

Moral and epistemic normativity: The guilty and the innocent

Richard Joyce

Penultimate draft of paper appearing in C. Cowie & R. Rowland (eds.),
Companions in Guilt Arguments in Metaethics (Routledge, forthcoming).

1. Introduction

The error-theoretic view in metaethics is the position that moral discourse essentially misdescribes reality. This position has three components.

First, in order to misdescribe reality, a discourse must be in the business of *describing* reality. One might object that moral discourse is not in the business of describing but rather of *evaluating* reality (actions, persons, states of affairs, etc.). However, of course the error theorist doesn't deny that moral discourse is a way of evaluating reality, but she maintains that moral discourse accomplishes this end via speakers committing themselves to reality being a certain way: when one evaluates the act of torturing innocents as morally wrong, for example, one commits oneself to that action having the property of wrongness. This contrasts with a purely noncognitivist view of moral judgment.

Second (following from the first), if one is describing or committing oneself to the world being a certain way, then the possibility arises that the world fails to play along. This is what the error theorist thinks about moral discourse: that reality just isn't furnished with the properties and relations that would be required in order to render moral judgments *true*; thus users of moral discourse are making a mistake. This contrasts with the view of the "success theorist," who maintains that the world does contain the properties required to render some (not *all*, obviously) moral judgments true.¹

The third component is indicated by the word "essentially." The mistake that moral discourse makes is central to the very conceptual framework, such that were we to attempt to expunge the erroneous component, thereby committing ourselves to the world being a way that the world actually is, then the resulting discourse wouldn't count literally as a "moral" discourse anymore. (Analogy: our erroneous "unicorn discourse" was essentially committed to horse-like animals with a single horn; when we discovered that the world contains no such creatures, we were not free to expunge the problematic "horn clause" and declare that we had been talking about actual horses all along.) This contrasts with the view of the revisionist success theorist, who accepts that although we often do make mistakes when employing moral discourse (or even, conceivably, *always* do), the conceptual framework that would result from the erroneous elements being excised would (i) still warrant the label "morality" and (ii) sometimes accurately describe reality.

When discussing the error-theoretic view in metaethics it's often useful to compare another kind of error theory with which we're more familiar: atheism. It's

¹ Some success theorists are realists, some are constructivists. Often it's hard to know where to draw the line.

useful on this occasion because it allows us to observe a basic point: what it takes to be an atheist is one thing; what arguments might lead one to that position are something else. All it takes to be an atheist is to disbelieve in any gods. (It doesn't, incidentally, involve any claim to knowledge on this score.) But there are potentially many routes to this disbelief, and I hazard to suggest that the typical atheist is likely to think that the grounds for disbelief are over-determined. One atheist might be impressed by arguments that leave another atheist cold.

And so it is with moral error theory. All it takes to be an error theorist is to endorse the above three theses. But there are potentially many routes to this disbelief, and one moral error theorist might be impressed by reasons that leave another moral error theorist cold. In what follows, this observation (which is, oddly, often overlooked) will be important.²

Opponents of the moral error-theoretic view typically fall into one of two camps: some accuse the error theorist of making a conceptual mistake, some accuse her of making an ontological mistake. The former think that the error theorist has misidentified what we commit ourselves to when we engage in moral discourse—perhaps claiming that the error theorist foists on moral speakers extravagant and *recherché* commitments that they do not make (or, at least, *need not* make). The latter accept that the error theorist has done decent job of identifying the commitments of moral discourse, but think that the error theorist has misidentified what the world is like: in fact, reality does contain the properties and relations necessary to render many moral judgments true. (In the past, I have called these the “concessive” and “head-on” strategies, respectively.³)

The subject of this chapter—the so-called *companions in guilt argument* (CGA) against the moral error theorist—can be thought of as a version of the latter strategy. The proponent of the CGA claims that the very quality that the error theorist ascribes to morality—a quality supposedly sufficiently problematic to sink it—is one that she willingly accepts elsewhere in other kinds of discourse about which she is a success theorist; therefore the error theorist must be mistaken, by her own lights, in thinking that committing oneself to this quality renders one's judgments untrue.

My goal in this chapter is to defend the moral error theory from the CGA. In section 2 I will discuss the inherent limitations of the CGA, showing that at best it undermines one kind of argument for the error-theoretic position, but not the metaethical position itself. In section 3 I will show that the CGA doesn't even achieve that.

² It is worth mentioning a disanalogy: since all it takes to be an atheist is to disbelieve in any gods, then a noncognitivist about religious discourse counts as an atheist. But I don't think many actual atheists follow this route; most (I strongly suspect) are error theorists, maintaining that religious discourse satisfies the three components outlined.

³ Joyce & Kirchin 2010; Joyce 2011.

2. Restricting the companions in guilt argument

Let me begin this section by contrasting the ways that the two excellent editors of this collection have in the past presented the CGA. And then I will make the impolitic move of expressing a preference!

Richard Rowland, who is in favor of the CGA, presents it as follows (2013: 1):

- (1) According to the moral error theory, there are no categorical normative reasons.
- (2) If there are no categorical normative reasons, then there are no epistemic reasons for belief.
- (3) But there are epistemic reasons for belief.
- (4) So there are categorical normative reasons (2, 3).
- (5) So the error theory is false (1, 4).

Christopher Cowie, who is opposed to the CGA, presents it as follows (2014a: 408):

1. *Parity premise*: if the arguments for the moral error theory are sufficient to establish its truth, then those arguments (or appropriate analogues of them) are also sufficient to establish the truth of the epistemic error theory.
2. *Epistemic existence premise*: the epistemic error theory is false.
3. So, the arguments for the moral error theory are not sufficient to establish that the moral error theory is true.

There are several differences in presentation (for a start, some of Cowie's premises compress several of Rowland's), but the one to which I'd like to draw attention is that Rowland thinks that the CGA refutes the moral error theory, whereas Cowie thinks that it refutes *the arguments for* that position.

