

Moral naturalism and categorical reasons

Terence Cuneo

Suppose you recognize that lying to your spouse simply to save face is wrong. Under what conditions does the wrongness of this act provide you with a reason not to lie? Proponents of the Humean theory of reasons answer that it does so only when and because you have desires of the right sorts. According to the Humean view, the fact that your lying is wrong does not, as such, provide you with a reason to act; it does so only because you also care about such things as doing what is right or the wellbeing of those who might be adversely affected by your lying, such as your spouse. If you were to lack desires such as these, however, you would also lack a reason not to lie. To which Humeans typically add: most of us do in fact have such desires and, hence, such reasons. There is a reliable contingent connection between an act's being wrong and our having a reason not to perform it.

Philosophers who identify themselves as both moral realists and naturalists tend to embrace the Humean theory of reasons.¹ These philosophers offer arguments of various sorts for their position. Philippa Foot, for example, argues that accepting Humeanism saves realists from believing the Kantian "illusion" that all failures of morality are also failures of rationality.² Mark Schroeder argues, by contrast, that since the Humean theory provides the most streamlined account of the conditions under which we have reasons to act, it should be accepted.³ Peter Railton contends, somewhat differently, that "for all its faults," the Humean theory "is the clearest notion we have of what it is for an agent to have reasons to

I am grateful to Tyler Doggett, Don Loeb, and the editors of this volume for their comments.

¹ See, for example, Boyd (1988); Brink (1989 and 1992); Railton (1986 and 1992); and Schroeder (2007). Copp (1999) also expresses openness to the view.

² Foot (2002: 167). Foot subsequently recanted her earlier commitment to the Humean view.

³ Schroeder (2007). Schroeder's version of the Humean theory, I should add, differs considerably from the versions of Humeanism defended by philosophers such as Railton and Brink.

act."⁴ None of these philosophers, it should be noted, argues that there is something specific about a naturalist approach to philosophy that recommends Humeanism. Still, they all believe that their position is better off because it embraces the Humean theory. A commitment to the Humean theory, say these philosophers, is a virtue of their view.

It is a striking fact that others who sympathize with a naturalist approach to philosophy disagree. In *The Myth of Morality* and *The Evolution of Morality*, for example, Richard Joyce contends that, far from being a virtue, a commitment to the Humean theory is a serious liability.⁵ The problem with Humean moral naturalism – or "moral naturalism" for short – according to Joyce, is that anything worth calling a moral system incorporates a commitment to there being reasons that apply to us no matter what desires we happen to have. In so far as moral naturalists fail to accommodate the existence of these so-called categorical reasons, they fail to offer us a description of anything worth calling a moral system.

I have some sympathy with this complaint. Despite this sympathy, I propose in this essay to take up the cudgels in favor of moral naturalism. In particular, I wish to defend moral naturalism from Joyce's contention that any philosopher with broadly naturalist inclinations should reject moral naturalism in favor of the error theory, according to which there are no moral facts at all. Joyce's argument merits careful consideration. In my judgment, however, moral naturalists have at their disposal a reply to Joyce's argument that is both adequate and economical. According to this reply, arguments such as Joyce's suffer from an arbitrariness problem in so far as they arbitrarily weight certain features of ordinary moral practices while discounting others. Once we get clear on the desiderata that a good moral theory should satisfy, I further claim, we can see that not only do arguments such as Joyce's fail to provide sufficient reasons for rejecting moral naturalism, but also that moral naturalism is preferable to the error theoretic position that Joyce claims that naturalists should accept.

My aim in this essay, then, is to defend moral naturalism. This might seem a curious thing to do, since I have elsewhere defended a version of "robust" moral realism that rejects the Humean theory.⁶ But my wish to defend moral naturalism on this occasion is the result of neither a change of heart nor philosophical perversity. Rather, it is rooted in the conviction that it is instructive to think about what metaethical view one would accept were one to become convinced that one's favored view is false. My

⁴ Railton (1986: 166). ⁵ See Joyce (2001 and 2006).

⁶ In Cuneo (2007: ch. 7).

own favored view, as I have indicated, is a version of robust realism that accepts the existence of categorical moral reasons. Were I to become convinced that robust realism is false, however, I would accept a version of moral naturalism. Even philosophers need contingency plans!

6.1 THE CASE FOR CATEGORICAL REASONS

Moral naturalists, we have seen, accept the claim that:

The moral wrongness of an act does not, as such, provide an agent with a reason not to do it.

At first glance, this claim might take us aback. How could the wrongness of an act fail to provide you with a reason not to do it?

Moral naturalists have an answer to this question. The answer consists in distinguishing the claim that:

(A) Morality is categorically applicable: a person can be morally obligated to do something even if doing it fails to serve any of her desires
from:

(B) Morality is categorically reason-giving: a person has a reason to perform her moral duty even if doing so fails to serve any of her desires.⁷

Moral naturalists claim that while (A) is true, (B) is false.

To help us see how this could be so, moral naturalists such as Foot draw parallels between morality and social practices such as etiquette.⁸ It is plausible to believe, Foot contends, that:

(A') Etiquette is categorically applicable: a person can be obligated by the practice of etiquette to do something even if doing it fails to serve any of her desires.

But it is not at all plausible to believe:

(B') Etiquette is categorically reason-giving: a person has a reason to conform to the requirements of etiquette even if doing so fails to serve any of her desires.

