Metaethical pluralism: 
How both moral naturalism and moral skepticism may be permissible positions

Richard Joyce

Penultimate draft of the chapter appearing in S. Nuccetelli and G. Seay (eds.),

Introduction: Some distinctions

This paper concerns the relation between two metaethical theses: moral naturalism and moral skepticism. It is important that we distinguish both from a couple of methodological principles with which they might be confused. Let us give the label “Cartesian skepticism” to the method of subjecting to doubt everything for which it is possible to do so—usually by introducing alternative hypotheses that are consistent with all available evidence (e.g., brains in vats). Let us give the label “global naturalism” to the principle that requires of any item which we admit into our ontology that it “fits” (in some manner or cluster of manners to be specified) with our naturalistic scientific worldview. One might be both a Cartesian skeptic and a global naturalist, if the latter principle is something that has survived the former test procedure. Alternatively, one might have adopted global naturalism for some other reason, while having little patience with the Cartesian method of doubt.

Moral naturalism is the metaethical view that moral entities (e.g., properties like goodness and evil) fit within the scientific image of the world. The moral naturalist will probably be a global naturalist, but need not be: It is consistent with allowing non-natural entities into one’s ontology that one happens to think that moral properties are of the natural variety.

Moral skepticism denies that moral entities fit within our scientific worldview. One way of denying moral naturalism is to be a moral error theorist: to hold that our moral discourse attempts to make reference to moral properties, but these properties do not exist. Another way of denying moral naturalism is to be a noncognitivist: to hold that our moral discourse was never really in the business of referring to moral facts or properties in the first place, and ipso facto such facts or properties are not naturalistic. In this paper, the label “moral skepticism” denotes the disjunction of these two theses. Neither the error theorist nor the noncognitivist must be committed to global naturalism, but usually will be; indeed, this commitment will often be a motivating factor of their metaethical views. The error theorist who is a global naturalist will typically deny the existence of moral properties precisely because (he thinks) these properties fail to fit with a naturalistic worldview. Likewise, those who turn to noncognitivism often do so because they cannot see how moral properties could find a place in a naturalistic order.

(There is a third way of denying moral naturalism which is not a kind of moral skepticism at all. The moral non-naturalist denies that moral entities fit within our naturalistic worldview but does not judge this a mark against their existence or our epistemic access to them; moral

---

1 The usual error theoretic strategy here is to argue that moral properties have features that no naturalistic property can have (e.g., some strong kind of practical authority) and thus the moral naturalist’s characterization of moral properties must be inadequate. (See Mackie 1977; Joyce 2001.) Alternatively, the error theorist might accept the moral naturalist’s characterization of the moral (for the sake of argument) and argue that there are still no such properties. For a study of the latter strategy, see Joyce 2011.
facts are acceptable non-natural facts. I shall not be discussing this metaethical view in this paper.)

“Skepticism” is a pliable term, and the way I am using it here will not jibe with the way all others use it. I have already stipulated how the term will be employed, but it is probably useful to make a couple of further clarifications. First, I have described both the error theorist and the noncognitivist as denying something: that moral entities are to be admitted into our ontology. This stands in contrast to a classical kind of skeptic who neither affirms nor denies but simply doubts. The skeptic I have in mind doubts X in a strong sense of affirming the denial of X (and is thus, in the classical vocabulary, a negative dogmatist). Second, my moral skeptic is not playing the Cartesian game of doubting moral naturalism simply because alternative hypotheses can be conjured. My moral skeptic has probably accepted the naturalistic worldview—indeed, employs it as a premise for skepticism. Thus the skeptic takes himself or herself to have positive grounds for doubting moral naturalism beyond the unimpressive pastime of imagining farfetched (but unfalsifiable) scenarios in which moral properties are absent. This is a naturalistic skepticism, not a Cartesian skepticism.

Moral naturalism and moral skepticism are certainly not contradictories; they do not exhaust the space of metaethical possibilities. (I have already mentioned one alternative: non-naturalism.) But it is usual to think of them as contraries, for surely to embrace one position is to reject the other. In this paper I want to investigate this latter assumption closely. I will explore the possibility of a certain kind of metaethical pluralism, whereby seemingly contrary metaethical positions enjoy equal claims to legitimacy—and not just in the sense that we are not (yet) in a position to know which position is correct, but in the more interesting sense that there is no fact of the matter about which is correct.

**Lewis’ pluralism**

I take as my point of departure a series of intriguing and characteristically astute comments made by David Lewis at the close of his 1989 paper “Dispositional theories of value (part II).” In the course of this paper Lewis has developed and advocated a kind of moral naturalism: a dispositional theory according to which values are those things which we are disposed, under certain idealized conditions, to desire to desire. Worried that this theory makes values contingent—that we feel uneasy about a theory that allows that benevolence (say) is, but might not have been, a value—Lewis toys with identifying values with what we are necessarily disposed to desire to desire. This amendment would probably soothe the worry, he acknowledges, but it would be at a price: There is likely nothing that we are necessarily disposed to desire to desire; thus to identify moral properties with this dispositional property would leave us with a moral error theory.

It is not Lewis’ dispositional theory per se that interests me here, but what he goes on to say next. The version with the necessity operator is, he thinks, that which “best captures what it would take for something to perfectly deserve the name ‘value’” (Lewis 1989: 136). But

---

2 It is difficult to choose a label that denotes the disjunction of the error theory and noncognitivism. “Moral anti-realism” would serve in some people’s books, but I prefer to reserve this term to cover the error theory, noncognitivism, and a kind of moral naturalism according to which moral facts are in some manner constructed by human attitudes (as opposed to being objective features of the world). To my terminological tastes, some moral naturalists are moral realists and some are moral anti-realists. The disjunction of the error theory and noncognitivism is called “skepticism about moral truth” by Walter Sinnott-Armstrong in his taxonomy of moral skepticisms (Sinnott-Armstrong 2006: 11).
this perfect deserver leaves us with skepticism. However, “there are plenty of imperfect deservers of the name” (ibid.).