The argument for error theory that both authors have in mind is one offered by John Mackie (1977) and developed by others such as Jonas Olson (2014) and myself (Joyce 2001). According to this argument, moral normativity is imbued with a certain kind of strong authority. Moral prescriptions aren't pieces of advice on how to satisfy our desires; they often seem to require us to perform or omit actions irrespective of our desires. One way of trying to make sense of this practical authority is to claim that morality implies that we can have *reasons* to act irrespective of our desires. And these reasons would be, moreover, non-institutional—not the mere construct of a human-made normative system.⁴ This is the conceptual step of this error-theoretic argument. The ontological step is then to argue that there are no such reasons. Mackie call his version of this two-step argument “the argument from queerness”; he thinks that moral authority is just too *strange* to countenance within a naturalistic worldview.⁵ Though he doesn't mention reasons in his primary presentation of his argument

⁴ See discussion in Joyce & Kirchin 2010: xviii-xix.

⁵ I'm much inclined to think that it's high time to retire the term “queerness” in this context, but I'll continue (somewhat reluctantly) to use it when explicitly denoting Mackie's argument.

(1977: 38-42), elsewhere he suggests that “to say that they [objective prescriptions] are intrinsically action-guiding is to say that the reasons that they give for doing or for not doing something are independent of that agent’s desires or purposes” (Mackie 1982: 115). I will call this version of the two-step argument—with reasons made explicit—“the argument from reasons” for moral error theory.

The CGA aims to show that the argument from reasons would, by parity of reasoning, suffice to prove that we never have any reasons *for belief*—which is supposedly an unacceptable or even self-undermining result.⁶ I will not get to the epistemic premises of the CGA till much later in this paper; first I want to focus just on premise (1) of Rowland’s version. How reasonable is his claim that according to the error theorist there are no categorical normative reasons? One could quibble over whether the phrase “categorical normative reasons” quite captures all that might be bothering the error theorist, but let that pass for now. My main concern is that the premise conflates the error-theoretic position with an argument for it.

To see what I mean, reflect again on the comparison of atheism, and consider one well-known argument for that position: the argument from evil. In a nutshell, this argument goes like this: God is by definition omniscient, omnipotent, and benevolent; therefore God would know about evil and would be able to and would want to eliminate evil; but evil evidently has not been eliminated; therefore God doesn’t exist. It’s about as pithy an argument against theism as you’re likely to find.

I submit that Rowland’s premise (1) is analogous to this:

X: According to the atheist, the existence of evil shows that God doesn’t exist.

(X) is clearly false. Certainly an atheist *might* think that the existence of evil shows that God doesn’t exist, but the atheist need not endorse this argument. I myself have no settled opinion on whether the argument from evil is sound—I confess that I haven’t thought much about it since I wrote an essay on the topic as an undergraduate—yet this doesn’t shake my firm confidence in atheism, for that disbelief doesn’t depend on the argument from evil in particular. The atheist *as such* doesn’t endorse any particular argument for the non-existence of God. If one thought otherwise—if one thought that the atheist *as such* endorsed the argument from evil, then a refutation of that argument would suffice to show that atheism is false—which is clearly an unacceptable result. And yet the analogous conclusion to Rowland’s CGA has such strong ambitions: “So the error theory is false.”

Cowie’s presentation of the CGA seems better, since it makes reference to *the arguments for* the moral error theory. However, it would be better still, in my opinion, if the premise in question were more explicit in specifying a particular argument, leading to the following conclusion of the CGA:

⁶ The CGA might focus on some other kind of “innocent” normative system instead. For example, Guy Fletcher (2017) proposes *prudence* as the innocent companion. In this chapter I discuss only the analogy with epistemic norms, but I believe that the arguments I offer will, *mutatis mutandis*, apply to other versions of the CGA.

3. So, *the argument from reasons* for the moral error theory is not sufficient to establish that the moral error theory is true.

All of this might seem like pointless hair-splitting if the argument from reasons were the only viable ground one might have for endorsing the moral error theory. But this is not the case. I will argue in what follows that the argument from reasons is neither necessary to establish the moral error theory, nor (and this is perhaps more surprising) sufficient.

Let us begin by returning again to the comparison of atheism, and asking “On what grounds might someone be an atheist?” I’m in a bit of a bind in addressing this question here, since, on the one hand, the main point I want to make is that there are many different answers and the matter is quite complicated, but, on the other hand, it is just an illustration and I don’t want to go on for pages and pages—so what follows will be swift but dense.

THE CASE FOR ATHEISM:

Gods are ascribed some strange qualities, unlike anything else. Some of the attributes ascribed to gods are deeply mysterious (*how* did God create the universe?—the answers are evasive and vague), or even potentially incoherent (such as the paradoxes of omnipotence), and some seem contrary to what we can apparently plainly see (e.g., the problem of evil). One might try to argue that gods don’t need to be ascribed such strange qualities, but on reflection these qualities seem rather the whole point of the conceptual framework. We need strong evidence to support belief in strange things. But in the case of gods there is no such evidence; there is no phenomenon we know of that requires a theistic explanation; everything that is put forward as evidence is explicable in non-theistic terms. It is, however, not surprising that people should believe in gods, since there are plausible explanations (psychological, anthropological, and evolutionary) as to why such belief systems might emerge and seem compelling (even if they are erroneous). Religious belief systems often contradict each other (concerning, e.g., how many gods there are), which in particular draws attention to the question of by what means people would have epistemic access to these facts (though that question would be pressing even if there were only one unanimous religion). The atheist is unlikely to have much respect for an epistemology of *faith* or be impressed by claims about being the “chosen people” (of God or the gods). What is more likely to impress the atheist is the observation that for over two thousand years the best intellects available have tried to demonstrate the existence of gods and failed.

I call this “the case for atheism” merely for ease of reference, but in fact it’s just a rough and incomplete sketch that’s intended to represent a fairly mundane body of considerations speaking in favor of atheism. I have no intention of trying to advocate any of these arguments here; I am not even claiming that they are sound. My point is just to illustrate that the case for atheism is best conceived of as a web of loosely interlocking and complementary arguments, which jointly make disbelief more reasonable than either belief or agnostic fence-sitting.

