptance of a completely general moral theory of the form 'all and only actions which are **** are wrong'.

If this is a correct thesis about moral epistemology, then in any case in which a subject is justified in accepting the premises of a moral-descriptive *modus ponens* argument, it will have to be because she has derived them from this general moral theory, along with auxiliary non-moral assumptions – as in the Decalogue example. And hence these non-moral assumptions will be available to justify the conclusion of the moral-descriptive *modus ponens* argument. Perhaps Lenman accepts this overall picture of moral epistemology. Perhaps he even thinks that everyone should accept it. That may be what guides his thinking, here. But this is an extraordinarily controversial thesis about moral epistemology – it is very implausible that most ordinary people even accept any perfectly general generalization of the form of R2, and even less plausible that their ordinary moral views are justified on that basis.

If I had to speculate, I would guess that Lenman may be being misled by considerations about supervenience. It *is* very plausible that the supervenience of the moral on the non-moral is an *a priori* constraint on competence with moral terms, and it is a familiar fact that strong supervenience entails the truth of some generalization of the form of R2 – provided that we are sufficiently liberal with what can be

substituted for 'contraventions of the Decalogue'. But it is simply a mistake to think that this entails what Lenman needs; it is perfectly coherent to believe that the moral supervenes on the non-moral and even to think that things are wrong *because* of their non-moral features, without knowing just which non-moral features make something wrong. Indeed, the whole point of introducing talk of supervenience into philosophy in the first place, was that we often find ourselves in this kind of situation: we know that there is a supervenience basis, but we don't know exactly what it is. So considerations about supervenience lend no support to the very strong assumptions about moral epistemology that Lenman needs in order for his solution to the wishful thinking problem to work.

What we've been observing about Lenman's solution, is not that it is *wrong*, but that it relies on an extremely strong assumption about the structure of moral epistemology which – contra Lenman – is not at all obviously plausible independent of a commitment to noncognitivism, or even independent of the wishful thinking problem, for that matter. What his solution shows, is that noncognitivists can solve the wishful thinking problem by taking on this kind of very strong picture about the structure of moral epistemology. And if you find it independently plausible, as Lenman does, that everyone accepts a comprehensive moral theory and bases their judgments about cases on that theory, then you might be happy with this solution. But if this is the only way around the wishful thinking problem, then it is fair to say that the wishful thinking problem places very sharp constraints on moral epistemology – constraints that we might intelligibly wonder whether noncognitivists can do without.

3.1 what about the second fork?

So it seems that neither Enoch's nor Lenman's suggestions do quite what they are presented as doing. But it does seem to me that it is worth thinking about the merits of the second fork of the dilemma. A first observation is that 'wishful thinking' is something of a persuasive definition. True cases of wishful thinking are cases of wanting it to be the case that *p*, and coming to believe that *p*. That is clearly a bad way of going, and deserves a special name. But it is not obvious that the connection between P2 and C looks like this, unless it turns out that 'lying is wrong' expresses the state of wanting the souls of liars to be punished in the afterlife. So it could be that some cases of getting to descriptive conclusions from, among other things, a desire-like attitude, are not as bad as the paradigm cases of wishful thinking, and hence the name of the problem is itself efficacious in dissuading us from the second fork. If so, it would seem, we should be cautious.

A second observation is that even though Dorr describes cases in which Edgar starts by accepting P1 and comes to accept P2, his acceptance of P1 is not, itself, irrelevant to the justification for C. But though it is clear that accepting P2 is just having a desire-like attitude, it is not clear what sort of state is involved with accepting P1, until an adequate expressivist solution to the embedding problem and account of logical inconsistency and logical entailment is on the table. Whatever such a state turns out to be like, it will have to have the property that it can be involved in joint inconsistency with both beliefs and desire-like attitudes at the same time. It must turn out, for example, that it is inconsistent to be in the state expressed by P1, the desire-like attitude expressed by P2, and the belief expressed by the negation of C. It is very puzzling how there could be any state that could make this so, and that is an important part of why an especially hard part of the Frege-Geach Problem in the first place is to give an account that deals adequately with mixed moral-descriptive conditionals.[8]

So perhaps noncognitivists should embrace the second fork of the dilemma, and argue that this is nevertheless sufficiently different from the ordinary cases of wishful thinking that it is not at all obvious that whatever is so bad about such ordinary cases carries over. Making good on this strategy would require making good on an explanation of precisely what is so objectionable about paradigm cases of wishful thinking, and an explanation of why those objectionable features don't arise after all in superficially resembling cases like Edgar's. To make good on such explanations, we would need a much more complete understanding of both wishful thinking and of noncognitivist theories of moral thought and inference.

3.2   tools for progress

In this paper, we've been considering just one significant problem for noncognitivism in epistemology – the wishful thinking problem. I have focused on it because it is relatively new and interesting, and because the responses to date on behalf of noncognitivism have been intriguing but less than convincing and it is not clear what satisfactory view could come out of it. But there are a variety of other significant problems for noncognitivists in epistemology, and it is worth thinking about the wishful thinking problem in this broader context. For example, all along in the discussion in this chapter, I have followed both Enoch and Lenman in assuming that within a noncognitivist framework, it will make sense to talk about 'evidence' for P1 and P2, and assumptions that 'justify' Edgar in accepting one or the other.

