3
In Defense of Moral Error Theory

Jonas Olson

1. Introduction

My aim in this essay is largely defensive. I aim to discuss some problems for moral error theory and to offer plausible solutions. A full positive defense of moral error theory would require substantial investigations of rival metaethical views, but that is beyond the scope of this essay. I will, however, try to motivate moral error theory and to clarify its commitments.

Moral error theorists typically accept two claims – one conceptual and one ontological – about moral facts. The *conceptual claim* is that moral facts are or entail facts about categorical reasons (and correspondingly that moral claims are or entail claims about categorical reasons); the *ontological* claim is that there are no categorical reasons – and consequently no moral facts – in reality. I accept this version of moral error theory and I try to unpack what it amounts to in Section 2.[1] In the course of doing so I consider two preliminary objections: that moral error theory is (probably) false because its implications are intuitively unacceptable (what I call the Moorean objection) and that the general motivation for moral error theory is self-undermining in that it rests on a hidden appeal to norms.

The above characterization seems to entail the standard formulation of moral error theory, according to which first-order moral claims are uniformly false. Critics have argued that the standard formulation is incoherent since – by the law of excluded middle – the negation of a false claim is true. Hence if 'Torture is wrong' is false, 'Torture is not wrong' is true. Contrary to what moral error theorists contend, then, moral error theory seems to carry first-order moral implications that by the theory's own lights are uniformly false. In Section 3 I suggest a formulation that is consistent with the standard formulation of moral error theory, free of first-order moral implications, and subject to no logical difficulties.

In Section 4 I consider and rebut Stephen Finlay's recent attack on moral error theory. According to Finlay the conceptual claim is false because all moral claims – and indeed all normative claims – are, or should be

understood as, relativized to some moral standard or system of ends. Moral error theorists thus attribute to ordinary speakers an error that simply isn't there. I argue that Finlay's view has some very implausible implications and that it does not avoid commitment to various forms of error theory. This becomes especially clear when we focus on fundamental moral claims.

In Section 5 I consider the worry that error theorists' rejection of categorical reasons proves too much; in particular, the worry that error theorists' qualms about categorical reasons apply equally to claims about *hypothetical reasons*, that is, claims to the effect that there is reason to take the means to one's ends. In my view error theorists such as Mackie and Joyce have failed to pay due consideration to this problem. What the challenge establishes, I submit, is that error theorists cannot just take for granted that hypothetical reasons are metaphysically unproblematic; they must offer an account of hypothetical reasons that shows that they are. I argue that the only plausible account available to error theorists is one according to which claims about hypothetical reasons reduce to non-normative claims about relations between means and ends.

2. Motivating moral error theory

Ever since John Mackie's seminal discussion, standard arguments for moral error theory are routinely lumped together under the label 'arguments from queerness' (Mackie, 1977: ch. 1). In my view some of these arguments have considerably more force than others. The most acute of Mackie's queerness worries about moral facts is not that moral facts – that is, facts to the effect that some agent morally ought to do or not to do some action; that there are moral reasons for some agent to do or not to do some action; that some action is morally permissible; that some institution, character trait, or what have you, is morally good or bad; and the like – would be intrinsically motivating in the sense of exerting a motivational pull on anyone who takes herself to be aware of them. This worry presupposes a version of *motivational internalism*.[2] But, as many critics have pointed out, motivational internalism is after all a highly controversial view (e.g. Brink, 1984; Dworkin, 1996). In other words, it is far from clear that it is part of ordinary speakers' conceptions of moral facts that they exert a motivational pull on anyone who takes oneself to be aware of them.[3]

Richard Garner and other commentators have noted that the most acute of Mackie's queerness worries is, rather, that moral facts would have to be, as Mackie said, *objectively prescriptive*. What makes moral facts queer is that they make demands from which we cannot escape (Finlay, 2008; Garner, 1990; Joyce, 2001; Robertson, 2008).[4]

Ronald Dworkin has complained that Mackie's talk about the objective prescriptivity or 'inbuilt to-be-pursuedness' of moral facts is overly metaphorical (1996: 114). I agree that Mackie's discussion is sometimes opaque,

and I will therefore try to unpack what it is that Mackie and other moral error theorists object to.[5]

As Mackie and other error theorists have noted, there is a sense in which we are all familiar with objective prescriptivity as instantiated in the real world.[6] For instance, it is a familiar fact that chess players ought not to move the rook diagonally and that there are reasons for soccer players not to play the ball to their own goalkeeper when under pressure. But these are not examples of the kind of objective prescriptivity Mackie objected to. Mackie did not deny that there are rules and standards according to which certain agents in certain situations ought or have reason to behave in certain ways (Mackie, 1977, pp. 25–7).

The kind of objective prescriptivity Mackie did object to is one that involves *categorical* reasons. To say that there are categorical reasons for some agent, A, to behave in some way, Φ, is to say that there is reason for A to Φ irrespective of whether A's Φing would promote satisfaction or realization of some of A's desires or aims, or promote fulfillment of some role A occupies, or comply with the rules of some activity A is engaged in. Suppose, for instance, that torturing animals for fun is morally wrong and that donating 20 per cent of one's income to charity is morally required. It seems commonsensical that there would then be reasons for any agent not to torture animals for fun and to donate 20 per cent of her income to charity, even if doing so would not satisfy or realize one of her desires or aims, or promote fulfillment of some role she occupies, or comply with the rules of some activity she is engaged in. In other words, moral facts entail facts about categorical reasons and moral claims entail claims about categorical reasons.

Elsewhere I have distinguished between *transcendent* and *immanent norms* (Olson, forthcoming). The former apply to agents categorically; their reason-giving force transcends particular aims, activities, or roles. Immanent norms, by contrast, are those whose reason-giving force depends on agents' engagement in certain goal-oriented or rule-governed activities or their occupation of certain roles, such as institutional or professional roles; the reason-giving force of immanent norms does not transcend goal-oriented or rule-governed activities or roles, which is why immanent norms imply merely *non-categorical* reasons.[7] Another way of putting it is to say that, while immanent norms determine correct behavior according to rules or fixed standards, it does not follow that there are *categorical reasons* to comply with these norms. For transcendent norms, it does follow that there are categorical reasons for compliance.

As mentioned above, it is a plausible conjecture that on the commonsense conception of moral norms these are examples of transcendent norms, whereas the norms of, for instance, chess, soccer, grammar, and etiquette are prime examples of immanent norms. To say that a norm is a moral norm is to say that there are reasons for any agent to comply with that norm, irrespective of her desires, ends, or roles.[8] To say that some norm is a norm

of etiquette or grammar, by contrast, is not to say that there are categorical reasons to comply with it, but rather to say that some sort of behavior would be incorrect relative to a certain standard of etiquette or relative to the rules of grammar.[9] In my terminology, *norms* are transcendent or immanent and *reasons* are categorical or non-categorical.