It is true that some of these considerations might, if successful, be sufficient as stand-alone arguments. You’ll have noticed, for example, that the problem of evil makes an appearance in the atheist’s case, and it is tempting to suppose that if this argument were sound then none of the other considerations would be necessary. But I’m inclined to think that even this is something of an idealization of how

philosophical argumentation proceeds. Philosophical arguments rarely land a knock-out blow; there's always room for disagreement and negotiation. So even someone who's highly sympathetic toward the argument from evil should have the epistemic modesty to accept that the argument *might* be overturned by, say, some clever theodicy that hasn't yet been properly articulated, and hence should accept that the argument at best shows that its conclusion is *probable*; for higher credence, additional arguments to the same conclusion are needed. And if the atheist were to lose confidence in the argument from evil, it's probably no great loss to the overall case. It's certainly possible that the subtraction of that plank might be enough to shift someone's credence from atheism to agnosticism (or even to theism), but it's much more likely that the atheist considers the case for disbelief sufficiently strong that even without the argument from evil atheism remains solidly reasonable.

The general case for moral error theory is very similar, I think.

THE CASE FOR MORAL ERROR THEORY:

Moral properties are ascribed strange qualities, unlike anything else. Some of the qualities of moral properties are deeply mysterious (how, exactly, does a situation place a *demand* on someone independent of his/her desires?), or even potentially incoherent (such as the notion of free will that underlies moral culpability), and some seem contrary to what we can plainly see (e.g., that reasonable people will converge in their moral outlooks). One might try to argue that moral properties don't need to be ascribed such strange qualities, but on reflection these qualities seem rather the whole point of the conceptual framework. We need strong evidence to support belief in strange things. But in the case of morality there is no such evidence; there is no phenomenon we know of that requires a moral explanation; everything that is put forward as evidence is explicable in non-moral terms. It is, however, not surprising that people should believe in morality, since there are plausible explanations (psychological, anthropological, and evolutionary) as to why such moral systems might emerge and seem compelling (even if they are erroneous). Moral belief systems often contradict each other (concerning, e.g., the rights of the individual versus the good of the community), which in particular draws attention to the question of by what means people would have epistemic access to these facts (though that question would be pressing even if there were only one unanimous morality). The error theorist is unlikely to be impressed by an epistemology of *intuition*. What is more likely to impress her is the observation that for over two thousand years the best intellects available have tried to illuminate the nature of morality and failed.

Again, my calling this "*the* case" for moral error theory isn't to be taken too literally; it's largely for ease of referring back to it. And, again, my intention isn't to advocate this dense cluster of considerations, or even claim that they are persuasive (though it so happens that I find them all very persuasive). The point is to see that the case for the moral error theory is a complex of interlocking and complementary arguments. You'll notice, however, that reference to *reasons* doesn't appear anywhere, so where does the so-called argument from reasons belong in the case for moral error theory?

One of the puzzling aspects of moral normativity is that it seems to have a kind of practical authority: morality tells us what to do (and what not to do, and what kind of people to be, etc.) in ways that we can't just beg off if we don't care or if it doesn't suit us. Someone who says "Yeah, but I just don't care about morality, so why shouldn't I steal [e.g.] if I can get away with it?" might meet with the response

“Because it’s just *wrong!*—Even if you don’t care and can get away with it, you just mustn’t do it!” A challenge for metaethicists is to try to articulate this inchoate insistence more clearly. References to *reasons* enter at this stage. “You have a reason not to ϕ ” is an attempt to make some sense of the practical nature of “You morally ought not to ϕ .” And since the moral “ought”-claim is categorical (not advice on how the addressee may satisfy his/her ends), then the reason will also be categorical. The argument from reasons then proceeds to the ontological step of asking whether such categorical reasons for actions exist, and presents arguments in support of the negative answer.

One thing I’d like to stress is that the reference to reasons has (at least to my mind) the status of a *hypothesis* within the argument.⁷ The error theorist perceives that moral properties are imbued with *some* kind of unusual practical relevance, and seeks to tease open this quality in order to examine it more thoroughly. The idea that moral properties imply *reasons* is an obvious avenue to explore. The argument from reasons kicks in at this point, declaring that if moral properties imply reasons, and the kind of practical reasons implied (i.e., categorical and objective) do not exist, then there are no moral properties.

But maybe the hypothesis turns out to be an unhelpful one; maybe even the error theorist comes to see that reference to *reasons* turns out not to be an illuminating way of understanding the practical nature of moral authority. In this case, it is available to the error theorist to retreat and regroup: “Okay, forget about reasons, then,” she says, “there is still something deeply puzzling about the practical authority with which moral properties are imbued, to the extent that disbelief seems reasonable.” Perhaps, indeed, *nothing* very illuminating can be said about this authority, but if so then it is just more grist for the error theorist’s mill. After all, it isn’t clear that it is the dialectical responsibility of the error theorist in particular to try to articulate something that may be, at bottom, conceptually confused and indefensible. As Nadeem Hussain nicely puts it:

[P]art of what might attract one to an error theory about the moral in the first place is the thought that there is something deeply mysterious about moral concepts and the moral properties they supposedly pick out. Morality, one thinks, is an ideology, and mystification is the life-blood of ideologies. Surely it would be no surprise, then, if some fundamental unclarity is essential to morality’s ideological role. Given this essential unclarity, no surprise, then, if moral concepts seem to systematically escape analysis. (2004: 155-156).

The point I am emphasizing is that before we get to “According to the moral error theory, there are no categorical normative reasons” there are *two* argumentational decisions that the error theorist must make. First, she must decide that it is the unusual “practical authority” of moral normativity that is its problematic feature. Next, she

⁷ This may not have been clear to me when I wrote *The Myth of Morality* in 2001, but it is explicitly acknowledged in Joyce 2008.

must decide that reference to “categorical reasons” is a reasonable way of trying to unpack this authority. The moral error theorist *as such* makes neither of these moves.

The CGA, then, would at most refute a particular argument for the moral error theory; it would hardly show that the position itself is unreasonable, and nor would it show that there are no sound arguments for that position. But even if its ambitions were scaled down in this manner, the CGA may still be highly important. Advocating the argument from reasons may not be *necessary* for being a moral error theorist, but it’s certainly a well-known argument and one that moral error theorists (including myself!) have advocated. So the CGA, if successful, would remove an important support from the error theorist’s case. In the following section I shall argue that the CGA does not achieve even this more modest ambition.