But it is very obscure what account can be given of evidence for P2, on an expressivist view, or of what it is to be justified in one's moral views, as opposed to being justified in an ordinary descriptive belief.

---

[8] See, for example, Hale [1993], [Kölbel 2002], and Schroeder [2008], especially chapters 7-10.

This is not to say that these concepts are beyond the reach of noncognitivists, but they are certainly ones a better understanding of the noncognitivist approach to which it is natural to expect will lead us to further resources for getting our heads around the wishful thinking problem.

Consider the following comparison. Mathematicians worked for a long time with only an intuitive notion of continuity, because they didn't have the resources available to define it properly, which required the epsilon-delta definition of a limit. Discussing problems which turn on the notions of evidence and justification is like working with the intuitive notion of continuity — something like 'can you trace it without taking your finger off of the graph?'. Absent the right tools for understanding how justification and evidence really work in a full-fledged noncognitivist framework, it may in some cases be hard to tell whether we have a problem about justification, or a place in which our hand-wavy intuitive notion is breaking down. So trying to solve the wishful thinking problem directly, without developing a more mature underlying theory of justification and evidence, may simply be premature.

The problem is that epistemological notions promise to be some of the hardest to get straight on what the right noncognitivist account of them should be — and not least because these epistemological notions are controversial in their own right. This is because if someone knows something only if they truly believe it, then providing an expressivist account of knowledge will be at least as hard as providing an expressivist account of belief and providing an expressivist account of truth, put together — but noncognitivist accounts of belief and of truth are themselves complicated topics, each in their own right. And that still leaves out whatever further condition is required to turn true belief into knowledge, which epistemologists have had enough trouble understanding, even on the assumption that noncognitivism is false. So all told, there are excellent reasons to expect the epistemological issues facing expressivism to be very difficult, as well as being particularly difficult to resolve without first more general issues facing noncognitivism about logic, truth, and belief.

references

Ayer, A.J. [1936]. *Language, Truth, and Logic*. New York: Dover.

Blackburn, Simon [1984]. *Spreading the Word*. Oxford: Oxford University Press.

_____ [1993]. *Essays in Quasi-Realism*. Oxford: Oxford University Press.

_____ [1998]. *Ruling Passions*. Oxford: Oxford University Press.

Budolfson, Mark [unpublished]. 'Non-Cognitivism and Rational Inference.' Unpublished paper.

Carnap, Rudolf [1935]. *Philosophy and Logical Syntax*. Bristol, UK: Thoemmes Press.

Dorr, Cian [2002]. 'Non-cognitivism and Wishful Thinking.' *Noûs* 36(1): 97-103.

Dreier, James [2006]. 'Negation for Expressivists: A Collection of Problems with a Suggestion for their Solution.' *Oxford Studies in Metaethics*, volume 1. Oxford: Oxford University Press, 217-233.

Enoch, David [2003]. 'How Noncognitivists Can Avoid Wishful Thinking.' *Southern Journal of Philosophy* 41: 527-545.

Gibbard, Allan [1990]. *Wise Choices, Apt Feelings.* Cambridge, MA: Harvard University Press.

———— [2003]. *Thinking How to Live.* Cambridge, MA: Harvard University Press.

Hale, Bob [1993]. 'Can There Be a Logic of Attitudes?' In Haldane and Wright, eds., *Reality, Representation, and Projection.* New York: Oxford University Press.

Hare, R.M. [1952]. *The Language of Morals.* Oxford: Oxford University Press.

Horgan, Terry, and Mark Timmons [2006]. 'Cognitivist Expressivism.' In Horgan and Timmons, eds., *Metaethics After Moore.* Oxford: Oxford University Press.

Kölbel, Max [2002]. *Truth Without Objectivity.* New York: Routledge.

Lenman, James [2003]. 'Noncognitivism and Wishfulness.' *Ethical Theory and Moral Practice* 6: 265-274.

Ridge, Michael [2007]. 'Ecumenical Expressivism: The Best of Both Worlds.' *Oxford Studies in Metaethics* 2: 51-76.

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

———— [2009]. 'Hybrid Expressivism: Virtues and Vices.' *Ethics* 119(2): 257-309.

———— [forthcoming]. *Noncognitivism in Ethics.* Forthcoming from Routledge.

———— [unpublished]. 'Does Expressivism Have Subjectivist Consequences?' Unpublished paper.

Smith, Michael [1994a]. 'Why Expressivists About Value Should Love Minimalism About Truth.' *Analysis* 54(1): 1-12.

———— [1994b]. 'Minimalism, Truth-Aptitude, and Belief.' *Analysis* 54(1): 21-26.

Stevenson, C.L. [1937]. 'The Emotive Meaning of Ethical Terms.' Reprinted in Stevenson [1963], 10-31.

———— [1944]. *Ethics and Language.* Oxford: Oxford University Press.