The reason why we have little reason to believe (B') is that there appear to be cases in which the requirements of etiquette apply to an agent, but she has no reason to conform to them. For example, etiquette might require that you eat at a relaxed pace, not shoveling large quantities of food into your mouth. If you are eating dinner alone in your study, however, you

might have no reason whatsoever to pay attention to these requirements. You might even enjoy, as Joyce points out, being boorish.⁹ Moral naturalists believe that morality is similar. Moral requirements, the moral naturalists maintain, apply to all of us simply in virtue of the fact that we are normal, mature human beings. We cannot escape these requirements simply because we lack desires of certain kinds. But these requirements do not, as such, generate reasons for us to act. If we lack desires of the right sort, the obligations apply but do not count in favor of us conforming to them.

In the two books mentioned earlier, Joyce acknowledges the distinction between the categorical applicability and the reason-giving force of morality. However, he contends, moral naturalists are mistaken to think that (A) is true but (B) is not. For nothing, Joyce claims, could count as a moral system if it failed to incorporate a commitment to both (A) and (B). In this regard, Joyce says, morality is deeply disanalogous to social practices such as etiquette.

Why is that? In his earlier book, *The Myth of Morality*, Joyce argues that the clue lies with our practices of holding others accountable. “Just consider,” Joyce writes, “our moral condemnation of the Nazis.” We do not let wrongdoers such as the Nazis off the moral hook — acknowledging that they had no reason to treat others with minimal respect — if we discover that they don’t care about morality or the welfare of those adversely affected by their actions. To the contrary:

The manner in which we condemn Nazis, ignoring any unusual desires or interests that they may have, is not a peripheral element of moral discourse; it presents a kind of reprehension that is central. A system of values in which there was no place for condemning Nazi actions simply would not count as a moral system.¹⁰

When spelled out more fully, Joyce’s thought is this: our practice of holding wrongdoers morally accountable appears to presuppose that wrongdoers have moral reasons not to act as they did. If there were such reasons, however, then they would have to be categorical; otherwise wrongdoers could get off the hook by sincerely denying that they care about morality or the welfare of those they harmed. But the lack of such desires is precisely the type of consideration that those of us who engage in practices of holding others accountable find morally irrelevant. To the extent, then, that our practices of holding others morally accountable actually reveal our moral commitments, we have strong reason to believe that we are committed to

⁷ I borrow the basis for this formulation from Shaffer-Landau (2009).

⁸ Foot (2002). Not all moral naturalists are enamored with the appeal to etiquette. For reservations, see Railton (1992).

⁹ Joyce (2006: 202).

¹⁰ Joyce (2001: 43).

the existence of categorical moral reasons (and that these reasons carry considerable weight). Indeed, our commitment to there being such reasons is so deeply entrenched in our moral practices, Joyce adds, that any "system of values that leaves out categorical imperatives" would lack the authority that we expect of morality, and simply "not count as a 'morality' at all."¹¹

This is a contentious claim. We are, after all, familiar with cases in which we discover that certain claims that are deeply entrenched in a given practice are false. In some cases, we give up on the practice; the claims that we have discovered to be false, we determine, are so central to it that there is not enough left of the practice to salvage. Witness the fate of various practices of soothsaying, such as those that involve consulting the entrails of animals, for example. In other cases, however, upon discovering that various claims that are entrenched in a practice are false, we revise the practice, noting that while we must fundamentally alter our understanding of the concepts it employs, the alterations are worth it. When we became convinced that relativity theory is true, for example, we did not jettison our practice of ascribing simultaneity relations between events; we simply altered our understanding of what we are doing when making these ascriptions. Joyce is of the view that our commitment to categorical reasons is so deeply entrenched in ordinary moral discourse and practice that, were we to surrender it, whatever is left would simply not count as morality.

But one might wonder why we should believe this. Why not concede that a commitment to categorical reasons appears to be central to our moral lives but also contend that, were we to discover there are none, morality would continue to go on much as it did, since we could – and perhaps must – live with the alteration? Why not believe that our moral practices are fluid in this regard, like our practices of ascribing simultaneity relations between events?

In his later book, *The Evolution of Morality*, Joyce tackles this issue at greater length.¹² In this discussion, Joyce acknowledges that, in the abstract, we probably have no way to decide whether a commitment to categorical reasons is a non-negotiable aspect of moral thought and practice. Philosophers simply have no settled procedures for determining such

¹¹ Joyce (2001: 177).

¹² Joyce (2001: 97–157) broaches the issue but not at any length and with far less diffidence than in *The Evolution of Morality*. In *The Evolution of Morality*, I should add, Joyce voices reservations about using the conceptuality of categorical reasons, preferring to talk of "moral clout" and "practical oomph." For the sake of expediency, I will formulate his view in terms of categorical reasons; nothing of importance, I believe, will be distorted.

matters. Even so, Joyce contends, there is a way to make progress on the issue before us. The first step is to identify the central roles played by our moral concepts, such as our concepts of the putative moral rights, responsibilities, and obligations that comprise our moral practices. The next step is to ask whether these concepts could play these roles if moral naturalism were true. If they could not, Joyce continues, then we have grounds for doubting that a given system of putative rights, responsibilities, and obligations counts as a moral system at all.¹³

And what exactly are the roles that moral concepts would have to play in order to be worthy of the name? The role to which Joyce devotes most attention is what we might call their *silencing function*.¹⁴ An example might best illustrate what Joyce has in mind. Suppose, once more, that I am faced with the choice whether to lie to my spouse. I know that doing so would be wrong. But I also know that lying would save me from considerable embarrassment. If I genuinely understand, however, that lying on this occasion is wrong, then this should have the effect of putting out of mind rationalizations that might tempt me to lie. It should shrink the space of deliberative possibilities so that I do not engage in any sort of cost–benefit calculation. Of course I may go ahead and lie anyway. But the price will be an awareness that I deserve to be the object of disapprobation.