3. Refuting the companions in guilt argument

We are now considering a restricted version of the CGA: one that concludes that the argument from reasons fails to establish that the moral error theory is true. Here is how I stated the latter argument (though not under this title) in *The Myth of Morality* (2001: 77):

THE ARGUMENT FROM REASONS

- 1) If x morally ought to ϕ , then x ought to ϕ regardless of what his desires and interests are.
- 2) If x morally ought to ϕ , then x has a reason for ϕ ing.
- 3) Therefore, if x morally ought to ϕ , then x can have a reason for ϕ ing regardless of what his desires and interests are.
- 4) But there is no sense to be made of such reasons.
- 5) Therefore, x is never under a moral obligation.

The focus of our attention here should be on premise (4). First, let me say that the phrase “there is no sense to be made” should not be over-interpreted; the proposition “But no such reasons exist” is all that is really meant. In what follows I’ll adopt Rowland’s label for the kind of reasons in question: “categorical normative reasons.” The proponent of the CGA takes this premise to have the unacceptable implication that we never have any categorical reason to believe anything (not even this). Premise (4) may have this implication either directly or indirectly; I will discuss these possibilities in turn.

Premise (4) might be thought to have this consequence directly if it is interpreted as making the entirely general claim that no categorical normative reasons exist *anywhere*. But it should be obvious that the argument requires nothing so strong. The argument from reasons is restricted to discussing reasons *for action*, so premise (4) is simply the claim that no categorical normative reasons *for action* exist. Consideration of the kind of arguments that the error theorist is likely to use in support of (4) should make this apparent. She doesn’t merely throw up her hands and declare that such reasons are too weird to countenance. Rather, she carefully and critically examines

theories of practical rationality and comes to the conclusion that some version of instrumentalism—where an agent’s reasons for action are inextricably linked to his desires and interests—has more going for it than any version of non-instrumentalism. (I won’t rehearse the arguments that might lead to this conclusion here; for current purposes it doesn’t matter what they are, just that there *are* such arguments that one might find persuasive.) But the fact that one might be a committed instrumentalist about practical reasons exerts zero pressure on one to be an instrumentalist about other areas. Instrumentalism about doxastic reasons is, on the face of it (and possibly at bottom, too) quite implausible. Instrumentalism about warrant for emotions is also easy to resist. The proponent of the argument from reasons for the moral error theory need not think otherwise.

Suppose, then, that reflection on epistemic norms leads us to see that there *are*—that there *must* be—such things as categorical normative reasons in that sphere.⁸ Perhaps we are forced to this conclusion because the very act of denying it is self-undermining: we would have, *ex hypothesi*, no reason to believe anyone who advocates an epistemic error theory (see Cuneo 2007; Shah 2011). Perhaps even the moral error theorist can’t resist this conclusion. But I cannot see why the moral error theorist should even particularly *want* to resist the conclusion. Even if she has vigorously and rigorously advocated the argument from reasons, the fact that there exist categorical normative reasons *in the epistemic sphere* needn’t undermine her advocacy one iota.

One might respond that if the moral error theorist admits the existence of categorical normative reasons in the epistemic sphere, then she has opened the door to an ontological category the resistance to which motivates her moral skepticism. “If you’ve acknowledged the existence of such reasons *there*,” the response goes, “then you’ve lost your incentive and grounds for resisting such reasons *here*.”

However, I think that this response badly misunderstands the dialectic. It’s true that one might be an error theorist because some quality of morality just seems ontologically weird, in which case showing that the error theorist admits this very quality elsewhere would be telling. But the error-theoretic argument we are currently considering has gone well beyond simply finding something weird. Perhaps there’s nothing particularly “weird” about categorical normative reasons *per se*; it’s simply that, when it comes to practical rationality, instrumentalism is true and non-instrumentalism is false (and moral concepts presuppose the truth of practical non-instrumentalism). Indeed, in *The Myth of Morality* this was a leading motivation for my interpreting Mackie’s argument from queerness in terms of the argument from reasons: the hope that focusing on *reasons* would provide more analytical traction and take the debate beyond merely throwing up one’s hands in bewilderment.

Another way that premise (4) might have the unacceptable implication that we never have categorical reasons for belief is indirectly. Even if one recognizes that premise (4) is restricted to the denial of categorical normative reasons *for actions*, one might claim that nevertheless the arguments that the error theorist uses to establish

⁸ I’m not actually convinced that this is true, but I’m willing to accept it *arguendo*.

this premise have analogies that would establish an analogous denial of categorical normative reasons *for belief*.

I think this is an entirely implausible claim. One way of demonstrating this implausibility would be to go through the arguments that the moral error theorist might use to establish practical instrumentalism, which lie behind premise (4); all that would then be required is to show that these arguments have considerably less credibility if applied in the epistemic sphere; this would suffice to unravel the CGA. But this is not the strategy that I wish to pursue on this occasion. Rather, I want to take a step back from premise (4) and reflect more broadly on the role of the argument from reasons within the case for the moral error theory. In section 2 of this chapter I argued that endorsing the argument from reasons is not necessary for the moral error theorist. Now I want to show that the argument need not be conceived of as *sufficient* for establishing the error theory, either. This, it will be seen, also suffices to unravel the CGA.

Recall that the argument from reasons can be thought of as a way of understanding Mackie's famous argument from queerness, and while one might be tempted to see the argument from queerness as having the aspiration to be *sufficient* to establish the moral error theory, it is clear that Mackie doesn't see it this way. For a start, he also deploys the argument from relativity. More revealingly, though, even after both the arguments from relativity and queerness have been deployed, he evidently does not think that the error-theoretic result has yet been established. He states that the error-theoretic conclusion is so counterintuitive that it should be accepted only if a further proviso can be satisfied: that we can explain how moral beliefs, though false, have "become established and ... so resistant to criticism" (1977: 42). Why does Mackie think this?