Strictly speaking, nothing shall get the name without deserving it perfectly. Strictly speaking, genuine values would have to meet an impossible condition, so it is an error to think there are any. Loosely speaking, the name may go to a claimant that deserves it imperfectly. Loosely speaking, there are values, lots of them. (136-7)

He goes on:

What to make of the situation is mainly a matter of temperament. You can bang the drum about how philosophy has uncovered a terrible secret: there are no values! … Or you can think it better for public safety to keep quiet and hope people will go on as before. Or you can declare that there are no values, but that nevertheless it is legitimate—and not just expedient—for us to carry on with value-talk, since we can make it all go smoothly if we just give the name of value to claimants that don’t quite deserve it. … Or you can think it an empty question whether there are values: say what you please, speak strictly or loosely. When it comes to deserving a name, there’s better and worse but who’s to say how good is good enough? Or you can think it clear that the imperfect deservers of the name are good enough, but only just, and say that although there are values we are still terribly wrong about them. Or you can calmly say that value (like simultaneity) is not quite as some of us sometimes thought. Myself, I prefer the calm and conservative responses. But as far as the analysis of value goes, they’re all much of a muchness. (137)

Perhaps the moral error theorist has a different temperament to that of Lewis. I confess that I have long been drawn to the error theoretic version of moral skepticism, and it would be untruthful to claim that this is entirely the product of having been persuaded by solid philosophical argumentation. Maybe it is a mischievous streak in me; perhaps it is an excitement about striding into uncharted intellectual territory having cast off familiar assumptions. Who knows? I wouldn’t attempt to publicly analyze the antecedents of my skeptic-leaning temperament any more than Lewis sought to openly scrutinize his own conservative preferences. The important points are (i) that temperaments opposed to Lewis’ exist (exhibit A = myself), and (ii) that Lewis thinks our temperaments determine “what to make of the situation.” Thus he makes one thing of the situation—moral naturalism—while I make another—error theoretic skepticism. Is this a reasonable conclusion? And, if so, then where (if anywhere) do we go from here?

Lewis is certainly not saying that all there is to the debate between the moral naturalist and the moral error theorist are their respective temperaments. Both parties must first have some decent arguments on their side. For example, we can assume that no amount of “temperament” favoring the existence of magic will make a naturalistic endorsement of magic reasonable. Likewise, no amount of “temperament” opposed to the existence of big hairy spiders is going to make doubting their existence epistemologically acceptable. From the perspective of global naturalism, some things clearly exist (spiders) and some things clearly do not (magic). Where things get interesting is in the gray area, where there are decent arguments both for and against the existence of the item, and here, Lewis thinks, temperaments might matter.

Let us approach the problem with reference to Ramsey sentences. Suppose we are wondering whether Φ exists. We construct a list of platitudinous desiderata of what we think
Φ is like: “Φ is P₁,” “Φ is P₂,” etc. (Of course, the sentences do not need to be simple atomic predications, but it makes expression easier.) Let us say we come up with twenty such desiderata. We then conjoin these sentences, replace all mention of Φ with the same variable, and bind that variable with an existential quantifier:

\[ \exists x (x \text{ is } P₁ \land x \text{ is } P₂ \land \ldots \land x \text{ is } P₂₀) \]

This allows us to say that Φ exists if and only if the existential claim (the Ramsey sentence) is true. There are two ways in which it can be false: if there is nothing that satisfies all twenty predicates, or if there is more than one thing that satisfies all twenty predicates. (See Lewis 1970.)

But one should not be too strict about this, for there is a requirement to accommodate the fact that we can be mistaken about certain qualities of Φ without it automatically following that Φ does not exist. We gave the name “Earth” to the thing we live upon and at one time reckoned it flat (or at least a good many people reckoned it flat); but the discovery that the thing we live upon is a big ball was not taken to be the discovery that we do not live upon Earth. It was once widely thought that gorillas are aggressive brutes, but the discovery that they’re in fact gentle social creatures was not taken to be the discovery that gorillas do not exist. Examples abound. We make mistakes, big and small, without facing skepticism at every turn.

Applying this to Lewis’ argument: One of the desiderata for “value” pertains to non-contingency. (I shan’t pause for further specificity.) But nothing that satisfies the other desiderata (some of which can be systematized into that which we desire to desire) also satisfies non-contingency, and this is why Lewis claims that “genuine values would have to meet an impossible condition.” But, he thinks, there are imperfect claimants—ones that satisfy nearly all that we want. Just as a spherical Earth was an imperfect claimant of all that we might have wanted to say about the Earth (but close enough to avoid Earth skepticism), and gentle gorillas were imperfect claimants of all that we might have wanted to say about gorillas (but close enough to avoid gorilla skepticism), so too, Lewis thinks, his dispositional theory of value is an imperfect claimant of all that we might have wanted to say about value, but close enough to avoid moral skepticism.

But Lewis realizes that his dispositional theory is not close enough to silence all debate on the matter; it remains in a gray area. And the grayness persists because we have no methodology for making decisions concerning indeterminate cases. Lewis writes: “What it takes to deserve this name, not perfectly but well enough, was never settled” (1989: 136); and: “When it comes to deserving a name, there’s better and worse but who’s to say how good is good enough?” (1989: 137). Who indeed?