Error theorists do not object to the existence of immanent norms and non-categorical reasons. There is nothing metaphysically queer about the fact that there is (conclusive) non-categorical reason for chess players not to move the rook diagonally, since this is just the fact that moving the rook diagonally is incorrect according to the rules of chess; there is nothing metaphysically queer about the fact that there is (non-conclusive) reason for soccer players not to play the ball to their own goalkeeper when under pressure, since this is just the fact that such play tends to give the opposing team opportunities to score (and preventing the opposing team from scoring is one of the goals in soccer). Similarly, there is nothing metaphysically queer about the fact that there is non-categorical reason for a soldier to comply with the orders of a general, since this is just the fact that complying with the orders of those superior in military rank is part of the role of being a soldier. Note that a soldier might not *desire* to comply with the general's order, and he might have no ends that would be served by his compliance. The same goes for chess players and soccer players; they might not desire to play by the rules and they need not even desire to win. That is why I add that error theorists can recognize non-categorical reasons that depend on agents' *roles* and goal-oriented or rule-governed *activities*. Agents can occupy roles they have no desire to fulfill and engage in activities they have no desire to succeed in.

Moral norms and moral reasons, as we have seen, are a different a matter. The reason-giving force of moral norms transcends agents' desires, aims, and roles. One way of unpacking the popular view that moral facts are *non-natural* is in terms of categorical reasons. On this interpretation, what non-naturalist realists mean to capture in claiming that moral facts are non-natural is precisely that these facts are or entail categorical reasons.[10] By contrast, facts about, for example, etiquette and rules of grammar are natural since they do not entail categorical reasons.

Following others (e.g. Miller, 2003; Smith, 1994a), we can call the claim that moral facts are or entail categorical reasons (and correspondingly that moral claims are or entail claims about categorical reasons) *the conceptual claim*. Moral error theorists accept the conceptual claim, but they also accept *the ontological claim* that there are no such reasons in reality. Some naturalist realists aim to demystify moral facts by denying the conceptual claim (e.g. Brink, 1984). I shall consider and reject one such recent attempt in Section 4 below.

Other realist critics of moral error theory accept the conceptual claim but deny the error theorist's ontological claim.[11] The problem for these realists

is precisely to explain how there can be facts that *in themselves*, that is, irrespectively of the desires, aims, roles, or activities of human beings and other agents, require, or *count in favor of*, certain forms of behavior.[12] A popular realist rejoinder is to adopt a 'partners in guilt (or innocence)' strategy and claim that moral facts are not metaphysically queerer than, for example, mathematical and logical facts, or facts about set theory (cf. Scanlon, 1998: 62–4).[13] The latter kinds of facts about abstracta may be metaphysically problematic in a number of ways, but they do not display the feature that moral error theorists find especially queer about purported moral facts – they do not entail categorical reasons.

Someone might object that, for example, logical facts do entail categorical reasons for belief. An example might be that the fact that *p* and *if p then q* entail *q* entails that, if one believes *p* and *if p then q*, there is reason to believe *q* or give up at least one of the prior beliefs. The error theorist's response is that the reason here is non-categorical, since the claim that, if one believes *p* and *if p then q*, there is reason to believe *q*, or give up at least one of the prior beliefs, simply amounts to the claim that, according to the *modus ponens* rule, if one believes *p* and *if p then q*, it is correct to believe *q*, or give up at least one of the prior beliefs (cf. Olson, forthcoming). The *modus ponens* rule is an example of a rule that tells agents what there is reason to do *qua* (occupying the roles of) reasoners, or *qua* engaging in the activity of reasoning. But such rules do not give categorical reasons to comply with them. By contrast, when we make moral claims we do not merely mean to state or express correct moral rules for behavior; we mean to say that there are categorical reasons to comply with these rules.

2.1. Two initial objections: the Moorean argument and the hidden appeal to norms

At this point one might object that metaphysical doubts about transcendent norms and categorical reasons are based on pretty advanced, or at least controversial, philosophical theorizing. And are we not comparatively more certain that some actions – for example, torturing animals or children for fun – really are morally wrong than we are that reality harbors no categorical reasons and consequently no moral truths? Since it marshals commonsense against philosophical theorizing, let us call this argument *the Moorean argument* against moral error theory.[14]

But metaphysical qualms about categorical reasons are not the sole cornerstone of the case for moral error theory. Moral error theorists often give debunking explanations of why we humans tend to believe that there are moral facts (Joyce, 2001; 2006; Mackie, 1977, pp. 105–24). One important ingredient in these debunking explanations is the evolutionary advantages of moral beliefs. For instance, moral norms against stealing, harming, cheating, and so on tend to promote senses of trust and security, which facilitate cooperation, which in turn raise prospects of survival. As Mackie said, in

human evolutionary history morality serves as a 'device for counteracting limited sympathies' (1977, p. 107).[15]

Belief in transcendent norms and correlative categorical reasons is useful in other respects too: it puts pressure on individual agents and makes them less likely to succumb to temptations to maximize expected short-term egoistic or parochial benefits. In short, morality persists in the world of human thinking partly because of its socially useful coordinating and regulative functions.

In addition, there are plausible hypotheses, which are congenial to moral error theory, as to how and why belief in moral facts originates in the individual human mind. Shaun Nichols (2004) argues that belief in moral norms originates partly because of the linkage to affect. Witnessing suffering in others tends to give rise to intense distress in most human beings, and this is at least part of the explanation for why most people are strongly motivated to enforce and comply with norms against harming innocents, such as animals and children. Reactive distress causally explains beliefs to the effect that violations of norms against harming are generalizably wrong (Nichols, 2004, p. 180). This clearly echoes Hume's famous dictum that moral judgement stems from a 'productive faculty, [that] gilding or staining all natural objects with the colours, borrowed from internal sentiment, raises, in a manner, a new creation' (1998, p. 163).

These are rough sketches of attempts at debunking explanations of moral beliefs. Other writers have offered highly impressive and detailed elaborations, and I won't delve deeper into the matter here (see, e.g., Joyce, 2006; Nichols, 2004). Suffice it to say that these elaborations have enough plausibility to undermine the Moorean argument. For, once we take these debunking explanations into consideration, it is far from clear that we are more certain that some actions – such as torturing animals or children for fun – really are morally wrong than we are that there are no categorical reasons and consequently no moral truths (Joyce, 2010; Mackie, 1977, p. 42). Proponents of the Moorean argument might protest that we *are* comparatively more certain that certain actions really are morally wrong than we are about the correctness of debunking explanations of these beliefs. But proponents of debunking theories à la Mackie, Joyce, and Nichols have the upper hand here, since these theories predict that certain moral beliefs will be held with a high degree of certainty, and also explain why this is so. The explanation is simply that the regulative and coordinating functions they facilitate are of such vital importance to us.

It is fairly obvious that the argument against categorical reasons that proceeds via Mackie's queerness worry and debunking explanations of moral beliefs is based on an appeal to Occam's Razor. The gist of the argument, after all, is that error theory offers a *theoretically simpler* and hence *preferable* explanation of the phenomena to be explained (i.e. moral thought and talk) than do competing realist explanations.[16] But appeals to Occam's Razor and

considerations of theoretical simplicity seem to be appeals to *norms*. And consequently the moral error theorist's argument against the existence of some norms, such as moral norms, seems to involve a hidden appeal to other norms, which makes it smack of self-defeat (cf. Sayre-McCord, 1988, p. 277f.).