Joyce's thought, then, appears to be that a constellation of putative rights, responsibilities, and obligations counts as a moral system only if our concepts of them would play a silencing function of this sort. The question to put to moral naturalism is whether it can account for the silencing function of moral concepts. Joyce thinks it is clear that moral naturalism cannot.

For consider the following two cases.¹⁵ Imagine, first, that I have become totally morally disaffected. I no longer care about complying with moral demands. Imagine, furthermore, that my relationship with my spouse has so deteriorated that I am no longer concerned about her wellbeing. I am now faced with the choice to lie to her, knowing that if I do, this will remain undetected. If I were aware that moral naturalism were true, then morality would not play a silencing function for me. To the contrary, in the case described, I would plainly see that I have no reason whatsoever to accede to morality's demands.

Imagine, now, a somewhat different case in which I am not totally morally disaffected. My commitment to morality and the wellbeing of my

¹³ See Joyce (2006: 201–202).

¹⁴ See Joyce (2006: 42).

¹⁵ See Joyce (2006: 6–4). I have modified Joyce's cases for my own purposes.

spouse is, however, weak. I am once again faced with the prospect of lying to her to save face. If I recognize that moral naturalism is true, I could either strengthen my resolve to do the right thing or surrender the desires that provide reasons for me to act morally. Either way, moral considerations do not play a silencing function for me. To the contrary, my recognition of the moral facts is simply an invitation to engage in exactly the sort of cost-benefit analysis that ordinary morality says should be ruled out.

From cases such as these, Joyce concludes that moral naturalism cannot accommodate the silencing function of moral concepts. Since he also takes the silencing function of morality to be non-negotiable, Joyce maintains that we have strong reason to believe that moral naturalism has failed to offer us an account of anything that deserves to be called a moral system. The system moral naturalism describes, says Joyce, is simply “too wimpy to be mistaken for morality.”¹⁶

I am dubious about this. Later I will explain why. For present purposes, however, I propose that we not probe the line of argument we just considered. Instead, I suggest that we take a maximally concessive approach toward the sort of case Joyce wishes to develop. We can do this by making the following four concessions.

First, let’s concede for argument’s sake that we have good reason to believe that at the heart of ordinary moral thinking lies a commitment to the reality of categorical moral reasons. Accordingly, let us assume, for present purposes, that we have decent reasons to believe that:

Necessarily, if there are moral facts, then there are categorical moral reasons.

Second, let us concede, without actually reviewing the case offered in its favor, a claim about which Joyce and many moral naturalists agree, namely:

There are no categorical reasons of any sort.¹⁷

Third, let us agree that:

While there are no categorical reasons, there are non-moral Humean reasons, such as prudential ones. These reasons apply to agents in virtue of the fact that they have desires of the appropriate sorts.

In his discussion of moral naturalism, Joyce admits that this view is “under serious pressure” from its critics.¹⁸ Still, since both moral

naturalists and Joyce accept this claim, it will do no harm to grant it for present purposes.

Finally, Joyce argues that if one is a naturalist, then one’s metaethical options are limited. According to Joyce, if one is a naturalist, then one should accept one of two views. Either one should accept moral naturalism or the error theory, according to which:

There are no moral facts (in the realist’s sense or in any other).

And:

Ordinary moral thought and discourse purport to represent moral facts but fail to do so, as there are none. In this respect, they are deeply and systematically mistaken.

At various places in *The Myth of Morality*, Joyce explains why he believes error theory to be more plausible than other forms of moral antirealism, such as expressivism.¹⁹ Let us suppose that these explanations are convincing. We can then grant, for the purposes of argument, this fourth claim:

If there are no moral facts, then the moral error theory is true.

This last claim represents yet another claim that is common ground between Joyce and the moral naturalists. For both agree that moral thought and discourse purport to represent moral reality. As such, both agree that if there were no moral facts, then ordinary moral thought and discourse would be in deep and systematic error.

It will be helpful to have a name for the position that incorporates these four claims. Let us call such a position *the mixed view*. Advocates of the mixed view hold that there are normative reasons aplenty, albeit no moral ones, for such reasons would have to be categorical, and these do not exist. At any rate, with the rudiments of the mixed view before us, we are now in a position to present the central argument that Joyce offers against moral naturalism and how moral naturalists might respond to it.

6.2 THE CASE AGAINST MORAL NATURALISM AND THE MORAL NATURALISTS’ RESPONSE

Moral naturalism of the sort in which I am interested comprises two claims. The first is that there are moral facts. As such, moral naturalism

¹⁶ Joyce (2006: 208).

¹⁷ Joyce (2001: chs. 4 and 5) presents an argument for this claim.

¹⁸ Joyce (2001: 52).