Elsewhere I have toyed with the idea that what determines this matter is a complex counterfactual about how we would, if called upon, as a matter of fact decide (Joyce 2006: 201). But this idea was never meant to help us out as a decision procedure, and it is, in any case, problematic. Who, exactly, is this “we,” and in what circumstances are we being called upon to make this decision? Since I harbor no optimism that in making such decisions we collectively or individually follow any particular rules, overt or tacit, I’m inclined to think that our pronouncements on such matters can be influenced by the most arbitrary of stimuli (e.g.,

---

3 I have on occasion made what I take to be the equivalent point via a thought experiment concerning translation. (See Joyce 2001: 3, 26-7; 2006: 71.)
the plot of a popular movie). Thus in circumstances A we’ll decide that we’ve had a false belief about \( \Phi \), correct that error and carry on believing in \( \Phi \) (i.e., we’ll be \( \Phi \) naturalists); while in circumstances B (which are not dramatically different from A) we’ll decide that our belief in the existence of \( \Phi \) has been a huge mistake (i.e., we’ll be \( \Phi \) skeptics). This to some extent problematizes the counterfactual “If we were to decide on the matter, we would decide thus-and-so.”

Subsequently, I suggested that what determines the question is how we use the concepts (Joyce 2006: 201; 2007: 65). Roughly: If concept \( \Phi \)—understood as satisfying twenty desiderata—has been used in ways \( U_1, U_2, \ldots U_n \), but the best imperfect claimant (call it \( \Phi^* \), which satisfies, say, eighteen of those desiderata) cannot be put to the same range of uses, then this is sufficient grounds for denying that \( \Phi^* \) is “close enough” to count as a revision of \( \Phi \).

I continue to think that this appeal to usage promises to make some headway toward solving some of the Lewisian Is-it-close-enough?-type questions. Regardless of whether one’s temperament favors Lewis’ “calm and conservative” preference for dispositional moral naturalism, if it turns out that the dispositional property in question just isn’t going to pull its pragmatic weight in everyday life—if it can’t do the work of a full-blooded all-you-could-want moral property—then this would count decisively against this form of moral naturalism. Likewise, irrespective of whether one’s mischievous temperament favors error theoretic moral skepticism, if it turns out that Lewis’ dispositional property can be pressed into service for all the practical roles that moral concepts play, then the day would go to the moral naturalist and his imperfect claimant.

One interesting and possibly surprising consequence of conceptualizing the problem in this manner is that it makes the debate between the moral naturalist and the moral error theorist at bottom an empirical debate. We ask, first, “What is morality used for?” Then we ask “Could concept such-and-such [the imperfect claimant] be used in this manner?” Both are difficult and complex matters, but they are both empirical questions, and the former, at least, seems reasonably tractable. Yet the remaining problems are more than just the difficulty of our gaining epistemic access to these empirical truths. The problem is that the indeterminacy simply reiterates in a new place. Suppose we have used concept \( \Phi \) for ten purposes—\( U_1, U_2, \ldots U_{10} \) (idealizing horribly here, of course)—and suppose that the best imperfect claimant (call it \( \Phi^* \)) can be used in, say, eight of those ways. We can’t use \( \Phi^* \) for everything that we used to use \( \Phi \), but we can use it for most things. Well, is that close enough? I feel at this point we can only reiterate Lewis’ question: “Who’s to say?” If we can find no good answer, then we have made not only the surprising discovery that the debate between the moral naturalist and the moral skeptic bottoms out in empirical disputes, but a second surprising discovery that there may be no fact of the matter about who is correct.

**Carnap’s pragmatism**

Lewis is evidently content to live with this indeterminacy between moral naturalism and moral skepticism. But we shouldn’t give in just yet; there is an obvious avenue to survey: an appeal to pragmatism. If there is an indeterminacy in what we mean by “value” (say)—in that both the moral skeptic and the moral naturalist make defensible claims about what is necessary or sufficient for something to be a value—then let the matter be decided by which
is most practically expedient; let us choose a precisification on the basis of which best serves our purposes.

Many philosophers, it seems to me, are drawn to this method of selecting sides when indeterminacy threatens, and, moreover, the common assumption regarding the present case is that an appeal to pragmatism will favor the moral naturalist over the skeptic. My aim is to call into question both moves. First, the appeal to pragmatism is fraught with difficulties and does not obviously represent an escape route from indeterminacy. Second, I think the assumption that an appeal to pragmatism (were it to prove workable for this purpose) would favor the moral naturalist can be subjected to pressure. But before arguing these claims I want to bring the noncognitivist onto the stage, and I will do so via a discussion of Rudolf Carnap. Carnap’s views are useful for our purposes since as well as advocating noncognitivism he espouses an interesting pragmatic methodology.

That the Carnapian notion of explication is relevant to our discussion is immediately evident from the characterization Carnap gives it in *Logical Foundation of Probability*: “By the procedure of explication we mean the transformation of an inexact, prescientific concept, the explicandum, into a new exact concept, the explicatum” (Carnap 1950a: 3). The test of a good explication is not whether it is “true” or “correct” (Carnap says such claims make “no good sense” [ibid. 4]); but rather “whether the proposed solution is satisfactory” (ibid.)—where this satisfactoriness is in part a pragmatic notion, centrally depending on whether the proposed explicatum is fruitful and simple.

The pragmatism embodied in his notion of explication is plain elsewhere in Carnap’s philosophy. In “Empiricism, semantics, and ontology” (1950b) he analyzes many traditional philosophical questions as “external questions.” The question “Does the world exist,” for example, is really a practical question of whether we should adopt the “thing language” according to which objects exist at a space and time within the world. In determining the practical pay-off, “[t]he efficiency, fruitfulness, and simplicity of the use of the thing language may be among the decisive factors” (ibid. 208). He says the same about the external questions “Do propositions exist?”, “Do numbers exist?” and so on. According to Carnap, these questions should be interpreted as: “Are our experiences such that the use of the linguistic forms in question will be expedient and fruitful?” (ibid. 213).