In response, the moral error theorist should concede that appeals to Occam's Razor and considerations of theoretical simplicity are indeed appeals to norms. But these are immanent rather than transcendent norms. To say that a theory T offers a theoretically simpler explanation of some phenomenon than a distinct theory T' is not to say that the comparative simplicity of T is a *categorical* reason to prefer T to T'. It is just to say that T is in one respect preferable to T' according to a standard of theory assessment commonly accepted by many philosophers, naturalists and non-naturalists alike, and commonly adopted in many natural and social sciences, to wit, that T is preferable to T' if T makes fewer problematic assumptions without loss of explanatory power. This is the case with moral error theory as compared with realism. The greater theoretical simplicity of the former as compared with the latter is therefore a non-categorical reason to prefer moral error theory to realism. Appeals to norms such as Occam's Razor are hence unproblematic from the moral error theorist's naturalist perspective.

I hope that what has been said so far makes moral error theory seem, if not a promising theory, then at least not a dead end in metaethics. That much suffices as a rationale for my defensive project in the remainder of the essay. I shall consider three challenges to moral error theory, starting with the most basic one, according to which the standard formulation of moral error theory is incoherent.

3. Formulating moral error theory

It is routinely said that, according to moral error theory, first-order moral claims are uniformly false. A first-order moral claim is a claim that entails that some agent morally ought to do or not to do some action; that there are moral reasons for some agent to do or not to do some action; that some action is morally permissible; that some institution, character trait, or what have you, is morally good or bad; and the like. But this raises the question of what to say about the truth-values of negated first-order moral claims, which leads to two worries: Is the standard formulation of moral error theory coherent?[17] Can it be maintained that moral error theory lacks first-order moral implications?

Mackie insisted that his error theory is purely a second-order view and as such logically independent of any first-order moral view (1977, pp. 15–17). But this can be doubted. According to the standard interpretation of Mackie's error theory, a first-order moral claim like 'Torture is morally wrong' is false. According to the law of excluded middle it follows that its negation, 'Torture

is not morally wrong', is true. That torture is not morally wrong would seem to imply that torture is morally permissible. More generally, then, the apparent upshot is that, contrary to Mackie's contention, moral error theory does have first-order moral implications. And rather vulgar ones at that; if moral error theory is true, any action turns out to be morally permissible!

But it seems that we can also derive an opposite conclusion. According to moral error theory, 'Torture is morally permissible' is false. According to the law of excluded middle it follows that torture is not morally permissible, which seems to entail that torture is morally impermissible. More generally, then, the apparent upshot is that any action is morally impermissible! This may not be a vulgar first-order moral implication, but it is surely absurd. It also transpires that the standard formulation of moral error theory leads to a straightforward logical contradiction, since we have derived that it is true that, for instance, torture is morally permissible (since any action is morally permissible) and that it is false that torture is morally permissible (since any action is morally impermissible). Ronald Dworkin has argued that this demonstrates the impossibility, indeed the incoherence, of being 'sceptical about value [...] all the way down' (1996, p. 91).

Walter Sinnott-Armstrong has suggested the following way out of the predicament: the scope of moral error theory is to be restricted, to the effect that only *positive* first-order moral claims are deemed uniformly false (2006, pp. 34–6). A positive first-order moral claim is defined as a claim that entails something about what some agent morally ought to do or not to do, what there are moral reasons for some agent to do or not to do, and so on and so forth; or what would be morally good or bad, or morally desirable or undesirable, and so on. It says nothing about mere permissibility.

Restricting moral error theory to positive first-order moral claims only rids moral error theory from incoherence and from the absurd implication that anything is morally impermissible. But one may object that it remains the case that a negative first-order moral claim such as 'Torture is not morally wrong' entails 'Torture is morally permissible,' since it seems to be a platitude that any action that is not morally wrong is morally permissible. In other words, moral error theory would still imply vulgar first-order moral nihilism, according to which anything is morally permissible. But Mackie's contention that his error theory is purely a second-order view and as such logically independent of any first-order moral view must be taken to include the first-order moral view that anything is morally permissible. In other words, Mackie's moral error theory holds that no first-order moral claims are true, and claims about moral permissibility are no exception.[18]

A better way out is to deny that the implications from 'not wrong' to 'permissible' and from 'not permissible' to 'wrong' are conceptual, and maintain instead that they are instances of conversational implicature. To illustrate, 'not wrong' conversationally implicates 'permissible,' because normally when we claim that something is not wrong we speak from within

a system of moral norms, or moral standards for short. According to most moral standards, any action that is not wrong according to that standard is permissible according to that standard.[19] General compliance with Gricean maxims that bid us to make our statements relevant and not overly informative (Grice, 1989, p. 26ff.) ensures that we do not normally state explicitly that we speak from within some moral standard when we claim that something is not wrong. But the implicature from 'not wrong' to 'permissible' is cancellable. The error theorist can declare that torture is not wrong and go on to signal that she is not speaking from within a moral standard. She might say something like the following: 'Torture is not wrong. But neither is it permissible. There are no moral properties and facts and consequently no action has moral status.' This would cancel the implicature from 'not wrong' to 'permissible.' (Analogous reasoning, of course, demonstrates why the error theorist's claim that torture is not morally permissible does not commit him to the view that torture is morally impermissible and hence morally wrong.) On this view, error theory has neither the vulgar implication that anything is permissible nor the absurd implication that anything is impermissible.

But one might object that the problems remain. The law of excluded middle entails that if 'Torture is wrong' is false, then 'Torture is not wrong' is true. If the latter claim is a first-order moral claim, the standard formulation of moral error theory still has first-order moral implications, that is, implications that by its own lights are false.

In response, recall that, according to our above definition, first-order moral claims are claims that entail that some agent morally ought to do or not to do some action; that some action is morally permissible; that some institution, character trait, or what have you, is morally good or bad; and so on. Now, according to the view on offer, a negated claim like 'Torture is not wrong' does not *entail* that torture is permissible; it merely conversationally implicates that it is, since the implicature from 'not wrong' to 'permissible' is cancellable. Likewise, 'Torture is not morally permissible' does not entail that torture is impermissible and hence wrong; it merely conversationally implicates that torture is impermissible and hence wrong. Thus negated atomic claims involving moral terms are not strictly speaking first-order moral claims, but some such claims conversationally implicate first-order moral claims.[20] Since claims like 'Torture is not wrong' are true, we cannot derive that their negations (such as 'Torture is wrong') are true. This saves the standard formulation of moral error theory from the threat of incoherence and from implausible first-order moral implications. I shall continue to say, then, that according to moral error theory first-order moral claims are uniformly false.[21]

Having defended moral error theory against the most basic challenge, I turn in the next section to the challenge that the theory is ill motivated, since the error it claims to identify in ordinary moral discourse is a chimera.