¹⁹ See Joyce (2001: ch. 1). I am working with the account of the error theory that Joyce offers in *The Myth of Morality*.

implies that there are facts such as *that lying simply to save face is wrong* and *that Hitler was wicked*. (Naturalists add that these facts are in some interesting sense natural facts. What they have in mind by this is difficult to pinpoint. But the matter need not concern us here.) The second claim is that the Humean theory of reasons is true. According to the Humean theory, recall, nothing is a reason for an agent unless and because she has desires of the appropriate types. Against moral naturalism of this variety, Joyce offers:

The Categoricity Argument

- (1) Necessarily, if there are moral facts, then there are categorical reasons.
- (2) There are no categorical reasons.
- (3) So, there are no moral facts.
- (4) If there are no moral facts, then moral naturalism is false.
- (5) So, moral naturalism is false.²⁰

We have conceded that we have decent reasons to accept premise (1). And we have granted that both Joyce and the moral naturalists accept (2) and (4), the argument's remaining two premises. Obviously, however, moral naturalists cannot accept the argument's conclusion. What, then, should a moral naturalist say in response?

As a first step, let us call to mind a proposition that I have singled out for attention several times, namely:

It is wrong to lie to one's spouse simply to save face.

To this proposition, we could add others like it, such as:

It is wrong to break a promise simply because one feels like it.

And:

It is wrong to kill another human being simply because it gives one pleasure.

Call propositions such as these the Stock Moral Truism. These propositions are truisms not simply because anything worth calling a moral system would have to include them. They deserve the name because they are also widely accepted and appear to be obviously true. For imagine someone were to reject the claim that it is wrong to kill another human being

simply because it gives one pleasure. (Such a person need not be committed to the claim that the killing is morally permissible. He might instead say that such killings, like all our activities, are simply not the sort of thing that can have moral properties.) This would be a strong prima facie reason to believe that this person lacked the concept of wrongness, suffered from some sort of serious cognitive malformation, received a deeply inadequate moral education, was not sane, was kidding, was attempting to deceive us, was deeply confused, or the like. Admittedly, such a denial would constitute only a prima facie reason to believe one or more of these things. It might be that, although this person is sincere and perfectly competent with the concepts in question, he believes that our concept of wrongness is subtly incoherent. His denial, I imagine, would still strike many of us as strange. But it would be comprehensible. And, conceivably, it might even be compelling. In that case, the Stock Moral Truism might still appear true, although we would find ourselves with reasons not to accept them.

With this in mind, let's now return to the Categoricity Argument, noting the following feature: defenders of the argument take the appearances of ordinary moral thought and practice very seriously. The reason why we should accept the argument's first premise, says Joyce, is because it appears to make the best sense of how we conduct ourselves in the moral domain, such as when we hold wrongdoers morally accountable. Indeed, we've seen that Joyce believes that this premise is constitutive of our moral practices in the sense that anything worth calling a moral system must presuppose it. Yet, it must be said, the defenders of the Categoricity Argument, such as Joyce, do not take the appearances of other dimensions of ordinary moral thought and practice with the same degree of seriousness. After all, they reject the Stock Moral Truism. They reject the Stock Moral Truism in spite of the fact that ordinary moral thought and practice also seem to commit us to them. Indeed, according to some, these truism are constitutive of our moral practices in the very same sense that premise (1) of the Categoricity Argument is supposed to be: they are non-negotiable features of our ordinary moral practices. A system that failed to incorporate these apparent truism would simply not count as a moral system at all.

Herein lies the tension I want to exploit. For suppose we agree, for argument's sake, that there are no categorical reasons. Why, at the outset of theorizing, would we take this to be a reason for rejecting the Stock Moral Truism rather than the first premise of the Categoricity Argument? After all, both the first premise of this argument and the Stock Moral Truism

²⁰ See Joyce (2001: ch. 2) and (2006: ch. 6).

appear to be truths deeply embedded in ordinary moral thought and practice. Indeed, I have just noted that they appear to be non-negotiable features of anything that could count as moral thought and practice. Accordingly, it is difficult to see why a person should argue that:

A: It is conceptually necessary that, if there are moral facts, then there are categorical reasons. But there are no categorical reasons. So, there are no moral facts. However, if there are no moral facts, then an agent cannot exhibit a moral demerit such as having acted wrongly. Suppose, now, that in ordinary conditions, an agent in fact fails to intend to honor her promise or kills another human simply because she feels like doing these things. It follows that the Stock Moral Truths are false: for in ordinary conditions, an agent fails to intend to honor her promise or kills another human being simply because she feels like it yet she does not thereby exhibit the moral demerit of having acted wrongly.

Rather than contend instead that:

B: It is conceptually necessary that, in ordinary conditions, if an agent fails to intend to honor her promise or kills another human being simply because she feels like doing these things, then she thereby exhibits the moral demerit of having acted wrongly. Suppose that, in ordinary conditions, an agent in fact fails to intend to honor her promise or kills another human simply because she feels like doing these things. It follows that she exhibits the moral demerit of having acted wrongly. But if she exhibits the moral demerit of having acted wrongly, then moral facts exist. And since she does exhibit such a demerit, moral facts exist. There are, however, no categorical reasons. So, the first premise of the Categoricality Argument is false: there are moral facts, but there are no categorical reasons.

It was G. E. Moore who used a strategy similar to this in his reply to the skeptic about the external world.²¹ In his reply to the skeptic, Moore pointed out that we have no better reason to accept the skeptic's principles rather than the claims about the external world than the skeptic rejects; the former recommend themselves no more than the latter. If so, Moore claimed, the skeptic's case can be neutralized. We would have no more reason to reject the reasons we have to believe in the existence of the external world rather than the skeptic's principles. Although moral naturalists have held much of what Moore had to say at arm's length, I suggest that in this context they call upon Moore as an ally. For the Moorean strategy helps us to see that the Categoricality Argument suffers from an arbitrariness problem. Advocates of the argument help themselves to

certain deeply entrenched components of our moral thought and practice while jettisoning others.

