Enough with the errors! A final reply to Finlay
Richard Joyce

Unpublished, 2012

The most disappointing aspect of Stephen Finlay's latest response to my work is that he has elected not to develop further his hypothesis that metaethical theories are geographically sensitive. This thesis, which shall henceforth be called "Finlayism," boldly extrapolates from the fact that several well known moral error theorists come from Australasia (Mackie raised in Sydney, myself in Auckland) to the conclusion that there is something in the multiculturalism of Australia and New Zealand that makes it a natural breeding ground for moral skepticism. (Now that it's mentioned, I do recall that playing with the Maori boys at school always left me feeling vaguely nihilistic.) It is a shame that we could not have seen Finlayism applied more widely, taking in other metaethical theories and other regions. It is well-known that metaethicists with similar views are often located within one thousand miles of each other (and note that this is actually *less* than the distance between Sydney and Auckland), and Finlayism offers insight into understanding this remarkable "clumping" phenomenon. Kudos should go to the *Australasian Journal of Philosophy* for having the courage to publish this controversial piece of philosophical pioneering.

The ramifications of Finlayism are significant, for if a philosopher's allegiance to a metaethical theory can be explained by reference to the surrounding sociological, economic, and cultural forces, then of course there is no need to take his or her arguments entirely seriously. "Oh, well, he *would* think that," one can say, "he's from Sydney." Finlayism thus promises to cut down dramatically on the amount of tiresome argumentation that currently clutters metaethics. Instead we should focus attention on the views of those metaethicists who are fortunate enough to live in places where such sociological factors do not sully thinking. If Finlay's insight is anything to go by, then Southern California may be one such place, for it was the move from New Zealand to Los Angeles that allowed Finlay to escape the insidious forces dragging him towards moral skepticism, and instead enabled him to see the shining truth of a metaethical theory that is not to be explained by reference to upbringing or geography (and is certainly not associated with Southern California): namely, moral relativism.

Before concluding on this upbeat note, however, there is (as Lieutenant Columbo used to say) just one thing I don't understand. It would appear that the nihilistic antipodean climate has clouded my judgment to the extent that there are still some aspects of Finlay's refutation of the error theory that leave me puzzled.

In my 2011*a* paper I complained that Finlay (in his 2008 paper) had misunderstood the error theoretic argument advocated by John Mackie (1977) and myself (2001): interpreting it as revolving around the question of absolutism when in fact it revolves around something else. In his 2011 rejoinder, Finlay first insists that he meant "absolute" in a vernacular manner, not in its metaethical usage that stands opposite relativistic theories; but then he also maintains that my argument is committed to metaethical absolutism anyway, and thus he feels free to use the metaethical sense after all. This is all rather confusing, so I'd like to try to clear up the

role of absolutism in the argument.[1]
Consider the following outline of an argument:

1. Moral concepts are absolutist (i.e., non relativistic).
2. But absolutist moral concepts are indefensible.
3. Therefore moral discourse is bankrupt.

Someone who agrees with premise 2 but dislikes the conclusion will attack premise 1. One way of doing so is to claim that although many people may think that morality is absolutist, its being so is not an essential feature: The idea of *relativistic* moral concepts is not an oxymoron.

This 1-2-3 argument, however, is not Mackie's and is not mine. It has obviously confused matters that at a certain point in *Myth of Morality* I did argue emphatically against moral relativism, but that was something of a sideshow to the main event. It has also confused matters that Mackie occasionally uses the word "absolute" during his presentation (in a manner that will be explained below). Still, it was very surprising to me that a critic should articulate my (and Mackie's) arguments entirely in terms of whether moral values are "absolutist." Rather, the "poisoned presupposition" of moral discourse, in my (and Mackie's) opinion, has more to do with the special authority of moral norms and imperatives. Though using a variety of terms, Mackie's favorites are "objective prescriptions" and "objective values." In an attempt to capture this sense of special authority, I have used the ugly phrase "non-institutional categorical imperatives." Thus, if we are to strip arguments down to their bones, then the one that Finlay should be attacking goes more like this:

A. Conceptually, morality requires non-institutional categorical imperatives.
B. But such things are indefensible.
C. Therefore, moral discourse is bankrupt.

Before going further with this debate, I should like to reiterate how narrow-minded it is to think that the moral error theory stands or falls entirely on this A-B-C argument. Finlay goes so far as to say that he considers "morality provisionally vindicated if Mackie's and Joyce's arguments are refuted" (2011: 536). I confess to feeling quite a weight of responsibility on my shoulders: With Mackie no longer with us, apparently my spruced-up PhD dissertation (which is what *Myth of Morality* is) is all that determines whether or not Hitler was really evil.

Thankfully for my shoulder muscles, the burden is illusory. The above argument could be totally unsound (and quite possibly is) and the error theory could still be persuasive. Indeed, the above argument is not what really motivates my allegiance to moral skepticism at all. (And nor was it playing with Maori kids at school.) If anything, I am drawn to the position after surveying over 2000 years of attempts to vindicate morality and forming the opinion that it is all best described with the phrase "desperate scrabbling." The apparent implication of Finlay's claim (that morality is provisionally vindicated if the A-B-C argument can be defeated) is that everything was going pretty well for morality until Mackie came along in 1977 and stuck his monkey wrench in. But while others may see it that way, I do not. The fact that millions of people have believed in morality over thousands of years and across

---

[1] The sequence goes: Mackie (1977); Joyce (2001); Finlay (2008); Joyce (2011*a*); Finlay (2011); and now this.

numerous cultures does not incline me toward thinking that the error theorist has any burden of proof to overcome. One comes by that convenient methodological principle only by assuming that people are smart when it comes to morality, and it seems to me that in this domain people are about as clever as goats.