4. Defending the conceptual claim: the error in 'The error in the error theory'[22]

Although many philosophers accept the conceptual claim, it hasn't gone unchallenged.[23] In his recent article 'The Error in the Error Theory' (2008), Stephen Finlay argues that moral claims – and indeed all normative claims – are, or should be understood as, relativized to some (contextually implicit) end or system of ends.[24] According to this view, for a fact, F, to be a reason to Φ, relative to an end, E, is for F to explain why Φing would be conducive to E (2006, p. 8). Whether some fact is a reason is thus independent of agents' aims, desires, and roles. It appears, then, that Finlay's view does not rule out the notion of categorical reasons as characterized in Section 2 above. But Finlay adds that whether a reason *matters* to an agent does depend on the agent's attitudes, in particular her cares or concerns (2006, p. 17). There might be moral reasons for an agent to donate 20 per cent of her income to charity irrespective of her attitudes, but these reasons will matter to her just in case her donating 20 per cent of her income to charity would conduce to satisfaction of her cares and concerns.

I want to resist the distinction between a fact being a categorical reason and that fact mattering normatively.[25] I believe it is of the essence of a categorical reason to matter normatively (i.e. to count in favor of, to demand), irrespective of agents' attitudes. If F is a categorical reason for some agent, A, to Φ, then F matters normatively to A irrespective of whether A has the relevant cares or concerns, because even if she does not have them she *ought* to have them. So let me add explicitly to the characterization in Section 2 that to say that F is a categorical reason is to say that F matters normatively, irrespective of agents' desires, aims, or roles. With this addition in place, it is clear that Finlay rejects the notion of categorical reasons. As Finlay sometimes puts it, moral claims lack 'absolute authority' (2008, p. 351f.). According to Finlay, then, the error in the error theory is that it attributes to ordinary moral discourse an error that simply isn't there; ordinary moral claims are not and do not entail claims about categorical reasons, so the error theorists' conceptual claim is false.

I shall argue that the view that all moral claims are relativized to some end has some very implausible implications and that it does not avoid commitment to various forms of error theory. This becomes especially clear when we focus on fundamental moral claims.

4.1. On the disputation evidence for the conceptual claim

Finlay seeks to undermine various sources of evidence for the conceptual claim (2008, pp. 352–60). I shall comment on one such source, since this ties in with my arguments against Finlay's relativistic view to be offered in Section 4.2. We tend to pursue moral arguments even with people whom we take not to share our fundamental moral views, and we do so with the

objective of convincing them that we are right and they are wrong. This suggests that we do take moral judgements to be absolutist rather than relativistic. Following Finlay (2008, p. 355), we can call this 'disputation evidence' for the conceptual claim.

Finlay makes two points in response. First, he claims, 'most moral discourse takes place between people who share their fundamental moral values, and assume that they share these values' (p. 356). Second, Finlay claims that, to the extent that disputation between speakers who do not share fundamental moral values does occur, withholding relativizations of moral judgements is to be seen as a pragmatic device to win the opponent over. Withholding the moral standards or system of ends to which one's moral judgements are relativized 'is a rhetorical way of expressing the *expectation* (demand) that the audience subscribes to the speaker's ends or standards' (p. 357, Finlay's italics).[26]

Finlay's first point underestimates the prevalence of fundamental moral disagreement in many current societies. Even a cursory glance at public political debate in many countries will reveal fundamental moral disagreements between conservatives and feminists; socialists and neoliberals; cosmopolitans and nationalists; and so on. Moreover, fundamental moral disagreement between, for example, 'ethical vegetarians' (who believe that animal suffering is on a par morally with human suffering) and speciesists (who believe that humans are especially valuable *qua* being humans), and between 'pro-choice' and 'pro-life' activists regarding abortion, are not uncommon in everyday conversations.[27] In fact, we need not step outside the confines of academic moral philosophy to find many cases of fundamental moral disagreement between utilitarians and deontologists; Rawlsians and Nozickians; anarchists and communitarians; and so on. Finlay asks us to 'survey the moral judgements made on television or radio talk shows and news broadcasts, and try to recall the last time [we] engaged in moral discourse with someone like Charles Manson or a neo-Nazi' (2008, p. 356). But why assume that the person with whom you have a fundamental moral disagreement is such a depraved character? She might, rather, be a utilitarian, a Nozickian, a liberal, a conservative, a socialist, a nationalist, an ethical vegetarian, or a 'pro-life' activist.[28]

Finlay's second point backfires. The idea that moral judgements are partly rhetorical devices used to put pressure on people to behave in certain ways is congenial to both moral error theory and Finlay's relativist theory, but it fits better with the former. First, it fits well with the already mentioned hypothesis that part of the reason why moral thought and talk evolved is that their coordinating and regulative functions are highly useful from an evolutionary perspective (recall Mackie's view of morality as 'a device for counteracting limited sympathies'). It is a plausible conjecture that moral discourse fulfills these functions better if moral claims entail claims about categorical reasons than if they are reduced to claims about what would conduce to some end (cf. Joyce, 2006; Olson, 2010).[29]

Second, the most straightforward explanation of why moral claims have the kind of rhetorical force that demand certain behavior is that the conceptual claim is true: moral claims have rhetorical force *because* they are or entail claims about categorical reasons. Compare the following two claims:

(1) 'It is bad manners to eat peas with a spoon.'
(2) 'It is morally wrong to cheat on your tax declaration.'

In both (1) and (2) the standard or end to which the claims are supposedly relativized are withheld. But (1) and (2) differ in that (2) has a lot more rhetorical force than (1). Finlay's proffered explanation is that '[m]oral standards or ends are of pressing concern to [us], [and this explains] why we are much more serious and intransigent about our moral appraisals than we are about our appraisals of manners' (2008, p. 354). But one would expect the difference in seriousness and intransigence between moral claims and etiquette claims to be reflected in the concepts we use to make them. The conceptual claim makes good on this expectation: the fact that moral standards or ends are of especially pressing concern to us explains why moral claims entail claims about categorical reasons.

Furthermore, if moral claims and etiquette claims were of the same status, in so far as both kinds of claims reduce to claims about what would conduce to some end or accord with some standard, it is hard to see how moral claims could *maintain* their greater rhetorical force – someone who does not care about the relevant standard or end could waive (2) just as easily as someone who does not care about table manners could waive (1). The conceptual claim provides a straightforward explanation of why moral claims maintain greater rhetorical force than etiquette claims. It also explains straightforwardly why (2) cannot be waived as easily as (1).[30] This is simply because, unlike etiquette claims, moral claims entail claims about categorical reasons.

4.2. Against Finlay's relativist theory

Finlay holds that the *essential application conditions* for moral terms, that is, 'the criteria on which a [moral] concept or term is applied', are relational, even in the use of those who avowedly accept the conceptual claim: 'An action is judged to be *morally wrong* if and only if it is supposed that it frustrates certain ends or violates certain standards' (Finlay, 2008, p. 365).

Taken in one sense, Finlay's claim about essential application conditions for *moral wrongness* is entirely innocuous. Any ordinary moralizer who judges, for example, a particular action wrong will agree that that particular action violates the moral standard she endorses at the time of her utterance.[31] To cut any ice, then, Finlay's contention must be that *all* moral claims, and not just moral claims about particular actions, are relativized to standards.

It is a plain fact that we make moral judgements not only about particular actions but also about other things, including persons, institutions, societies, and *moral standards*. For instance, one might judge that some utilitarian moral standard is correct and that deontological moral standards are incorrect, or that some utilitarian moral standard is more likely to be correct than deontological moral standards. But on Finlay's relativist theory such claims become problematic.