Carnap thus provides us two ways that we might try to break out of the indeterminacy surrounding “value” (and other moral terms). First, we might try to explicate the notion, imposing precision on an indeterminate concept in a manner guided by expediency. Second, we might treat the question “Do values exist?” as an external question, deciding whether to employ the “value language” on the basis of its utility as an instrument.

Before examining these possibilities, we should have Carnap’s own metaethical view before us. His years as a leading member of the Vienna Circle had made a moral skeptic of him—but of the noncognitivist rather than the error theoretic variety. By 1935 he was staunchly advocating a prescriptivist version of noncognitivism: “Most philosophers have been deceived [by grammatical form] into thinking that a value statement is really an assertive proposition and must be therefore true or false. … But actually a value statement is nothing else than a command in a misleading grammatical form. … It does not assert anything” (Carnap 1935: 24-5). His noncognitivism had first become evident in 1929, in a lecture delivered to the Bauhaus in Dessau, where (according to fragmentary notes) he claimed: “Valuation (Wertung) is not the cognition of a fact but a personal attitude” (cited in Mormann 2007: 133). Interestingly, in his *Aufbau*, a younger Carnap had espoused a very different
position, where values are a part of the world order, whose presence may be known and asserted (see §152 of Carnap 1967 [1928]). In what follows, however, I shall take the mature Carnap’s noncognitivism as representative of his view. (See Mormann 2007, for discussion of Carnap’s change of mind.)

When Lewis outlines the range of options (in the long passage quoted earlier), the noncognitivist is noticeably absent; all that is on his radar is a continuum of positions between the moral error theory and moral naturalism. It is clear why this is so. Lewis’ examination of “value” as a verb (roughly: desiring to desire) has provided him with license to treat “value” as a noun (roughly: that which we are disposed to desire to desire), which in effect is license to exclude noncognitivism. Thus the question that exercises him at the close of the paper is whether the world provides a property that is close enough to deserve this noun (albeit imperfectly). The noncognitivist, by contrast, will not let Lewis get that far. The noncognitivist endorses an account of “value” as a verb that does not provide license for treating “value” as a noun. Focusing on “valuing” as a linguistic act rather than a kind of mental attitude, Carnap tells us that value utterances are really commands, in which case the question of to what kind of entity these value utterances refer simply does not arise. When one says “X is good,” Carnap thinks, one is saying something along the lines of “Pursue X!” The goodness has evaporated on this analysis. Asking whether the world provides anything close enough to deserve the noun “goodness” is, in Carnap’s opinion, a pseudo-question, because concepts like goodness are pseudo-concepts and therefore not fit for analysis.

If there is no concept of value, then a fortiori there is no such concept to be made precise via explication. That is not to say, however, that explication can play no role in Carnap’s metaethics. The concept of value accused of being an inexplicable pseudo-concept is the one associated with the noun “value.” But Carnap certainly believes in the human activity of valuing, and the related concept (the one associated with the verb “to value”) will not also be accused of “pseudo-ness.” Moreover, Carnap also believes that human valuing will often take the form of issuing statements, and thus there is a noun (“value statement”) whose associated concept is available for analysis and, if necessary, explication. On the only occasion that Carnap does explicitly sketch an explication of something normative, the explicandum is, indeed, “value statement” (Schilpp 1963: 1009 ff.).

Now, it is not my intention to discuss Carnap’s proffered explication in detail, but rather make some general comments about it. First, it is interesting to note that he does not take himself to be deriving the distinction between noncognitive/evaluative utterances and descriptive/factual statements from any examination of natural language practices; rather, he is speaking of “possible kinds of meanings and the relations between these meanings” (ibid. 1003). His aim is to establish that a noncognitive language is possible, and “to use it as a basis for the philosophical discussion of value problems” (ibid.).

This last comment reveals that Carnap’s eye is, as usual, on the pragmatic pay-off, but one should be aware of his restricted view on this matter. The “fruitfulness” of explication is understood in terms of the explicatum being “brought into connection with other concepts on the basis of observed facts; in other words, the more it can be used for the formulation of laws” (Carnap 1950: 6). Thus, it is not any general psychological kind of practical usefulness that is being claimed for noncognitivism; indeed, Carnap opines that whether one is a noncognitivist or a moral naturalist “is relatively unimportant in its influence of practical life”
When he does claim some advantage for his noncognitive explication of a value statement over any cognitive rival, it is in terms of a very slight (almost, one might say, unnoticeable) additional discriminatory power in the former language: The noncognitive language can provide an interpretation of a certain sentence of the cognitive language, but not vice versa (ibid. 1004). However, the argument is unpersuasive, to say the least. Among its weaknesses is the fact that it claims this advantage regarding just one type of comparison; but who knows what advantages the cognitivist language might enjoy over the noncognitive language if we looked further afield?  

Let us conclude, then, that Carnap leaves the matter undecided. And he would, I think, agree with and even welcome this diagnosis. His Principle of Tolerance allows metalanguages to compete freely for our allegiance, to be judged not according to “truth” but according to fruitfulness. Permit the noncognitivist and the cognitivist to advocate their respective languages; Carnap is tolerant of the competition. The language that best suits our theoretical purposes will be the eventual victor. “Let us grant to those who work in any special field of investigation the freedom to use any form of expression which seems useful to them; the work in the field will sooner or later lead to the elimination of those forms which have no useful function” (1950b: 221).

I will sum up before moving on. I set out to explore the possibility of an irresolvable indeterminacy between moral naturalism and moral skepticism. Lewis located one potential node of indeterminacy: between moral naturalism (of a certain kind) and error theoretic moral skepticism. This is an indeterminacy over “value” as a noun. But before we get that far, we face the possibility of another node of indeterminacy: between cognitivism (which includes all forms of moral naturalism) and the other kind of moral skepticism, noncognitivism. This can be thought of as indeterminacy over “value” as a verb. (Quite possibly there is indeterminacy in the rival noncognitivist construals of valuing, too—between, say, Hans Reichenbach’s volitionism and A.J. Ayer’s emotivism.) Later I will further discuss the possible indeterminacy between naturalism and skepticism, but first I will investigate the potential undecidability of the debate between cognitivism and noncognitivism.