Or so it would seem. For it might be tempting to brush off the Moorean-style response as doing little more than highlighting that one man's ponens is another's tollens. But I think we should not let the relative familiarity of cases in which philosophers use similar premises to argue to incompatible conclusions to obscure the following fact: given the dialectical situation, proponents of the mixed view should be able to articulate principled reasons — ones that don't merely presuppose the truth of the position that they're trying to support — for why they embrace argument A rather than argument B. My claim is that the reasons aren't apparent.

To make this point, let me note that ethical theories, as Joyce himself argues, are themselves assessable along two main dimensions. In the first place, one can ask of a given theory whether it commits us to types of entities that we would rather not accept if we could get away with it because they would be highly mysterious or clutter our ontology. In the second place, one can ask of such a theory whether it does an adequate job of preserving deeply embedded features of ordinary moral thought and practice.²² Ideally, a good morally theory would want to satisfy both these desiderata, although as Joyce notes, they often tend to pull against one another.

Let us consider the first criterion of assessment. Does accepting the second argument just offered — what I've termed argument B — violate this criterion? It seems not. Were they to accept the existence of Humean moral reasons, advocates of the mixed view would not have thereby admitted the existence of a type of entity that they'd rather not accept if they could get away with it because it is highly mysterious. Nor, for that matter, would they have introduced new and unwanted complications into their view. For if moral naturalism were correct, moral reasons would simply be a species of Humean reasons. We have already seen, however, that advocates of the mixed view do not blanch in the face of Humean reasons; they have no reservations about admitting the existence of reasons of this type. If that is right, however, accepting the second argument introduces no type of entity that is new and objectionable into the picture.

Turn now to the second criterion of assessment, which concerns whether a theory does an adequate job of preserving deeply entrenched features of ordinary moral thought and practice. Does accepting argument B violate

²¹ See for example, the first two chapters of Moore 1953. Rowe (1993: 86) calls this strategy the "G. E. Moore shift."

²² Joyce (2007b: 5–6). Timmons (1999: ch. 1) makes similar claims.

this criterion? The issue in this case is more delicate. If we concede Joyce's contention that our moral practices appear to commit us to the existence of categorical moral reasons, then moral naturalists will have to grant that ordinary moral thought and practice is in error. But it is important to describe the error accurately.

In various places, Joyce writes as if the moral naturalists' claim that there are no categorical moral reasons implies that moral thought cannot play a silencing function. For example, toward the end of *The Evolution of Morality*, Joyce writes:

So the question we need to ask is whether moral discourse could carry on playing whatever role it does play if the connection between its prescriptions and the reasons people have to comply were merely a reliable contingent one. If not, then we have grounds for doubting that such a framework counts as a "moral" system at all.²³

Joyce continues, arguing that if moral naturalism were true, moral thinking could not play its fundamental roles:

Moral thinking has a function ... and deliberations in terms simply of what we want and need will not suffice. The moralization of our practical lives contributes to the satisfaction of our long-term interests and makes for more effective collective negotiation by supplying license for punishment, justification for likes and dislikes, and bonding individuals in a shared framework of decision-making. It is, I submit, precisely the purported authority and inescapability of moral prescriptions that enable them to perform these functions. Thus, a value system lacking practical clout could not so effectively play the social roles to which we put morality, and thus we could not use it as we use morality.²⁴

In passages such as these, Joyce seems to move from the moral naturalists' claim that:

All moral reasons are Humean to the further claim that:

If all moral reasons were Humean, then moral thought could not perform a silencing function.

But such an inference would be erroneous. Moral thought might tend to play a silencing function even if all moral reasons were Humean. For it might be that moral thought tends to misdescribe the nature of moral reasons, taking them to be categorical when they are really Humean.

Let me explain more fully what I have in mind. Moral naturalists believe that there are moral facts such as those reported by the Stock Moral Truisms. They also believe that there are moral reasons, although they hold that they are all Humean. Although moral naturalists believe that moral reasons are Humean, they also believe that many – and perhaps most – of us have moral reasons to conform to the requirements of morality. In so far as they accept that Joyce has accurately described ordinary moral thought and practice, however, they concede that moral thought and practice tend to incorporate a mistake. The mistake consists in taking the moral reasons that many of us have to be categorical when they are really Humean.

It is in this respect that moral naturalists (at least if they agree with Joyce's description of ordinary moral thought and practice) believe that ordinary moral thought and practice tend to incorporate an error. To soften the blow of this consequence of their views, moral naturalists might have various things to say about why it might be useful for us to misdescribe the nature of moral reasons.²⁵ Consider, for example, Mark Johnston's recent description of the correct way to learn pocket billiards. According to Johnston, the correct method involves seeing a white "ghost ball" touching your object ball at the point farthest from the intended pocket. One then shoots the cue ball directly at the ghost ball, and the object ball rolls in the pocket, if all goes well. This is the correct way to learn billiards, says Johnston, even though there is no ghost ball, and even though seeing it actually gives the wrong aiming point on the object ball.²⁶

In a similar fashion, moral naturalists might claim that thinking of some moral reasons as being categorical is not only something we find very natural, but also the best way to acquire moral reasons. Perhaps, for example, thinking of moral reasons in categorical terms helps us to care about morality and the welfare of others, thereby "setting the moral hook" in us. In this way, it is a useful device for forming and maintaining the sorts of desires that give us genuine reasons to act morally. Granted, there are important differences between the two cases just described: in the billiards case, we intentionally employ a method that we know involves false presuppositions;

²³ One thing that a moral naturalist might emphasize is that their views are typically committed to so-called externalist accounts of reference, according to which it is possible to refer to something even when we have mistaken beliefs about it. What matters, according to these views, is not that we have correct beliefs about that to which refer but that our concepts have the right sort of explanatory connection to them. See Boyd (1988).