Consider the following claim, which many utilitarians endorse:

(3) Utilitarian standard U – according to which an action is right if and only if it would bring about at least as great a balance of happiness over unhappiness as any other available alternative, and wrong otherwise – is the correct moral standard.

It should be uncontentious that (3) is a moral claim.[32] But utilitarians who utter (3) certainly don't mean to say that U is correct relative to some distinct moral standard or ends; they mean to say that U is the correct fundamental moral standard.

At this point there are two main options for relativists such as Finlay. One is to take fundamental moral claims like (3) to deviate from the general pattern of analysis in that they are not to be relativized to ends. Perhaps fundamental moral claims could be given an expressivist analysis, or perhaps they could be analyzed along the lines of error theory or fictionalism (cf. Finlay, 2009, p. 334f.). The drawback of this option is that it leads to an unhappily *disunified* metaethical theory. If expressivism, error theory, fictionalism, or some other non-relativist account gives a plausible analysis of fundamental moral claims one would expect that account to give an equally plausible analysis of non-fundamental moral claims, such as claims about the moral status of particular actions. Moreover, disunified theories are unattractive in that they invite a double load of critique. For example, a disunified theory that gives an expressivist analysis of fundamental moral claims and a relativist analysis of non-fundamental moral claims is vulnerable both to standard objections to expressivism and to relativism. These may not be conclusive criticisms, but they place a heavy burden of proof on defenders of disunified metaethical theories.

The second main option is to hold that fundamental moral claims do not deviate from the general pattern of analysis and maintain that they be relativized to themselves. An advantage of this option is that it leads to a unified metaethical theory. Finlay has recently made a suggestion along these lines (2009, p. 334).[33] The thought is that fundamental moral claims express tautologies. More specifically, any normative claim is implicitly or explicitly prefixed by an 'In order that *e*' clause, where *e* is some end. 'In order that *e*, it ought to be the case that one perform Φ' expresses the claim that, if one performs Φ, the likelihood that *e* be realized is greater than it would be if

some alternative to Φ were performed. The utilitarian fundamental moral claim that one ought not to perform actions that fail to maximize happiness is thus to be understood as the following tautological claim: 'In order that one does not perform actions that fail to maximize happiness, it ought to be the case that one not perform actions that fail to maximize happiness.' It is, of course, trivially true that if one does not perform actions that fail to maximize happiness the likelihood that one does not perform actions that fail to maximize happiness is greater than it would be if some alternative actions were performed. Let us call this suggested analysis of fundamental moral claims the 'tautology approach.'

The tautology approach has many troublesome implications. Here I shall briefly highlight four interrelated problems.[34]

(i) *No absolutely correct fundamental moral standard.* I said above that utilitarians who endorse (3) do not mean to say that U is correct relative to some *distinct* moral standard. Neither do they mean to say that U is correct relative to itself. It is trivially true that any fundamental moral standard is correct relative to itself, but utilitarians who endorse (3) mean to say something that is not trivially true, namely that U is correct in a non-relativized way, that is, that U is the *absolutely* correct fundamental moral standard.[35] But, according to the tautology approach, there is no absolutely correct fundamental moral standard. Hence the approach vindicates error theory about absolutely correct fundamental moral standards.

(ii) *No incorrect fundamental moral standard.* Ordinary speakers normally assume that it is possible to be mistaken about which fundamental moral standard is correct. They normally deem incorrect any fundamental moral standard that appears incompatible with the ones they endorse. For instance, an ethical vegetarian might believe that any fundamental moral standard that sanctions eating meat is incorrect; a 'pro-life' activist might believe that any fundamental moral standard that sanctions abortion is incorrect. But, according to the tautology approach, these beliefs are false.[36] As we saw in (i), any claim to the effect that some fundamental moral standard is correct is trivially true, so there is no such thing as an incorrect fundamental moral standard. Hence the tautology approach implies an error theory according to which claims to the effect that some fundamental moral standard is incorrect are uniformly false.

(iii) *No disagreement in asserted content.* What has been said in (ii) illustrates that speakers who apparently disagree about fundamental moral standards, such as utilitarians and deontologists or ethical vegetarians and speciesists, disagree at most 'in attitude' but not in what is asserted. This means that the common belief – that when speakers make incompatible fundamental moral claims they disagree in what they assert – is false.[37]

(iv) *No informative fundamental moral claims.* Many moral philosophers, as well as many ordinary speakers, believe that their fundamental moral claims are informative, often unobviously true, and perhaps even highly

controversial. But the tautology approach implies that these beliefs too are false.

Finlay might retort that attributing false beliefs about fundamental moral standards to ordinary speakers is not a big cost, since fundamental moral claims rarely appear in ordinary moral discourse. When they do, they function as conversation stoppers, the point of which is to demand motivation and action, rather than to convey semantic content (Finlay, 2009, p. 334).[38]

But this is unconvincing. First, as has already been indicated, it is not uncommon for ordinary speakers to appeal to fundamental moral standards in, for example, debates about ideology, vegetarianism, or abortion. It is, of course, debatable how frequently cases of fundamental moral disagreements occur. (Finlay suspects they occur a lot less frequently than I do.) But, irrespective of this empirical issue, it is clear that fundamental moral beliefs and disagreements are of crucial importance to many people. Many people take very seriously their doubts about whether the fundamental moral standard they accept is really correct. In asking such questions they do not doubt or ponder trivial truths. The tautology approach, then, implies error theory about possibly large, and definitely crucial, parts of ordinary moral discourse.

Second, and relatedly, I agree that fundamental moral claims often function to demand motivation and action, but it is implausible that they do not normally also function to convey semantic content. After all, many ordinary speakers, not just moral philosophers, are willing to engage in debates about fundamental moral standards. It is implausible that fundamental moral claims function merely as conversation stoppers in such debates. Open-minded participants typically hold their views about fundamental moral standards open to scrutiny and revision. As points (i)–(iv) have already suggested, they do not normally take them to be trivially true.[39]

Let us sum up. The tautology approach agrees with moral error theory in taking claims to the effect that some fundamental moral standard is absolutely correct to be uniformly false, and even goes beyond it in taking claims to the effect that some fundamental moral standard is incorrect also to be uniformly false. Furthermore, it attributes to most moral philosophers and users of ordinary moral discourse false beliefs about disagreement over fundamental moral standards and over the logical and epistemic status of fundamental moral claims – while they are normally taken to be informative, often unobvious, sometimes highly controversial and mutually inconsistent, they are all trivially true. Attributing all these errors to ordinary moral discourse seems more far-fetched than attributing error about moral metaphysics.

Relativists such as Finlay might, of course, attempt an alternative to the tautology approach to fundamental moral claims. But it seems that any such alternative view leads to a disunified metaethical theory. And, as suggested

above, defenders of disunified theories must accept a heavy burden of proof. Until relativists such as Finlay have elaborated a plausible analysis of fundamental moral claims, the case against the conceptual claim remains unconvincing. I conclude that Mackie's theory fits better than Finlay's with ordinary moral thought and talk.