**Cognitivism vs. noncognitivism**

Carnap, as we saw, invites open competition with his cognitivist rival. But the mere fact that one is sufficiently undecided between two theories of X (or sufficiently sportsmanlike) to be willing to allow them to “compete” in some manner for victory hardly means that there is a substantive indeterminacy about X. For there to be indeterminacy it must be the case that (A) this competition yields no winner, and (B) no other form of decision procedure is available.

---

4 Carnap does think that clearly demarcating value questions from factual questions is practically important, and that the failure to do so “leads to confusions and misunderstandings in discussion of moral problems in personal life or in political decisions” (in Schilpp 1963: 81)—but that is not the same as thinking that the distinction between noncognitivism and cognitivism makes a practical difference.

5 A couple of other weaknesses with Carnap’s arguments are the following: First, he seems to be concerned only with showing how purely noncognitive sentences (“pure optatives,” as he calls them) are possible; but the moral cognitivist need not deny this, but rather simply denies that moral utterances are instances of purely noncognitive sentences. Second, Carnap has only a very restricted kind of cognitivism in his sights when he performs this comparison.

6 They are metalanguages because the Principle of Tolerance is applied after sentences in the material mode of speech have been translated into sentences in the formal mode.
Could there really be no fact of the matter about metaethical cognitivism versus noncognitivism?

One might think that of course there is a fact of the matter. This debate is, essentially, about what we are doing when we value something, and the act of valuing is a phenomenon available to our investigation. This is why although there are plenty of advocates of an error theory about values, there are no error theorists about valuing. We do not doubt that valuing occurs, and this type of action has certain features which determine the answers to various questions we might ask about the phenomenon. So surely the debate between the cognitivist and the noncognitivist must have an answer, even if we’re not sure yet what it is.

But this line of reasoning is mistaken. Valuing may well be an actual phenomenon displayed before us, available to our scrutiny, but the phenomenon will yield specific answers only if we ask specific questions. There are different ways of framing the cognitivism/noncognitivism question, but the standard way is: “Are moral judgments assertions?” Here valuing is treated as a linguistic activity, and the question concerns what kind of speech act constitutes the activity. Specific enough?

Not necessarily. Questions arise as to what is meant by “assertion.” There are significant differences among the views of assertion advocated by C.S. Peirce (1934), Michael Dummett (1959), J.L. Austin (1962), John Searle (1969), Michael Slote (1979), Robert Brandom (1994), and Timothy Williamson (2000)—to name just a few. Can these differences be settled? Perhaps not. Perhaps different accounts of assertion work better in different theoretical environments, with there being no decisive facts to appeal to in order to settle either which is the best account of assertion overall or which is the best account of assertion to apply to the moral case.

Many accounts of assertion state or imply that the person making an assertion aims at the truth, or aims to state a fact (see Williams 1966; Dummett 1981: 300). So when someone claims “Stealing is wrong,” is he or she stating a fact? It depends what you mean by “fact”—of which, it comes as no surprise, there are competing theories. One might distinguish between “fact” in a metaphysically robust sense, and “fact” in some more minimalist sense. This, in turn, would give us a distinction between “assertion” in a robust sense and in a minimalist sense. Indeed, we find exactly this kind of distinction appearing in metaethical debates over quasi-realism. The quasi-realist eschews the metaphysics of moral realism but thinks that language with all the trappings of realism is still permissible (Blackburn 1984, 1993). Thus, even if there is no metaphysical property of wrongness, the quasi-realist is happy to carry on claiming not only “Stealing is wrong,” but “‘Stealing is wrong’ is true,” “There is a property of wrongness (and stealing has it),” “Stealing is wrong regardless of my attitude,” “I hereby assert that stealing is wrong,” and “I really, really mean it.” A typical quasi-realist (noncognitivist) view is that while we can claim all these things (i.e., truth, fact, property, assertion) for morality in a minimalist sense, we cannot do so in a robust sense. In this way the quasi-realist hopes to vindicate the realist trappings of moral discourse while continuing to distinguish the position from moral realism.

This situation is not simply one of philosophers arguing over the correct analysis of a concept whose extension all will substantially agree to; rather, the extension may very well differ significantly depending on which conception is preferred. It is possible that the question

---

7 Unless, of course, it follows from the endorsement of some grander skepticism, encompassing doubt about cats and dogs, furniture, other people, etc. Even an error theorist about all mental phenomena need not be an error theorist about valuing, since one can treat valuing as a species of linguistic act or behavioral response.
“Are moral judgments assertions?” should receive a positive answer if using a minimalistic conception of *assertion* but a negative answer if using a robust conception. But nor is this situation simply one of an ambiguous concept, for the parties to the dispute will usually insist that their preferred conception is uniquely correct. We will not, for example, typically find moral realists accepting that the minimalistic notion of *assertion* is legitimate for certain contexts; rather, they will generally despise the very idea (along with the minimalistic versions of truth, fact, etc.). Similarly, the quasi-realist may well claim that the robust version of *assertion* is a mere philosopher’s fancy, and perhaps not a well-formed idea at all. Perhaps this dispute could one day be settled; perhaps one of the disputants is, after all, uniquely correct. But it is also possible (though I am not arguing that it is the case) that neither is uniquely correct; perhaps there is nothing that determines the correctness of one conception over the other; perhaps the concept of *assertion* is simply indeterminate across the relevant cognitivist/noncognitivist gap.