One line of resistance to error-theoretic arguments (including the argument from queerness/reasons) claims that we should always reject an argument that has as a consequence that there is nothing morally wrong with torturing innocents (say), since our confidence that there *is* something morally wrong with torturing innocents is surely higher than our confidence in any philosophical argument in favor of error theory. (See Dworkin 1996: 117-118; Huemer 2005: 115-117; Enoch 2011: 119.) We may not even be in a position to pinpoint quite what's wrong with the argument from reasons, but given that its conclusion has such counterintuitive implications (allegedly), we may rest assured that it is unsound.⁹

It is open to the error theorist at this stage, then, to interrogate the status of the pro-moral intuitions that are playing a role in the argumentation. Where do these intuitions come from? What grounds do we have for according them *prima facie* weight? By what means would these intuitions put us in touch with the moral facts? The error theorist can now develop further skeptical arguments aimed at undermining

⁹ Kit Fine: "It may perhaps be conceded that the arguments of the skeptic appear to be utterly compelling; but the Mooreans among us will hold that the very plausibility of our ordinary beliefs is reason enough for supposing that there *must* be something wrong in the skeptic's arguments, even if we are unable to say what it is" (2001: 2).

these very intuitions: explaining them in a way that deprives them of dialectic force—that is, explaining them *away*.¹⁰ Mackie states that the conclusion to the argument from queerness should not be accepted until further arguments of this nature are provided.

Mackie attempts to satisfy this proviso by proposing a thesis of moral psychology that he calls “objectification” (but which more usually nowadays would be called “projectivism”). According to this view, our moral experience is the product of our “objectifying” subjective emotional responses, which encourages us to see the world as containing features that it might not contain. This is Hume’s view of the human mind’s “great propensity to spread itself on external objects” ([1740] 1978). Hume contrasts the faculty of taste with the faculty of reason, ascribing to the former the production of moral judgment, and claiming that taste, “gilding or staining all natural objects with the colours, borrowed from internal sentiment, raises in a manner a new creation” ([1751] 1983). If humans are psychologically predisposed to engage in moral objectification in this manner, then, the thinking goes, it is hardly surprising that the mistaken judgments of which the error theorist accuses us should be so ubiquitous and entrenched. Even passionate and stubborn resistance to the error-theoretic view is exactly what the error theory (when argued for in this manner) would predict.¹¹

Another argument that can play this role (either independently or in conjunction with the objectification thesis) emerges from reflection on the evolutionary origins of moral judgment—seeing it as an adaptation that evolved in order to reinforce cooperative tendencies rather than to track the facts of a moral realm (see Joyce 2006, 2016, 2017).¹²

What is shared by both these arguments—Mackie’s objectification thesis and the evolutionary debunking argument—is that they aim to offer a reasonable explanation of why humans have strong moral intuitions in a manner that is consistent with these intuitions being false. Both arguments also, it is worth incidentally remarking, depend ultimately on *empirical* claims. The key point I want to emphasize, though, is that the argument from reasons is not intended to stand alone; it works in concert with other arguments such as these.

In the interests of emphasizing this point still more, let us take the steps of this metaethical debate a little further. I have just portrayed the error theorist as potentially employing an evolutionary debunking argument in order to undermine any appeal to moral intuitions that her opponent might make. It is open to the opponent to reply that even though the evolutionary genealogy offered *appears* consistent with the error

¹⁰ See McPherson 2009.

¹¹ I discuss moral projectivism at some length in Joyce 2009 and with particular reference to the structure of Mackie’s error-theoretic argument in Joyce 2010. Mackie discusses projectivism more in Mackie 1980 than in Mackie 1977.

¹² Note that an argument that explains the origin of moral judgment in a manner that is consistent with the moral error theory isn’t itself an argument for the error-theoretic view. Rather, it can be framed as an argument for the view that moral judgments lack justification. However, such an argument can nevertheless be pressed into the service of the error theorist’s overall case.

theory, in fact moral facts do have a place in the genealogy. For example, if we explain moral judgment as the output of a psychological adaptation that evolved in order to reinforce cooperative tendencies, one might respond that moral goodness supervenes on certain types of cooperation, such that any explanation that includes cooperation *does* implicitly include moral properties. (For this strategy, see Enoch 2010; Wielenberg 2010; Brosnan 2011.) According to this view, even if moral judgments are the output of a faculty that evolved to reinforce cooperation, we may nevertheless consider this faculty as a generally reliable tracker of moral facts—that is, we may hold that our moral judgments are (sometimes) justified.

The would-be debunker needs to respond to this possibility of there being a “bridging principle” between the components of the naturalistic genealogy that she endorses and the moral properties whose existence she doubts. One strategy is to offer arguments aimed at undermining moral naturalism generally (see Joyce 2006: chapter 6).¹³ Another strategy is offered by Folke Tersman (2017), who points out that if the opponents of the debunking argument claim that the human faculty of moral judgment is generally reliable, despite its evolutionary origins, then they face the challenge of explaining why there is so much disagreement over moral matters. By comparison, sensory perception can be considered to be generally reliable even though there are often disagreements, because we can usually fairly easily explain those disagreements as resulting from interfering factors (e.g., poor lighting conditions). Can we do the same for moral judgments? *Some* moral disagreements are explicable in an analogous fashion—as the result of fairly obvious interfering factors (e.g., failure to reflect properly on consequences)—but it remains an object of much contention how many can be plausibly accounted for in this manner (see Levy 2003; Doris & Plakias 2008). This dispute also bottoms out in a need for more empirical information—this time concerning the nature of moral disagreement.

It should be clear that my goal is not to evaluate these arguments in any depth, but rather just to illustrate the point that the case for the moral error theory is no less complex than the case for atheism, and even something as seemingly independent as the argument from reasons need not be seen as either necessary or sufficient to establish the error-theoretic position. One thread that I have followed above sees the argument from reasons potentially requiring that moral intuitions be explained away, which can be accomplished via a genealogical debunking argument, which in turn can be supported if the nature of moral disagreement turns out to have certain characteristics. (Other threads might have been followed instead.)

How do these observations affect the CGA? Remember that we are still considering an appropriately restricted CGA, which I shall base on Cowie’s wording:

¹³ In fact this may be considered overkill, since all that the would-be debunker needs to undermine are particular versions of moral naturalism—namely, those that would relate moral facts to those naturalistic properties mentioned in the genealogical explanation. Thus (contra Das 2016) the debunking argument does not presuppose the falsity of moral naturalism generally.