5. Does the rejection of categorical reasons prove too much?

It was argued in Section 2 above that the most powerful argument from queerness targets categorical reasons. It is easy to see that the argument generalizes: those who accept it are committed to error theory not just about moral discourse but about any discourse that involves commitment to categorical reasons. Some critics have argued that this is an embarrassment for moral error theory. For instance, it has been argued that epistemic reasons should be – from the error theorist's perspective – equally problematic as moral reasons (Cuneo, 2007; Stratton-Lake, 2000). I discuss this issue elsewhere (Olson, forthcoming) and won't pursue it further here.[40]

It has also been argued that moral error theorists' argument against categorical reasons apply to claims about *hypothetical reasons*, that is, claims to the effect that there is reason to take the means to one's ends. This is a potential problem for many moral error theorists, who have wanted to accept hypothetical reasons. Consider Mackie:

> 'If you want *X*, do *Y*' (or 'You ought to do *Y*') will be a hypothetical imperative if it is *based on* the supposed fact that *Y* is, in the circumstances, the only (or the best) available means to *X*, that is, on a causal relation between *Y* and *X*. The reason for doing *Y lies in* its causal connection with the desired end, *X*. (1977, pp. 27–8, emphases added)

Later on, Mackie says that 'the reason for doing *Y* is *contingent upon* the desire for *X* by way of *Y*'s being a means to *X*' (p. 29, emphasis added), and later still that the desire for *X* '*creates* the reason for doing *Y*.' (p. 75, emphasis added)[41]

One might ask what exactly it means to say that hypothetical reasons are 'contingent upon' desires (Hampton, 1998). That is a fair question. And it is not answered by Mackie's claims that hypothetical reasons are 'based on' or 'created by' desires, or that they 'lie in' desires. Clearly, error theorists cannot hold that there is a transcendent norm to the effect that agents take (what they believe to be) the means to their ends, for that would mean that error theorists are committed to there being categorical reasons after all.

I said above that error theorists find it puzzling how there can be facts that count in favor of certain courses of behavior. But why would it be any the less puzzling for facts about desires, and facts about what would bring

about satisfaction of those desires, to count in favor of certain courses of behavior? In other words, if categorical reasons are metaphysically puzzling, why believe that hypothetical reasons are any the less metaphysically puzzling (Bedke, 2010)?

This is yet another fair question. In response, error theorists should deny that hypothetical reasons are properly understood in terms of the counting-in-favor-of relation. According to error theory, claims about hypothetical reasons are true only if they reduce to empirical claims about agents' desires and (actual or believed) efficient means of bringing about the satisfaction of these desires. So, for instance, the claim that there is hypothetical reason for some agent to *Y* can be true if and only if it reduces to the claim that doing *Y* will or is likely to bring about the satisfaction of some of the agent's desires.[42] Such claims are clearly dependent for their truth on agents' desires and ends. Hence hypothetical reasons, thus understood, are instances of what I called non-categorical reasons in Section 2 above. Note, however, that error theorists need not claim that all hypothetical reasons claims are reducible to empirical claims. Those that are not so reducible are false, just as categorical reasons claims are uniformly false.

It might be objected that reducing claims about hypothetical reasons to empirical claims about agents' desires and means to bringing about their satisfaction removes the normativity of claims about hypothetical reasons, since no mention is made of facts counting in favor of certain courses of behavior. That is true, but from the error theorists' perspective it is just as it should be; reducing claims about hypothetical reasons to empirical claims is the only way of saving them from being uniformly false.

A related objection is that, since claims to the effect that some action will or is likely to bring about the satisfaction of some desire are empirical and not normative, it is a violation of ordinary language to say that such claims are claims about *reasons* in any sense of the term. But this objection can be safely dismissed. 'Reason' is notoriously ambiguous and there is clearly a sense of the term that fits the proposed understanding of hypothetical reasons. For instance, we might say that there is reason for Sleepy to have an extra cup of black coffee this evening, meaning by this nothing more than that Sleepy desires to stay up late and were he to have an extra cup of black coffee he would be less likely to fall asleep early. To make it even clearer that such claims need not be normative, consider the fact that we might say that there was reason for Hitler to invade Britain during World War II, meaning by this nothing more than that Hitler wanted to win the war and had he invaded Britain he would have been more likely to do so. Thus there clearly is a usage of 'reason' in ordinary language according to which the term merely signifies connections between agents' desires and means to bringing about their satisfaction.

I conclude that moral error theory can meet the challenges considered in this paper.[43]

Notes

1. There are other ways of arriving at moral error theory, some of which are discussed in Joyce (forthcoming). They won't be considered here.
2. According to this version, when one judges that an action has a moral property one judges that it has a property that exerts motivational pull. This allows for the possibility of judging that an action has a moral property without being motivated to act (since the judgement might be mistaken, as it necessarily is, according to moral error theory), though it would be incoherent to judge that an action has a moral property and to judge simultaneously that one is not motivated to act. Thus Mackie's version of motivational internalism does not postulate a necessary connection between making a moral judgement and being motivated to act accordingly (cf. the next footnote).
3. Recently, critics have argued that Mackie made the mistake of mislocating this kind of queerness. Mackie claimed that moral *properties* and *facts* are queer because intrinsically motivating, but motivational internalism is often taken to be a view about a necessary connection between making (sincere) moral *judgements* and being motivated to act. In other words, Mackie should have located the queerness in moral *judgements* rather than in what they are about (Dreier, 2010; cf. Copp, 2010, p. 146). He should then have concluded either that there are no moral judgements or that moral judgements are not beliefs, but rather some kind of non-cognitive attitude. The former conclusion is wildly implausible, while the latter vindicates non-cognitivism. To Jamie Dreier, Mackie's 'mistake of mislocation' 'seems very strange' (2010, p. 82). But, on a plausible reading of Mackie, there is no mistake. When Mackie presents his queerness arguments (1977, pp. 38–42) he takes himself to have established already that that moral judgements are beliefs and hence that non-cognitivism is false (1977, p. 32f.). Now, Mackie obviously thought that some version of motivational internalism (see the previous footnote) is a conceptual truth, but, since he had already argued that moral judgements are beliefs and since beliefs are not necessarily motivating, where could he locate the motivational force, if not in the subject matter of moral beliefs, that is, in moral properties and facts? In other words, I take Mackie's idea to have been that our ordinary conception of a moral property is a conception of a property that 'makes [one] pursue' what one correctly judges to possess it (Mackie, 1977, p. 40). Given Mackie's dialectic, I fail to see that he mislocated queerness in the way Dreier and Copp suggest. However, as noted in the main text above, I agree with Mackie's critics that it is highly disputable whether intrinsic motivational pull is a feature of ordinary speakers' conceptions of moral facts and properties. *This* kind of queerness might well be a chimera.
4. Two clarifications: first, moral facts may be facts about moral permissibility. Such facts would not be objectively prescriptive but rather, as we might say, objectively *permissive*. Second, the fact that an agent ought morally to Φ would not necessarily entail that there are conclusive reasons for that agent to Φ. It is a common view that morality does not exhaust normativity; there are other normative reasons besides moral reasons, and the former may trump the latter. But see Tännsjö (2010) for a dissenting view.
5. One example of opacity in Mackie's discussion has already been mentioned: his failure to distinguish clearly between the claim that moral facts would be queer because intrinsically motivating and the claim that they would be queer because objectively prescriptive. Another example is the overly compressed discussion of why moral supervenience is troublesome for realists (1977, p. 40). I won't expand

on this point here; see, for instance, Sobel (2001) for a clarifying discussion of metaphysical qualms about moral supervenience.