I want now to return to the possibility of settling this dispute by an appeal to convenience and expedience, though this time I propose to understand these matters (albeit vaguely) in a broader, more psychological, sense than Carnap intended. My worries with this suggested decision procedure arise before we even begin to wonder about what the practical upshot of moral noncognitivism is. My concern arises because the question “Expedient to whom?” seems a perfectly fair yet unanswered response.

Many pragmatist suggestions speak in a cavalier fashion about “what is useful to our practical purposes” without pausing to wonder to whom the “our” refers; the common background assumption is that there is a convergence in practical requirements. But this assumption is, upon reflection, wildly implausible. What if noncognitivism suits some people while cognitivism suits others? Then one precisification of “assertion” (say) will be practically expedient for some, while an opposed precisification will be practically expedient for others. This is reminiscent of an old complaint against the Pragmatist school of philosophy, one that Bertrand Russell heatedly made in his essay on the topic (and on every possible occasion thereafter): “One gathers … that a Frenchman ought to believe in Catholicism, an American in the Monroe Doctrine, and an Arab in the Mahdi” (Russell 1910: 97). That Russell had an accurate or fair image of his target is unlikely (see Haack 1976), but here my worry concerns nothing so grand as Pragmatism as a school of philosophy or a theory of truth. My target is just one kind appeal to pragmatic considerations in the hope of settling a philosophical dispute, and here, it seems to me, Russell’s question remains a live one.

Regarding the practical upshot of noncognitivism’s being true or false, is it really reasonable to suspect that there will be divergence among individuals’ interests? Alasdair MacIntyre’s analysis of the Bloomsbury group might serve as an illustration (MacIntyre 1981). MacIntyre interprets the Bloomsbury circle as developing a peculiar moral discourse: Surrounded by a community that employed moral vocabulary in an assertoric way, the members of the Bloomsbury group (unwittingly) did otherwise. Impressed with G.E. Moore’s elevation of beauty, love, and knowledge to the pedestal of “intrinsic moral values,” this group of intellectuals used moral language as a manipulative tool to persuade others of the delicacy (and thus superiority) of one’s own aesthetic sensibility. John Maynard Keynes recalls that about these matters it was “useless and impossible to argue” (Keynes [1949] in Rosenbaum 1995: 87). He goes on:
In practice, victory was with those who could speak with the greatest appearance of clear, undoubting conviction and could use the best accents of infallibility. Moore at this time was a master of this method—greeting one’s remarks with a gasp of incredulity—*Do you really think that*, an expression of face as if to hear such a thing reduced him to state of wonder verging on imbecility, with his mouth wide open and wagging his head in the negative so violently that his hair shook. *Oh!* He would say, goggling at you as if either you or he must be mad; and no reply was possible. (ibid.)

MacIntyre’s thesis that noncognitivism was true of the Bloomsbury group while not true of the wider populace is an interesting but knotty claim—but it is not, in any case, our present concern. Rather, I am suggesting that it might have *suited the Bloomsbury group’s purposes if noncognitivism were true*—because, roughly, it would *validate* their practices—while not serving the purposes of others. Thus the proposal that we should settle any indeterminacy between noncognitivism and cognitivism—and thus between moral skepticism and moral naturalism—by reference to “what is in our practical interests” is seen to be seriously problematic.

**The benefits of being horribly wrong**

The pragmatic pay-offs of noncognitivism versus cognitivism may be inscrutable and divergent, but surely we can hope to do better at the other node of potential metaethical indeterminacy: moral naturalism versus the error theory? Here the naturalist might feel more confident that there is nothing to be said, pragmatically, in favor of the error theoretic position. Here the naturalist might feel satisfied with Lewis’ diagnosis, assured that if all that lies between moral naturalism and moral skepticism is “temperament,” then he, the naturalist, may declare victory.

However, in my opinion, if the argument has come to this point, the moral naturalist has no particular grounds for complacency, for there are, I think, considerable *pragmatic* claims to be made in favor of the moral error theory.

The moral error theory tells us that we have been massively mistaken about the world, about ourselves, and about the relation between the two. Finding out that one has been massively mistaken is, it can be assumed, an uncomfortable state, and being uncomfortable is unpleasant. Finding out, moreover, that our mistake has been as great as the moral error theorist would have us believe—ubiquitous across all human societies, through all history, pervading nearly every aspect of our social existence—is likely to make us uncomfortable to the point of positive distress. But one needs to look deeper. Being uncomfortable can be a useful state (after all, pain serves a good biological purposes), and distress can, ultimately, be instructive.

By way of leading you in to my thinking, consider what we have learned from experimental psychology about confabulation. There are various kinds of cognitive impairment (to memory, to perception, etc.) for which the subject will compensate by creating a false narrative to “fill in the gaps.” Someone suffering from Korsakoff’s syndrome fails to admit, even to herself, the severe memory loss she suffers, but rather invents elaborate stories to cover her confusion. Asked why she is in the hospital, for example, the patient might genuinely assert that she is visiting someone or applying for a job there. Sometimes stroke patients with partial paralysis will deny the paralysis, inventing sincere excuses for why that part of their body cannot perform actions. (See Hirstein 2005; Schnider 2008.) Fascinating as
these kinds of case study are, they are really just the lurid side of the phenomenon of confabulation, much of which is far more routine and commonplace. We all superimpose justifications and explanations onto our actions when the real sources of motivation are hidden from us; we all string together coherent life narratives from fragmented and distorted memories; we all confabulate. (See Wheatley 2009.)

This is a disconcerting truth when its full extent is apprehended (“I am confabulating nearly all the time!”). It forces us to adopt a new perspective on who we are and how we work. Experimental psychology can bring us to see that our intuitive everyday image of ourselves and our lives is, in many ways, badly mistaken. Yet, I claim, this unsettling apprehension is often a good thing. It is good not merely because it is true, but because it is useful. Of course, we tend to compartmentalize this analytical awareness, and quickly return to confabulating as soon as we relax our attention. And that’s fine. What I am speaking in favor of is the value of regularly returning to the critical perspective from which we realize that we are more opaque and baffling beings than everyday life presupposes.