1. *Parity premise*: if the argument from reasons for the moral error theory were sufficient to establish its truth, then that argument (or an appropriate analogue of it) would also be sufficient to establish the truth of the epistemic error theory.
2. *Epistemic existence premise*: the epistemic error theory is false.
3. So, the argument from reasons for the moral error theory is not sufficient to establish that the moral error theory is true.

One way of interpreting all I've been striving to establish in this section is that the conclusion of this argument, (3), is in fact true. The moral error theorist probably *doesn't* think that the argument from reasons is sufficient to establish the moral error theory—other supporting arguments are needed. Mackie, as we have seen, thinks that the conclusion of the argument from queerness should not be accepted until additional arguments can explain away the moral intuitions that make the conclusion seem so objectionable. If, then, we imagine an analogue of the argument from reasons aimed at establishing an *epistemic* error theory, then the same question arises: are there further arguments that would explain away the intuitions that render the epistemic error theory so objectionable? If the answer is “yes” in the moral case and “no” in the epistemic case, then (*ceteris paribus*) we may have grounds for accepting the moral error theory while not accepting the epistemic error theory. In my opinion this outcome is not at all unlikely.

Recall how Mackie tries to undermine moral intuitions by appeal to the objectification thesis, which I suggested might be supplemented with an evolutionary debunking argument. Are analogous arguments available in the epistemic case? Might it be that our judging that someone ought to believe that *p* (given the evidence to which she has been exposed) is a matter of our projecting our emotions onto the situation? Might it be that our tendency to judge that someone ought to believe that *p* is the output of a psychological adaptation that emerged because it provided our ancestors with some kind of non-truth-tracking reproductive advantage? There is a great deal less to be said in favor of a positive answer to these questions than to their moral counterparts.

Compare, for example, the two following claims:

- i. In evolutionary terms, the capacity for moral thinking emerged because it strengthened motivation to cooperate, not because it tracked moral truths. Even the moral error theorist can agree that this is why the capacity evolved.
- ii. In evolutionary terms, the capacity for thinking that people ought to believe certain things (in certain conditions) emerged because it allowed ϕ , not because it tracked epistemic truths. Even the epistemic error theorist can agree that this is why the capacity evolved.

The problem isn't just that it's difficult to know what “ ϕ ” might stand for; it's that the supporting arguments will proceed in completely different ways. The opponent of the moral error theory, as we saw, will usually respond to (i) by trying to establish some bridging principle between moral facts and the naturalistic components of the

genealogical account. The opponent of the epistemic error theory need not bother with anything analogous; she can respond to (ii) by pointing out that the epistemic error theory is self-defeating: if its proponents are correct, then we have no reason to believe anything they say about evolution or anything else.

We also saw earlier that the plausibility of the bridging principle offered by the moral success theorist may well depend on the nature of moral disagreement, since certain sorts of disagreement will be challenging to explain if one is also claiming that the human moral faculty evolved in such a way that it reliably tracks the moral facts. But the phenomenon of moral disagreement appears radically different from the phenomenon of epistemic disagreement. It's not clear to what extent there even *is* any phenomenon of the latter sort. We disagree about which beliefs we ought to hold, of course, but it's not clear that ordinary folk disagree about *the norms* that govern belief-formation. When we travel to foreign countries we expect to encounter unfamiliar moral norms; we do not expect to encounter unfamiliar epistemic norms. Even if there is disagreement of the latter sort, there's no particular reason to assume it bears the relevant similarities to moral disagreement.

So the prospects of establishing (i) and (ii) seem very different. This is indicative of the prospects of undermining moral intuitions being very different from the prospects of undermining epistemic intuitions. And this, in turn, is indicative of the prospects of establishing a moral error theory (using the argument from reasons) being very different from the prospects of establishing an epistemic error theory (using an analogue of the argument from reasons).

My focus up to this point has been relatively restricted: trying to show how certain specific arguments interlock in support of moral skepticism, and how analogous arguments do not interlock in an analogous way in support of epistemic skepticism. But there is plenty of room to be a great deal less restricted and show that there are *many* disanalogies between moral norms and epistemic norms that are relevant to the case for skepticism. I will offer a quick inventory of disanalogies—scattershot and incomplete—some of which are from *THE CASE FOR MORAL ERROR THEORY* sketched earlier and some of which I'm mentioning for the first time.

- An important disanalogy between the norms of morality and the norms of belief is (arguably) that the latter appear compatible with involuntarism in a manner that the former are not. And free will is puzzling, possibly nonexistent, and potentially incoherent.¹⁴
- Certain areas of moral discourse concern themselves with character traits (e.g., virtues and vices). But it is not clear that human psychology is in fact structured in the way necessary to make sense of these areas of moral concern (see Doris 2002; Joyce 2011: 171-178). This does not seem to be a problem for epistemic norms.

¹⁴ For the incoherence claim, see Strawson 1986. For the weaker (but still radical) claim that free will doesn't exist, see Pereboom 2001.

- Something that seems central to moral normativity, but not epistemic normativity, is desert. There's a strong tendency to think that certain moral transgressions warrant punishment, for example. But this relation of *desert* is mysterious and very hard to make sense of. There doesn't seem to be much in the way of an epistemic analogue.
- There seems something perplexing about the way that we resist deferring to experts about morality and resist accepting moral judgments on the basis of testimony (in the way that we do with regarding facts about, say, botany). There appears to be much less resistance to the idea of someone's being an expert about epistemic norms whose testimony we would accept.
- Independently of how a genealogical debunking argument might support the argument from queerness (as discussed earlier), such arguments cast the epistemological credentials of moral judgments into doubt, since they offer an explanation of the origin of moral thinking that appears compatible with its systematic falsehood (see Joyce 2006, 2016). Such genealogical debunking explanations for judgments concerning epistemic norms do not appear to be forthcoming.
- [A specific version of the previous] There is much to be said in favor of the hypothesis that, historically speaking, our basic moral framework has been shaped within a theistic context. (Indeed, perhaps this is the origin of all of it, contra the evolutionary thesis.) But this theistic context is utterly mistaken, potentially leaving us with "the survival of a concept outside the framework of thought that made it a really intelligible one," as G.E.M. Anscombe put it (1956: 6).¹⁵ There is no matching temptation to see epistemic normativity as the remnant of an outdated and broken metaphysical scheme.
- The prospects of offering a naturalization of epistemic normativity seem considerably brighter than the prospects for moral normativity. (Some of the reasons for this are implicit in the preceding bullet points.) There may, for example, be more to be said in favor of epistemic expressivism than there is for moral expressivism (Ingram 2017). There may be more to be said for a teleo-functional account of epistemic norms than there is for moral norms (Sullivan-Bissett 2017; Olson, this volume).¹⁶ And there are other possibilities, too (see, e.g., Kornblith 2002; Grimm 2009; Heathwood 2009; Steglich-Petersen 2011; Cowie 2014b, 2016; Glüer & Wikforss 2018).