6. Mackie (1977, pp. 25–27, 79–82); Joyce (2001, pp. 30–41). Joyce distinguishes between 'strong' and 'weak' categorical imperatives, and his version of error theory denies the existence of the former but grants the existence of the latter (p. 36). Joyce's strong categorical imperatives correspond to what I will call categorical reasons, while weak categorical imperatives correspond to what I will call non-categorical reasons. I prefer my terminology because, first, Joyce's terminology suggests, implausibly, that categoricity comes in degrees; secondly, it is not clear why weak categoricity would be less metaphysically queer.
7. To clarify: I take *norms* to be facts expressible by universally quantified sentences that imply that there are, for some class of agents in some set of circumstances, reasons to behave in a certain way, or that, for some class of agents in some set of circumstances, some form of behavior would be (in)correct or (im)permissible. *Reasons* I take to be facts that explain why some agent ought (*pro tanto*) to behave in certain ways, or why some form of behavior would be (in)correct or (im)permissible. I allow for the possibility that some norms are self-explaining – that some norm holds might itself be a reason to behave in certain ways. See also Olson (forthcoming).
8. As mentioned in footnote 2 above, some moral norms are norms of *permissibility*. Norms of moral permissibility are transcendent too. For instance, to say that homosexual activity is morally permissible is to say that one *may* engage in homosexual activity, irrespective of agents' aims, desires, or roles. Note also that some claims about moral permissibility entail claims about categorical reasons. The claim that homosexual activity is morally permissible entails the claim that there are categorical reasons not to prevent people from engaging in homosexual activity. Thanks to Christian Coons and Jussi Suikkanen for discussions here.
9. It is possible, of course, that some transcendent (e.g. moral) norms require compliance with some immanent (e.g. etiquette) norms.
10. A common charge against naturalistic realism is that it cannot account for the *normativity* of moral claims (e.g. Dancy, 2006a, pp. 132–8; Parfit, forthcoming). This is often taken to mean that moral naturalism cannot account for the fact that moral claims are or entail claims about categorical reasons. Cf. Mackie: '[Naturalism] leaves out the categorical quality of moral requirements' (1977, p. 33).
11. Among them are Nagel (1986), Scanlon (1998), and Shafer-Landau (2003; 2009). Dworkin (1996) spends a fair bit of time criticizing Mackie's claim that moral facts are queer because they would be intrinsically motivating. As I note above, this queerness worry is not particularly forceful. Dworkin is much swifter about Mackie's claim that moral facts are queer because they would be or entail non-categorical reasons. Dworkin says: 'There is nothing bizarre in the idea that a moral duty necessarily supplies a moral reason for action, however. That can be true only in virtue of what "duty" and "reason" mean' (p. 115). It is easy to see that Dworkin simply restates the conceptual claim. He does not attempt to answer the question of how there can be facts that in themselves, that is, independently of desires, aims, roles, or activities, of human beings and other agents, count in favor of certain behavior.
12. The 'counting in favor of' locution is currently the most popular way of spelling out the notion of a normative reason. See, for example, Scanlon (1998) and Parfit (2001).
13. For a thorough critique of using 'partners in guilt (innocence)' strategies in defense of normative realism, see Lillehammer (2007).

14. An early proponent of the Moorean argument against moral error theory was A.C. Ewing (1947, pp. 30–3). For recent versions, see Huemer (2005, pp. 115–17) and Shafer-Landau (2009).
15. Mackie cites Protagoras, Thomas Hobbes, David Hume, and G.J. Warnock as sources of inspiration.
16. A full positive defense of moral error theory would, of course, have to specify what is wrong with expressivist and naturalistic accounts of moral thought and talk. I cannot offer such a defense here. But I am inclined to believe that there are several aspects of ordinary moral discourse that even the most sophisticated versions of expressivism and naturalism cannot account for. This means that these views will have to be put forward as revisionary rather than descriptive metaethical theories, or they will be committed to some form of error theory. For recent critiques of expressivism along these lines, see Cuneo (2006), Bykvist & Olson (2009) and Olson (2010). For critiques of naturalism, see Timmons (1999, ch. 3), and Horgan and Timmons (2009).
17. This problem has been discussed by, for example, Pigden (2007), Sinnott-Armstrong (2006), Sobel (MS) and Tännsjö (2010).
18. Joyce (2001, pp. 6–9) suggests a version of moral error theory according to which moral claims are neither true nor false because they rest on false presuppositions (though he does not do so in the context of attempting to solve the error theorist's problem with negated moral claims). But Joyce gives no principled argument for why moral claims would be neither true nor false, rather than false. In general, I take claims that predicate non-instantiated properties of some individual or individuals to be false. For instance, a claim to the effect that some person is a witch (where being a witch involves being a woman with magical powers) is false. (Joyce, in fact, seems to admit this; 2001, p. 96.) The same, as another example, goes for a claim to the effect that acts of torture are morally wrong. This latter claim too is false, because it predicates a non-instantiated property of an action type. I assume a liberal account of properties, according to which there is a property P if there is in some natural language a predicate that purports to pick out P and P gives rise to no Russellian paradoxes. The predicate 'morally wrong' fits this description, so there is a property of moral wrongness, but error theorists maintain that it is metaphysically impossible for this property to be instantiated.
19. Some moral standards allow for moral dilemmas, in which one and the same action token is simultaneously not wrong and impermissible, or simultaneously not wrong and wrong.
20. Some, but not all. For instance, the claim that it is not the case that Dick believes torture is wrong does not conversationally implicate a first-order moral claim.
21. Pigden (2007) calls the problem of formulating moral error theory the *Doppelganger Problem*. My solution is similar to Pigden's (p. 453f.), barring some differences. Pigden does not appeal to conversational implicatures, and in the summary of his article he states that moral error theory should be formulated as the view that '*non-negative atomic moral judgements* are all false' (p. 455, Pigden's italics). On my view, however, 'non-negative' is a superfluous proviso since, as I say above, negated atomic claims involving moral terms are not strictly speaking first-order moral claims, but they may conversationally implicate first-order moral claims.
22. 'The Error in "The Error in the Error Theory"' is also the title of Joyce's forthcoming response to Finlay. I note with satisfaction that Joyce acknowledges that I 'beat him to the punch' in using this title (Joyce, forthcoming).
23. Foot (1972) delivered an early attack on the conceptual claim. Joyce (2001) responds to Foot; Finlay's 2008 article is largely a rejoinder to Joyce.