Consider this scenario. You are riding in a bus through a city, absentmindedly watching pedestrians and street scenes, when the thought strikes you: “What odd creatures we are.” And suddenly you undergo a Desmond-Morris-style gestalt shift: You think of the pedestrians under the description “hairless upright social apes”—you literally see them as hairless upright social apes. Suddenly the city appears like a swarming nest of primates, commonplace urban activities come to seem like bizarre arbitrary rituals, the projects and concerns that ordinarily occupy our minds are seen as weird and futile. And perhaps the oddest thought of all is: “Hey, I’m one of those naked apes, too!”

When one thinks along these lines, it’s not that one all of a sudden forms new beliefs, for we all know (those of us that are sensible, at least) that we are hairless upright social primates, with bodies and minds shaped by Pleistocene Africa. But this knowledge is compartmentalized during our everyday lives; it takes an effort to attain the critical distance to feel the truth of such beliefs, to be dumfounded at the sheer strangeness of it all.

And that feeling—a feeling of alienation from the customary images of oneself and the society one inhabits—is, I contend, a valuable perspective to adopt. Like a full-blooded encounter with the aesthetic sublime, it is not an entirely pleasant experience. The sublime can make one weak in the knees with confusion and awe. Yet we seek out such encounters (sometimes) because in that feeling we sense that we are in touch with something authentic, like we have stepped back from our familiar humdrum concerns and tasted the enormity of the universe and the fragility of our place in it.

This is the benefit I would like to advertise of seeing ordinary beliefs as horribly wrong. It is good to be epistemologically shaken. It reminds us of how shockingly ignorant we are and how mysterious everything really is. It is both a corrective to epistemic complacency and a spur to intense reflection and inquiry. Discovering that everyone is horribly wrong about something fundamental to our lives is not just sobering andvaluably humbling, but

---

8 Of course, it is not just the phenomenon of confabulation that shows this. The broader claim that experimental psychology has revealed that we are, in a multitude of ways, routinely mistaken about what is really going on with our motivations, emotions, and belief-formation processes, is far less contentious and far too well documented to require citations.

9 This claim that we have “minds shaped by Pleistocene Africa” should not be over-interpreted as an endorsement of any strong form of Evolutionary Psychology. The prehistorically forged human mind may be designed to be a plastic and generalized problem solver.
intellectually exciting. Endorsing an error theory puts hairs on your chest, epistemologically speaking.

Well, that’s just me. If the last few paragraphs fail to speak to you, then that’s okay; that, indeed, is my central point. I am not claiming that every person would benefit from recognizing his/her profound error and therefore we should all be error theorists. Remember that the present dialectic presupposes that there are already some decent arguments in favor of the moral error theory (and decent arguments for moral naturalism), so the pragmatic considerations just raised are supposed to supplement solid arguments, not stand in place of them. But nor am I advancing the amended thesis that every person would benefit from recognizing his/her profound error and therefore \textit{if we have decent arguments in favor of the error theory} then we should all be error theorists. I am in fact nowhere claiming that every person would benefit from recognizing his/her profound error. Perhaps some people would and some people would not. My point is (again) that even in the case of moral naturalism versus the moral error theory, the pragmatic cost-benefit analysis is complex, and \textit{there is no discernible reason to expect convergence}.

I can imagine critics complaining that even if there do exist such intellectual benefits to be gained from an apprehension of the depths of human epistemic fallibility, surely they pale beside the losses that will be incurred from an endorsement of a moral error theory. Surely (the complaint goes) a belief in the moral error theory will lead to a breakdown of good citizenship, to a loss of motivation to enter sincerely into cooperative ventures, to cheating one’s friends when the chips are down, to selfishness and debauchery, and so forth. But these vague fears are expressed far more often than any scrap of evidence in their favor is offered. While I agree that moral belief contributes to an individual’s “cooperative motivation” (to choose an umbrella term) (see Joyce 2006: chapter 4), the idea that moral belief is the \textit{only} thing that does so is a hypothesis as doubtful as it is depressing. And even if the hypothesis were true, the claim that a breakdown of cooperative motivation is the inevitable outcome of embracing the moral error theory implies the similarly doubtful hypothesis that moral belief is the only thing that \textit{can} motivate cooperation. No, there is no evidence that moral skepticism leads to crimes and misdemeanors. I have met many moral error theorists, and am yet to encounter one who wasn’t perfectly civilized or was any less trustworthy than anyone else. By contrast, if one reflects upon the worst mass criminals that humanity has produced, one is likely to encounter not skeptics but individuals moved by moralistic fervor. (See Hinckfuss 1987; Garner 2010.)

Of course, these observations don’t settle the matter. It is entirely possible that, despite my anecdotal evidence, in fact moral error theorists are always less flourishing and always inferior citizens when compared to moral believers. Or perhaps it goes the other way round. What I am seeking to establish is that we shouldn’t jump to any conclusions regarding these empirical matters, and in particular shouldn’t assume that there will be convergence among individuals one way or the other. Though there may well be practical costs to being a moral error theorist, I am content on this occasion to attempt to publicize some of the neglected practical benefits.

The idea that our interests are best served by a methodology that always favors theories which mesh with our common sense is, for my money, short-sighted and disappointing. Perhaps some individuals require that kind of constant reassurance—perhaps many do—but not all do. Some find the prospect that common sense is massively mistaken an object of contemplation that is, while anxiety-provoking (partly \textit{because} anxiety-provoking),
profoundly rewarding. Given the choice between a theory that vindicates some aspect of common sense and a theory that overturns it, some people will find the latter more valuable. I cannot resist here quoting Bertolt Brecht’s gloriously scathing assessment: “I’m not writing for the scum who want to have the cockles of their hearts warmed” (Willett 1957: 14).