²⁴ Johnston (2010: 17).

²⁵ Joyce (2006: 202).

²⁶ Joyce (2006: 208).

in the moral case, by contrast, most of us presumably do not. We simply find it natural (perhaps because of our moral education) to think of moral reasons as being categorical. In any event, the important point to see is that by admitting that moral reasons are Humean, moral naturalists needn't also claim that moral thought and discourse fail to play the types of function that Joyce believes count as components of a moral system.

Earlier I proposed that we engage Joyce's argument against moral naturalism in a concessive spirit, granting that we have decent reason to believe that moral thought and practice appear committed to the existence of categorical reasons. Proceeding in this fashion, we now see, comes at a price. For if Joyce is right, then moral naturalism does not unequivocally satisfy the second dimensions along which we assess moral theories. Moral naturalists must admit that moral thought tends to traffic in error of a certain kind, albeit of a fairly local sort.

Let us now look at the alternative available to naturalists, which Joyce says is the error theory of morality. Does the error theory satisfy the two criteria with which we assess moral theories? Well, it fares no worse than moral naturalism along the first dimension. It does not introduce any new and problematic entities. Things are different, however, when it comes to the second dimension. In this case, it fares considerably worse than moral naturalism. For unlike moral naturalism, the error theory does not countenance a fairly localized error about the way we tend to think of the nature of moral reasons. Rather, it claims that there simply are no moral facts. The Stock Moral Truism fails to express truths. In addition, it says that there are no moral reasons of any sort. Since, however, moral thought and discourse purport to represent such facts and reasons, they are, according to the error theory, in massive and systematic error. The error theory earns its name. If it were true, it is difficult to see how we could be more mistaken about moral matters.

Admittedly, most philosophical views come at a cost. The task of the theorist is usually to determine which costs are, on the whole, less onerous than others. But the case at hand, it seems to me, is fairly straightforward: the theoretical cost of accepting moral naturalism is considerably less than embracing the error theory. If a good moral theory saves well-entrenched commitments of ordinary moral thought and practice, it is better to accept a view according to which there are moral reasons that we tend to misdescribe rather than one according to which there are no moral facts or reasons at all.

In summary: suppose we assume that moral thought is committed to things having moral features of various kinds. If so, advocates of the mixed view such as Joyce who propound the Categoricality Argument are subject to a charge of arbitrariness, as they arbitrarily deploy some deeply embedded commitments of ordinary moral thought and practice while ignoring others. I've intimated that this accusation can, in principle, be adequately discharged. But, I've argued, in the case at hand it is not. The type of reason to which one might appeal to discharge the accusation – that is, that accepting argument A is preferable along the two dimensions of theory construction – is not apparent. In fact, the available reasons appear to point in the opposite direction, according to which we favor the acceptance of argument B rather than argument A. On the assumption that the charge stands, I conclude that we ought not to reject moral naturalism on the strength of anything like the Categoricality Argument.

At this point, I can imagine advocates of the mixed view becoming impatient. They will protest as follows: suppose we agree that moral naturalism and error theory are the naturalists' only two viable metaethical options. And suppose it is true that moral naturalism is in certain important respects less problematic than error theory since the nature of the error it countenances is considerably less severe than that admitted by the error theory. This, we can admit, is a point in favor of moral naturalism. Still, it does not follow that we should accept moral naturalism according to the two criteria of theory selection that we have employed. For in as much as moral naturalists reject the existence of categorical moral reasons, they fail to offer an account of a moral system at all. But the criteria for theory selection which we have employed is for assessing *moral* or metaethical theories – theories that describe moral systems. If so, moral naturalism may be better than error theory in some respects, but it is not a better metaethical theory, for it does not describe a moral system at all.

The cogency of this objection hangs on what exactly Joyce has in mind by a moral system. In one of the more memorable passages in *The Myth of Morality*, Joyce says this:

It is not necessary for me to claim that absolutely every piece of recognizably moral language implies the validity of categorical imperatives, only that a sufficient portion of them do, such that if we were to eliminate categorical imperatives and all that imply them from the discourse, whatever remained would no longer be recognizable as – could not play the role of – a *moral* discourse. Any system of values that leaves out categorical imperatives will lack the authority that we expect of morality, and any set of prescriptions failing to underwrite

this authority simply does not count as a “morality” at all. Moral discourse, in other words, is a house of cards, and the card at center bottom has “categorical imperative” written on it.²⁷

This passage is interesting because Joyce tends to slide between talking about what is necessary for a discourse or system of concepts to count as moral, on the one hand, and what it is for a system of values or set of prescriptions to count as moral, on the other. The same ambiguity, I believe, is present in much of what Joyce says in *The Evolution of Morality*.