24. By an 'end' Finlay means 'a possible aim for action or object of desire' (2006, p. 8). He also makes clear that his view amounts to 'a naturalistic reduction of the relation of "counting in favour of" to a relation specifiable in only nonnormative terms' (p. 8). But, as I argue in the main text above, I believe that the 'counting in favor of' relation cannot be reduced to a naturalistic relation. To say that F counts in favor of Φing is to say not merely that F explains why Φing would be conducive to some end, but also that F matters normatively.
25. Here I side with Shafer-Landau (2009) and Parfit (forthcoming) on what it is for a fact to be a categorical reason. Unlike Shafer-Landau and Parfit, however, I do not believe that there are any categorical reasons.
26. It's a familiar fact that we sometimes withhold relativizations to standards or ends for rhetorical purposes and in cases where the relativizations are obvious to the involved parties. Finlay points out that it would be strange for a rugby captain to prefix his advices about rugby tactics with an 'in order to win the game', or 'in order to score a try' (2008, p. 353). But a crucial disanalogy is that it *would not* be strange for a moralizer to make moral claims such as 'Irrespective of your desires, aims, roles, or activities, you ought not to torture animals for fun.' By contrast, it *would* be strange for a rugby captain to express his advice about tactics by saying something like 'Irrespective of the aim to win or score, and irrespective of your role as teammate, you ought to play so and so.' Were the moralizer to prefix his claim that one ought not to torture animals for fun with an 'in order to fulfill your desires', or 'in order to fulfill a certain role or comply with the rules of certain activities', the claim would likely change its character or lose a good deal of its rhetorical force (as I argue in the main text). Were the rugby captain to prefix his advice about rugby tactics with an 'in order to win the game', or 'in order to score a try', he would at most be unnecessarily explicit. Cf. Joyce (forthcoming).
27. It is a familiar fact that seemingly fundamental moral disagreement sometimes stems from non-moral, such as empirical or theological, disagreement (Finlay, 2008: pp. 356–8). But it would be implausible and uncharitable to consider all, or even most, cases of seemingly fundamental moral disagreement as stemming from non-moral disagreement. Furthermore, people sometimes doubt or wonder whether the fundamental moral standard they accept is correct. When people ask such questions they are not merely doubting or wondering whether some courses of behavior conduce to some end. (I get back to this in Section 4.2 below.)
28. Finlay argues that it is not enough merely to locate fundamental moral disagreement between speakers. In order to count as evidence, it must also be established that speakers recognize that they are involved in fundamental moral disagreement (2008, p. 356f.). I believe it is not uncommon for people to recognize that they are involved in fundamental moral disagreements. This often happens in ideological debates, for example.
29. There is the possibility that ordinary speakers believe falsely that moral claims do entail claims about categorical reasons when in fact they reduce to claims about what would conduce to some end. In other words, there is the possibility that ordinary speakers are systematically mistaken about the meaning of moral terms. But this view seems considerably less likely to be true than the view that ordinary speakers are systematically mistaken about moral metaphysics. (See, further, Section 4.2 below.)
30. Cf. Joyce's response to C.L. Stevenson's claim that moral claims are imperatives disguised as assertions (2001, pp. 14–15).

31. Even moral particularists will agree. They will add only that the standard in question is irreducibly situation-specific.
32. Might Finlay avoid the problem by denying that (3) is a moral claim? In addition to being blatantly ad hoc, this move would allow moral conclusions to be derived from non-moral premises. For instance, it follows from (3) that, if some possible action, Φ, would bring about a greater balance of happiness over unhappiness than some distinct alternative, ψ, then ψ is wrong. The claim that ψ is wrong, and that it is wrong because it would be suboptimal in this way, seems a clear example of a moral claim. But then Finlay's theory would violate Hume's Law, in that it would entail that some moral claims – for example, the claim that ψ is wrong – are entailed by some non-moral claims – for example, (3) in conjunction with some further non-moral premises.
33. Finlay acknowledges that this analysis of fundamental moral claims is 'preliminary' and 'speculative' (2009, p. 334).
34. Finlay himself considers some of them (2009, p. 334).
35. Similarly, as Matt Bedke pointed out, those who reject (3) do not mean to deny a trivial truth. They normally mean to deny that U is the absolutely correct fundamental moral standard.
36. A speciesist moral standard, S, is of course incorrect relative to a non-speciesist moral standard, NS. But the claim that S is incorrect relative to NS is not a claim to the effect that S is an incorrect *fundamental* moral standard. To maintain that S is an incorrect fundamental moral standard, the ethical vegetarian must make the false claim that S is incorrect relative to itself.
37. The tautology approach, of course, shares this problem with expressivism. Unlike the former, however, expressivism is not committed to the implausible view that any fundamental moral claim is trivially true.
38. Finlay takes this conversational function of fundamental moral claims to be 'quite compatible with their being tautologous' (2009, p. 334). Cf. 2009, p. 334, note 41.
39. Finlay acknowledges in a footnote that it is a 'serious objection' that 'since people don't ordinarily take themselves to be asserting end-relational propositions when they utter ought-sentences, it is most unlikely that they are' (2009, p. 335, note 41). The serious objection I press above is that, since people don't ordinarily take themselves to be asserting tautologies when they make fundamental moral claims, it is most unlikely that they are. Finlay postpones a full response to these objections to a future occasion, but advertises that his response will rely on 'distinguishing sharply between what we mean by our words and what we think we mean' (2009, p. 335, note 41). This amounts to an error theory according to which ordinary speakers are systematically mistaken about what they mean by (some of) their words.
40. Let me just comment on one such line of criticism. It is sometimes suggested that error theory is self-undermining because it offers reasons to believe that there are no reasons (see, e.g., Stratton-Lake, 2000). In response, it should be borne in mind that consistent versions of error theory offer arguments to the effect that error theory is true (i.e. that there are no categorical reasons); they do not offer offer arguments to the effect that there are reasons to believe that error theory is true. See Olson (2009, 177f.; forthcoming) for further discussion.
41. Mackie thinks that once we have dispensed with categorical reasons it will be of no particular consequence whether *Y* actually is a means to *X*, or whether the agent knows or merely believes (truly or falsely) that it is: 'In each of these cases, the statement that [the agent] has a reason, and ought to [*Y*], is a thoroughly intelligible implementation of the general meanings of the terms' (p. 77).

42. Note that doing *Y* need not *cause* the satisfaction of the relevant desire. Suppose Romeo desires to embrace Juliet. Romeo's embracing Juliet does not cause the satisfaction of Romeo's desire; it is rather that Romeo's embracing Juliet *brings it about* that his desire is satisfied. Another example is that omissions sometimes bring about satisfaction of desires. But omissions do not cause anything.
43. Earlier versions of this paper were presented at a seminar at Stockholm University; at *Filosofidagarna* in Lund, June 2009; at the *RoME* congress in Boulder, Colorado, August 2009; and at a workshop on naturalism in ethics and metaphysics at Leeds University, September 2009. I thank the participants, in particular Selim Berker, Ross Cameron, Christian Coons, David Copp, Daniel Elstein, Ulrike Heuer, Jonathan Ichikawa, Gerald Lang, Daniel Nolan, Jan Österberg, Karl Pettersson, Wlodek Rabinowicz, Ted Sider, Jussi Suikkanen, Pekka Väyrynen, and Ralph Wedgwood, for helpful discussions. I am especially grateful to Matt Bedke, Stephen Finlay, Jens Johansson, Niklas Möller, and the editor of this volume, for their generous feedback.