¹⁵ Anscombe is faulting only one dominant approach to morality: a law-based Kantian conception. She thinks that a pre-Christian Aristotelian approach is more promising. However, see the second item on this bullet list.

¹⁶ For what it's worth, I'm quite sympathetic to this line.

- If epistemic norms can be naturalized but moral norms cannot be, then there are serious epistemological questions about how we come to know moral norms which will not arise for epistemic norms. (Philip Kitcher writes that “in ethics, as in mathematics, the appeal to intuition is an epistemology of desperation” (2005: 176).)
- If epistemic norms can be naturalized but moral norms cannot be, then there are serious questions about how moral facts can play an explanatory role which will not arise for epistemic facts. And (arguably) if something plays no explanatory role in our global ontology, then doubt arises as to why we should believe that it exists.

With all these disanalogies between moral normativity and epistemic normativity, it would be reasonable to assume that my goal is to destroy the parity premise of the CGA. But in fact the matter is slightly more complicated. The error theorist need not object to the parity premise as it is worded in the restricted CGA above. In fact (as I noted) the moral error theorist can accept that whole argument, conclusion and all: the argument from reasons is *not* sufficient to establish that the moral error theory is true. But this is no refutation of the error-theoretic position, and it is not even a refutation of the argument from reasons, so long as that argument is conceived, as I have maintained it should be, as a plank in the overall case for the moral error theory rather than as an independent master argument.

The advocate of the CGA may respond by reconstructing the moral skeptic’s overall case for the moral error theory: acknowledging the connections between the argument from reasons and other skeptical arguments, seeing the argument from reasons as just one consideration among others, giving due weight to other problematic aspects of moral normativity, and so forth. (This is in effect what Cowie accomplishes in his original presentation of the parity premise, with his general reference to “*the arguments*” for the moral error theory.) The advocate of the CGA can then offer a revised argument claiming that this overall case will have an analogue in the epistemic realm, leading to an unacceptable epistemic error theory.

But now the problem really will be with the parity premise, for there is little reason to assume that the overall case for the moral error theory is going to have an analogue leading to an epistemic error theory, and, I hope I’ve shown, numerous grounds for thinking otherwise.

4. Conclusion

Having been in the business of defending the moral error theory for a few years now, I have been struck by how quick opponents often are to assume that there’s really only one little argument for that position (with perhaps a few variations) and that therefore casting that argument into doubt is all it takes to refute the whole error-theoretic viewpoint. Much of the companions-in-guilt literature is, in my opinion, a case in point.

Though we professional philosophers are trained to focus on particular arguments and pare them down to their most minimal versions in order to evaluate their working structure, the idea that we are seriously striving to formulate independent valid arguments with demonstrably true premises is, more often than not, a kind of methodological pretense; we all know that it's rare that anyone is persuaded to adopt a philosophical viewpoint by an opponent coming up with such an argument. Tersman, whose argument was mentioned earlier, nicely sums up in general terms the point I wish to underline:

I mentioned in the beginning Rawls' view that the justification of a philosophical position is "a matter of ... everything fitting together into one coherent view" (Rawls 1971: 21). We often pay lip service to this idea. However, in practice it is not always treated with the respect it deserves. Instead, the arguments relevant to a position are discussed one at a time, under the implicit assumption that each of them is intended to be able to establish the position by itself. When it is found, unsurprisingly, that it fails to do so one forgets about it and moves on to the next. This is understandable and may in some cases be practical, to avoid getting stuck. Yet, the approach in question incurs a risk, namely that one fails to see the import of the combined, cumulative support the arguments provide. (Tersman 2017: 770-771)

In light of this, does the proponent of the argument from reasons seriously think that she can demonstrate the truth of the premise that there are no categorical normative reasons for actions? Probably not (which is not to say that she'd turn down such a demonstration if one were available!); rather, she attempts to *raise doubt* about the existence of such reasons, thereby rendering the existence of moral truths *implausible*. Other arguments work alongside the argument from reasons, turning this implausibility into grounds for *disbelief* (or so the error theorist thinks).

An atheist, if asked to justify her disbelief when cornered at a dinner party, say, might well reach for the argument from evil. It's a succinct and accessible little argument, and it might be appropriate for that conversation; but it's unlikely that the argument alone is really what stands behind her atheism. If she's a professional philosopher, she might even try her hand at really trying to make the argument watertight; but, again, it's unlikely that the argument alone is really what stands behind her atheism. It seems to me that the moral error theorist's attitude toward the argument from reasons should be considered in a similar light.

References

- Anscombe, G.E.M. 1958. "Modern moral philosophy." *Philosophy* 33: 1-19.
Brosnan, K. 2011. "Do the evolutionary origins of our moral beliefs undermine moral knowledge?" *Biology and Philosophy* 26: 51-64.