Let us suppose that when Joyce speaks of a system of morality being a system of values and prescriptions he means that it is a system of putative rights, responsibilities, and obligations. If so, then Joyce’s claim is that if a system of putative rights, responsibilities, and obligations were not to include categorical reasons, then it could not count as a moral system. In principle, this might be correct. But we need an argument to believe it. Joyce’s argument, recall, is that moral discourse could not play a silencing function if the system it purports to describe were not to include categorical moral reasons. I have argued that this is not so. For all we reasonably believe, our moral thinking could carry on pretty much as it always has even if there were no categorical reasons. Of course ordinary thought about this system would be erroneous in certain respects. But it might be both difficult and undesirable to avoid the error.

Suppose, by contrast, that Joyce means by a system of morality a system of moral concepts. If so, Joyce’s contention is that a system of concepts could not count as moral if it did not include the concept of a categorical moral reason, which we assume has genuine application. Now, however, we need to distinguish. On the one hand, the system of concepts in question might be that employed by most ordinary people in their everyday moral deliberation. On the other, it might be that described by moral naturalists when theorizing about morality. If it is the former system, then we should not reject moral naturalism on the grounds that it offers a distorted picture of our folk moral concepts. The moral naturalists are not offering an account of such a system at all. And, so, it cannot be said that, if their view were true, our ordinary folk concepts would cease to play a silencing role. If it is the latter system of concepts that is in question, then there might be a point to press against the moral naturalists. For it might be said that moral naturalists have failed to capture enough of ordinary moral thought since, were moral naturalism accepted, then moral thought would fail to play a silencing role.

²⁷ Joyce (2001: 176–77).

Let us focus for a moment on this last counterfactual claim: were moral naturalism accepted, then moral thought would fail to play a silencing role. Is its truth enough to vindicate the accusation that moral naturalists have failed to provide an account of a moral system at all? I doubt it. I have no full-blown argument to establish this. But the following two considerations, I believe, should give us pause.

First, consider a case analogous to the one Joyce offers. Many have thought – Plato, Hobbes, and Kant come to mind – that moral obligations must be tied to a system of rewards and punishments, typically meted out by a divine being. The idea, presumably, is that something could count as a moral system only if conformance to the rights, responsibilities, and obligations that comprise it were ultimately in one’s best interest and lack of conformance were not. (This, incidentally, is not the same idea as silencing, which Joyce invokes. According to the system of rewards, moral considerations override all others. But they do so only because conforming to them is ultimately in one’s own self-interest. If this view were true, proper moral deliberation could involve cost–benefit calculations, which reveal that acting virtuously is in one’s ultimate interest. This, recall, is precisely the sort of reasoning that silencing is supposed to rule out.) It is a plausible conjecture that, by making these claims, these philosophers were not introducing novel elements to the ordinary moral thought of their time, but accurately describing important elements of it. If so, one could imagine someone arguing that any view that rejected the system of rewards position would fail to capture enough of ordinary moral thought since, were it accepted, then moral thought would fail to represent moral obligations as overriding in the sense that conforming to them ultimately redounds to one’s own wellbeing.

In fact, however, many today reject the system of rewards position. But we would not say of such people that, when they describe their view, they were describing something other than a moral system. For our moral practices appear to be supple enough to allow for difference of opinion as to whether moral considerations are tied to a system of rewards. If that is right, though, then the question to ask of Joyce’s position is why we should think that morality could survive the rejection of the system of rewards view, but not the rejection of categorical moral reasons. As best I can tell, we don’t have a good answer to this question. That moral considerations must play a silencing function and that anything worth calling virtue must ultimately contribute to one’s own wellbeing are both convictions deeply entrenched in ordinary moral thought. It is difficult to see why morality could survive the rejection of one conviction but not the other. Let me emphasize that, by drawing attention to this parallel, I do not

mean to suggest that just about anything could count as a moral system. There are limits to what could count as a moral system and, hence, the changes it could absorb, although it may be difficult to say in the abstract what these changes would be. Rather, the point is that we should distinguish between a description of a moral system *M* according to which *M* has various serious theoretical blemishes since it fails to capture deeply embedded features of ordinary morality and one that fails to describe a moral system at all. I am prepared to say that moral naturalism falls into the first category. But I see little reason to believe it falls into the latter, at least when the view is cast in the most charitable light.

Second, recall that the proposal we are considering is that, when elaborating their view, moral naturalists have failed to describe a moral system since were their view accepted, then moral thought would fail to play a silencing role. But one could reasonably worry that if this last counterfactual were true, it would imply that too many putatively metaethical views would not count as metaethical views at all, since they would fail to describe a system of moral concepts. Consider expressivism, for example. Expressivists believe that moral thought does not even involve predicating moral properties of things, let alone thinking that there are requirements that yield categorical moral reasons. If expressivism were accepted by the folk, then they would not think of moral considerations as generating categorical reasons which silence cost–benefit calculations. It follows that, according to the line of argument Joyce has offered, expressivism should be dismissed on the ground that it fails to offer us an account of a system of moral concepts at all.

Or consider moral fictionalism, a view that Joyce himself has defended.²⁸ Fictionalists such as Joyce propose that we should reform ordinary moral thought and discourse in such a way that those who participate in them engage in a type of pretense wherein they pretend there are categorical moral reasons. But pretending to believe and genuinely believing there are categorical moral reasons, as Joyce emphasizes, are different. They are different stances one can take toward a proposition. Now imagine a scenario under which fictionalism were widely accepted. Under this scenario, most would not believe but merely pretend to believe there are categorical moral reasons. Would pretending to believe there are moral reasons play the silencing function that Joyce believes is a non-negotiable dimension of genuinely moral thought?