**Quine’s sectarianism and ecumenicalism**

I have argued that the debate between the moral naturalist and the moral skeptic—which I broke down into (i) that between the moral cognitivist and the moral noncognitivist, and (ii) that between the moral naturalist and the moral error theorist—may be undecidable. Key concepts like *assertion* or *value* may be sufficiently indeterminate to permit all aforementioned theoretical possibilities. The prospect of avoiding this conclusion by an appeal to pragmatism has been explored and rejected; pragmatic solutions presuppose an implausible convergence in practical interests. It should be stressed that I have not argued that we do face this indeterminacy, only that we might. For all I have said here, it is possible that the moral naturalist will triumph *tout court*. Or perhaps the moral skeptic will. But let us suppose for now that these or similar arguments favoring undecidability are sound. Where do we go from here?

Having discussed Lewis and Carnap, it seems apt to end with the philosopher who best connects them: W.V. Quine. Quine famously argues for the underdetermination of scientific theory: that logically incompatible global theories may be equally consistent with all possible evidence and equally satisfactory with respect to internal theoretical virtues (e.g., simplicity and clarity) (Quine 1960; 1975). As to the “Where do we go from here?” question, Quine vacillated. In certain moods he espoused a *sectarian* response, according to which we should continue to endorse our favorite theory as robustly as ever, rejecting all alternatives as false or meaningless (Quine 1981: 21; 1986). But on other occasions he preferred an *ecumenical* response, whereby we accept that indeterminacy reveals all satisfactory though incompatible theories to be true (Quine 1981: 29; 1989).

We face a similar decision regarding metaethical indeterminacy. Suppose, like Lewis, one is inclined to endorse a form of moral naturalism. And suppose the possibility that I have discussed turns out to be a reality: that another person could, with equal legitimacy, endorse some form of moral skepticism. What attitude should the first person take?—to push on with moral naturalism, declaring moral skepticism bankrupt (i.e., sectarianism), or perhaps to take a more tolerant view of alternatives (i.e., ecumenicalism)? The moral skeptic, of course, faces the same dilemma.

Quine became aware of this conflict in his work, and tended to plump for sectarianism. He even inserted comments into later editions of *Theories and Things* in order to distance himself from the ecumenical noises from his earlier self. He diagnoses his own indecisiveness on this matter by saying: “The fantasy of irresolubly rival systems of the world is a thought

---

10 Readers might be aware that I have argued for the error theory in the past (Joyce 2001, 2006, 2011, forthcoming), and I dare say that in the future I will return to doing so. Despite my sympathetic attitude towards Lewis’ pluralistic position, I have not officially conceded that his “imperfect claimant” (or any other contender) is close enough to warrant our endorsement as a *morality*. But even if I were to make that concession, continuing to argue for an error theory at the level of metaethics is consistent in every way with favoring pluralism at the metametaethical level. (See note 12 below.)

11 Carnap taught Quine; Quine taught Lewis.
experiment out beyond where linguistic usage has been crystallized by use” (Quine 1990: 100).

Faced with the nodes of indeterminacy highlighted in this paper, I would counsel neither sectarianism nor ecumenicalism in particular, but rather what might be called “metaethical ambivalence.” This perspective begins with a kind of metametaethical enlightenment. The moral naturalist espouses moral naturalism, but this espousal reflects a mature decision, by which I mean that the moral naturalist doesn’t claim to have latched on to an incontrovertible realm of moral facts of which the skeptic is foolishly ignorant, but rather acknowledges that this moral naturalism has been achieved only via a non-mandatory piece of conceptual precisification. (This describes Lewis’ tolerant view.) Likewise, the moral skeptic champions moral skepticism, but this too is a sophisticated verdict: not the simple declaration that there are no moral values and that the naturalist is gullibly uncritical, but rather a decision that recognizes that this skepticism has been earned only by making certain non-obligatory but permissible conceptual clarifications.

This enlightened awareness of the legitimacy of one’s opponent does not, however, suffice for metaethical ambivalence. For the next step we can turn again to Quine, recalling that even when advocating sectarianism he does not encourage adherents of one scientific theory to blinker themselves uncritically to other possibilities. What is at stake for Quine is the permissibility of applying the truth predicate to alternative theories to one’s own; but even in those moods when he claims that one can call only one’s own favored theory “true,” he still allows that one can “oscillate between … theories for the sake of added perspective from which to triangulate on problems” (Quine 1990: 100). And this is, perhaps, indicative of the most sophisticated pluralism of all. The enlightened moral naturalist doesn’t merely (grudgingly) admit that the skeptic is warranted in his or her views, but is willing sometimes to adopt the skeptical position in order to gain the insights that come from recognizing that we live in a world without values. And the enlightened moral skeptic goes beyond (grudgingly) conceding that moral naturalism is reasonable, but will sometimes assume that perspective in order to gain whatever benefits come from enjoying epistemic access to a realm of moral facts. Such a metaethical ambivalence will require a temperament that favors neither the conservative nor debunking attitudes mentioned by Lewis, but rather an intellectual courage to cut oneself loose from the comforting familiarity of assuming that there is always a fact of the matter.

Victoria University of Wellington

REFERENCES:


12 Broadly parallel to David Wong’s *moral ambivalence* (2006: 20 ff). I take it that just as Wong’s moral ambivalence is ambivalence about moral systems—an ambivalence that occurs at the metaethical level—so too my metaethical ambivalence is ambivalence about metaethical theories—an ambivalence that occurs at a metametaethical level. For insightful discussion of doxastic ambivalence, see Roorda 1997.