- Cowie, C. 2014a. "Why companions in guilt arguments won't work." *Philosophical Quarterly* 64: 407-422.
- Cowie, C. 2014b. "In defence of epistemic instrumentalism about epistemic normativity." *Synthese* 191: 4003-4017.
- Cowie, C. 2016. "Good news for moral error theorists: A master argument against companions in guilt strategies." *Australasian Journal of Philosophy* 94: 115-130.
- Cuneo, T. 2007. *The Normative Web*. Oxford: Oxford University Press.
- Das, R. 2016. "Evolutionary debunking of morality: Epistemological or metaphysical?" *Philosophical Studies* 173: 417-435.
- Doris, J. 2002. *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.
- Doris, J. & Plakias, A. 2008. "How to argue about disagreement: Evaluative diversity and moral realism." In W. Sinnott-Armstrong (ed.), *Moral Psychology, Vol 2: The Cognitive Science of Morality: Intuition and Diversity*. Cambridge, MA: MIT Press: 303-332.
- Dworkin, R. 1996. "Objectivity and truth: You'd better believe it." *Philosophy and Public Affairs* 25: 87-139.
- Enoch, D. 2010. "The epistemological challenge to metanormative realism: How best to understand it and how to cope with it." *Philosophical Studies* 148: 413-438.
- Enoch, D. 2011. *Taking Morality Seriously*. Oxford: Oxford University Press.
- Fine, K. 2001. "The question of realism." *The Philosophers' Imprint* 1: 1-30.
- Fletcher, G. 2017. "Pain for the moral error theory? A new companions-in-guilt argument." *Australasian Journal of Philosophy*:
<https://doi.org/10.1080/00048402.2017.1386693>.
- Glüer, K. & Wikforss, Å. 2018. "Reasons for belief and normativity." In D. Star (ed.), *The Oxford Handbook of Reasons and Normativity*. Oxford: Oxford University Press. ##-###.
- Grimm, S. 2009. "Epistemic normativity." In A. Millar, A. Haddock, & D. Pritchard (eds.), *Epistemic Value*. Oxford: Oxford University Press: 243-264.
- Heathwood, C. 2009. "Moral and epistemic open-question arguments." *Philosophical Books* 50: 83-98.
- Huemer, M. 2005. *Ethical Intuitionism*. NY: Palgrave Macmillan.
- Hume, D. [1740] 1978. *A Treatise of Human Nature*. Oxford: Clarendon Press.
- Hume, D. [1751] 1983. *An Enquiry Concerning the Principles of Morals*. Cambridge, MA.: Hackett.
- Hussain, N. 2004. "The return of moral fictionalism." *Philosophical Perspectives* 18: 149-187.
- Ingram, S. 2017. "Epistemology shmeπισtemology: Moral error theory and epistemic expressivism." *Inquiry*:
<https://doi-org.helicon.vuw.ac.nz/10.1080/0020174X.2017.1291362>.
- Joyce, R. 2001. *The Myth of Morality*. Cambridge: Cambridge University Press.
- Joyce, R. 2006. *The Evolution of Morality*. Cambridge, MA.: MIT Press.
- Joyce, R. 2008. "Replies." *Philosophy and Phenomenological Research* 77: 245-267.

- Joyce, R. 2009. "Is moral projectivism empirically tractable?" *Ethical Theory and Moral Practice* 12: 53-75.
- Joyce, R. 2010. "Patterns of objectification." In R. Joyce & S. Kirchin (eds.), *A World Without Values: Essays on John Mackie's Moral Error Theory* (Springer Press): 35-53.
- Joyce, R. 2011. "The accidental error theorist." In R. Shafer-Landau (ed.), *Oxford Studies in Metaethics, Vol. 6* (Oxford University Press): 153-180.
- Joyce, R. 2016. "Evolution, truth-tracking, and moral skepticism." In R. Joyce, *Essays in Moral Skepticism* (Oxford University Press): 142-158.
- Joyce, R. 2017. "Human morality: From an empirical puzzle to a metaethical puzzle." In M. Ruse & R. Richards (eds.), *Cambridge Handbook of Evolutionary Ethics* (Cambridge University Press): 101-113.
- Joyce, R. & Kirchin, S. 2010. "Introduction." In R. Joyce & S. Kirchin (eds.), *A World Without Values: Essays on John Mackie's Moral Error Theory* (Springer Press) ix-xxiv.
- Kitcher, P. 2005. "Biology and ethics." In D. Copp (ed.), *The Oxford Handbook of Ethics*. Oxford: Oxford University Press: 163-185.
- Kornblith, H. 2002. *Knowledge and Its Place in Nature*. Oxford: Oxford University Press.
- Levy, N. 2003. "Descriptive relativism: Assessing the evidence." *Journal of Value Inquiry* 37: 165-177.
- Mackie, J. 1977. *Ethics: Inventing Right and Wrong*. Harmondsmith: Penguin.
- Mackie, J. 1980. *Hume's Moral Theory*. London: Routledge.
- Mackie, J. 1982. *The Miracle of Theism*. Oxford: Clarendon Press.
- McPherson, T. 2009. "Moorean arguments and moral revisionism." *Journal of Ethics and Social Philosophy* 3: 1-24.
- Olson, J. 2014. *Moral Error Theory: History, Critique, Defence*. Oxford: Oxford University Press.
- Pereboom, D. 2001. *Living Without Free Will*. Cambridge: Cambridge University Press.
- Rawls, J. 1971. *A Theory of Justice*. Cambridge, MA.: Harvard University Press.
- Rowland, R. 2013. "Moral error theory and the argument from epistemic reasons." *Journal of Ethics & Social Philosophy* 7: 1-24.
- Shah, N. 2011. "Can reasons for belief be debunked?" In A. Reisner & A. Steglich-Petersen (eds.), *Reasons for Belief*. Cambridge: Cambridge University Press: 94-107.
- Steglich-Petersen, A. 2011. "How to be a teleologist about epistemic reasons." In A. Reisner & A. Steglich-Petersen (eds.), *Reasons for Belief*. Cambridge: Cambridge University Press: 13-33.
- Strawson, G. 1986. *Freedom and Belief*. Oxford: Clarendon Press.
- Sullivan-Bissett, E. 2017. "Biological function and epistemic normativity." *Philosophical Explorations* 20: <https://doi.org/10.1080/13869795.2017.1287296>.
- Tersman, F. 2017. "Debunking and disagreement." *Noûs* 51: 754-774.

Wielenberg, E. 2010. "On the evolutionary debunking of morality." *Ethics* 120: 441-464.