It some cases it probably would. But, then again, in other cases, it probably wouldn't. For it is important to remember that, in Joyce's view, those who engage in the moral pretense know that they are doing so; they are not having the wool pulled over their eyes by an elite few who propagate something like Plato's noble lie. So, return one final time to the case with which we began: you are tempted to lie to your spouse simply to save yourself from considerable embarrassment. No one will find out if you do. Nor will doing so alter your long-term commitment to the fiction of morality. In this scenario, it is difficult to see why the thought that morality requires you not to lie – you know, after all, that you are merely pretending that there are moral reasons – would rule out your engaging in a cost–benefit analysis. To the contrary, I imagine that a moral thought of this sort would, for the clear-eyed fictionalist who is minimally committed to the moral fiction, simply be an invitation to engage in exactly the sort of cost–benefit analysis that Joyce believes is ruled out by genuine moral thought and discourse. Why is that? Well, the offense in question would not in any sense be a failure of rationality. Moreover, the clear-eyed fictionalist knows exactly why he often engages in moral pretense: it is for broadly instrumental reasons. He does so because he wishes to bolster self-control and increase social cohesion.²⁹ But, by hypothesis, lying on this occasion would threaten neither of these goods.

Earlier we saw that, in Joyce's view, parallel considerations were enough to rule out moral naturalism. Something could not be a genuinely moral thought if it invited an agent who is minimally committed to morality to engage in a cost–benefit analysis about whether to conform to putative moral demands. If this is right, however, then it would appear that moral fictionalism should also be dismissed on the ground that it fails to offer us an account of a system of moral concepts: the normative concepts that we deploy only in the fictive mood needn't play a genuinely silencing function.

Let me summarize. I have argued that there are three ways to understand the claim that, were moral naturalism true, it would not describe a moral system. Under the first interpretation, a moral system is a system of putative rights, responsibilities, and obligations. I have claimed that if these putative rights, responsibilities, and obligations were as moral naturalists say, this would give us no reason to believe that moral thought would fail to play a silencing function. Under the second interpretation, a moral system is a system of concepts, albeit it is that employed by ordinary

²⁸ See Joyce (2001: chs. 7–8). Actually, Joyce claims that the fictionalist position he defends is a version of expressivism. If so, then the considerations adduced earlier regarding expressivism apply to it.

²⁹ Compare the discussion of Joyce (2001: 71).

people when engaging in ordinary moral thought and discourse. I have contended that, if moral naturalism were true, Joyce's argument provides no reason to believe that these concepts would fail to play a silencing function. For moral naturalists do not claim that ordinary people think of moral reasons as being Humean. Under the third interpretation, a moral system is that system of concepts that ordinary people would use, were they to accept moral naturalism. When addressing this view, I have made two replies: first, we have reason to believe that our moral practices are supple. They exhibit enough flexibility to incorporate significant changes in moral thinking, even one in which we came to think of moral reasons as being Humean. Second, if it is true that moral thought must play a silencing function, then Joyce's argument would imply that a rather wide range of metaethical theories fail to describe a system of moral concepts. This is because were these views accepted, moral thought would not play a silencing function.

For a robust moral realist who believes in the reality of categorical moral reasons this last claim would be too good to be true. A single line of argument would be sufficient to dispose of moral naturalism, expressivism, moral fictionalism, moral subjectivism, and other views, since they all fail to describe a system of moral concepts. But – and this is the point I wish to press – it shouldn't be *that* easy to dismiss these rival views.

I suspect Joyce would agree. After all, Joyce seems to think that if a theory yields the result that we can at best pretend there are categorical moral reasons, then that is good enough. It does not on this account fail to count as a moral theory that describes a system of moral concepts. Should a similar courtesy be extended to moral naturalism it is difficult to see why one would withhold the favor.

CHAPTER 7

Does analytical moral naturalism rest on a mistake?

Susana Nuccetelli and Gary Seay

More than a century ago, G. E. Moore famously attempted to refute all versions of moral naturalism by offering an extended inference consisting of the open question argument followed by the charge that moral naturalism commits a "naturalistic fallacy." Although there is consensus that this extended inference fails to undermine *all* varieties of moral naturalism, the open question argument (OQA) is often vindicated as an argument against analytical moral naturalism. By contrast, the charge that analytical naturalism commits the naturalistic fallacy usually finds no takers at all. In this paper we argue that analytical naturalism of the sort recently proposed by Frank Jackson (1998, 2003) and Michael Smith (2000) does after all rest on a mistake – though perhaps not the one Moore had in mind when he made the naturalistic fallacy charge.

Analytical moral naturalism is roughly the doctrine that some moral predicates and sentences are semantically equivalent to predicates and sentences framed in non-moral terms. One attraction of analytical naturalism is that it promises to deliver a naturalistic account of the content of moral judgment that leaves no ground for objections inspired by the OQA – which argues, in brief, that no matter how much purely descriptive information is available about an action, it's still an open question whether that action is right or its end good, or whether we ought to perform it. Analytical naturalism attempts to accomplish this by resorting to conceptual analysis for moral predicates and sentences. Given this doctrine, it is at least possible that there are some such a priori or conceptual equivalences. If so, it is at least possible that the moral reduces to the natural in an a priori or analytical way. But our version of the OQA challenges this claim by showing that the reductions envisaged by these analytical naturalists are open to doubt on a priori grounds. We further

¹We wish to thank Brian McGuinness, Charles Landseman, and Lynn Tyrrell for helpful suggestions that have enabled us to improve this chapter.