Is Finlay's use of "absolutism" reasonable?

Pause to take in the big picture, because it's really quite remarkable. In his 2008 paper Finlay informed everyone what the argument of my book *Myth of Morality* was and then proceeded to criticize it. When I responded (in 2011*a*) that in fact this wasn't the argument of *Myth of Morality* at all, then Finlay's rejoinder is essentially to insist "Oh yes it is!" (in an indignant tone, no less) and to push on with his criticism. But his interpretation of the argument was incorrect in 2008, and it remains incorrect when he repeats it with exclamation marks in 2011.

Part of what emboldens Finlay to tell me my own argument (even after my protestations) is his ability to draw from *Myth of Morality* plenty of passages where I am indeed attacking conceptual moral relativism, and thus might be interpreted as defending the 1-2-3 argument. But his selection of quotes is disingenuous, for they all come from my Chapter 4 or later, where the argument has moved on from the central case for the error theory onto what I would consider subsidiary arguments (aimed at refuting moral rationalism). Finlay continues to intermingle quotes from Chapter 2 with sentences from Chapter 4, as if the very same argument were under discussion throughout. The term "absolute" in its metaethical sense doesn't appear in *Myth of Morality* till Chapter 4, by which time the central planks of the argument for error theory have already been laid down. (I went to quite tedious lengths in the 2011*a* article to explain the grand dialectic, so won't rehearse it again.) I happen now to think that my arguments in Chapter 4 against moral relativism are quite weak (I seem to recall thinking so at the time, too), but I don't take this fact to be devastating to the error theoretic argument.[2]

It is true that in Chapter 4 I spoke of moral absolutism as something "for which we face an error theory," but that's just metaethics-speak for saying that I don't believe in it. It should not be confused with the claim that moral absolutism plays some necessary role in the argument for the moral error theory. Finlay also writes (with an air of triumphant self-vindication): "Yes—Joyce *himself* identifies 'moral absolutism'…as a conceptual commitment of morality" (2011: 537). Quite so: *a* conceptual commitment (a claim pressed in the course of a subsidiary argument), but not *the* conceptual commitment upon which the central argument for the error theory turns.

The only quote that gives me a niggle of embarrassment is when I read myself saying that "All of the arguments of Chapter 2, to the conclusion that morality consists of categorical imperatives, point to morality's being non-relative" (2001: 95). That now doesn't seem correct to me at all, and so obviously some distinctions that seem evident to me now were yet

---

[2] My Chapter 4 argument against moral relativism seems to me now to be unnecessarily ambitious. What I am there trying to show is that moral imperatives cannot be identified with (a subset of) rational imperatives, since a relativistic construal seems correct for the latter but not for the former. In fact, all the argument requires is to identify the specific kind of relativism inherent in practical rationality (which, I argued in 2001, pertains to the opinions of an idealized version of the agent), and show that moral imperatives do not have *that* kind of relativity. As far as I my current thinking goes, showing that moral imperatives have *no* sort of relativity is not an essential premise of any argument I support.

to crystallize in my mind when I wrote those words. (Like most people reading back over something they wrote over ten years ago—an activity I generally avoid—I find myself cringing at many confusions, infelicities, and some truly horrible sentences.) Yet despite this confused remark, in *Myth of Morality* the argument for the error theory (involving the poisoned presupposition of non-institutional categorical imperatives) is sufficiently logically and spatially distinct from the later argument against moral rationalism (involving a premise insisting on moral absolutism) that that someone should conflate them surprises me.

But Finlay actually *doesn't* think the aforementioned comment is confused, for he considers that in advocating the A-B-C argument I *am* pushing for morality's being conceptually non-relativistic/non-relational.[3] He does this by assuming what he calls "a double marriage" of views:

> (i) relational views of value's nature to relativist or nonabsolutist views of its authority, and (ii) nonrelational or absolutist views of value's nature to nonrelativist or absolutist views of its authority. (2008: 350)

It turns out that this is a key pivot of disagreement between us, for I feel no inclination to accept this marriage of views (though I grant that the grounds for their divorce weren't properly recognized by me in 2001). This is more than a terminological quibble, since this double marriage is what gives Finlay license to assume that what is at issue in this debate is absolutism *in its metaethical sense*, and thus determines what arguments he thinks are pertinent. (In his 2008 paper, after assuming this double marriage he immediately embarks on a section entitled "The Evidence for Absolutism.") But I think he's wrong about this, and that he has thus misinterpreted the evidence that has been offered in favor of the error theoretic view and identified the wrong arguments with which to oppose the view.

Finlay claims that he previously assumed the double marriage "partly because Joyce and Mackie seemed to assume it too" (2011: 539)—seemingly forgetting that in an earlier footnote (2008: 351) he had admitted that Mackie might not endorse the marriage of views at all. In this footnote Finlay grants the possibility that the theses linked in the double marriage may come apart (something he acknowledges again in footnote 11 of his 2011 rejoinder), yet I confess that I cannot follow his reasoning for why this crucial possibility can be put aside. Finlay's footnote concessions that the theses can come apart make it hard to understand his declaration that he previously assumed the double marriage "with so little fuss partly because it seemed to me obvious" (2011: 540). In any case, in his 2011 rejoinder Finlay doesn't simply assume the double marriage, but argues for it. The argument he seems to use is to observe that non-institutional categorical imperatives (NICIs) contrast with both hypothetical imperatives (HIs) and institutional categorical imperatives (ICIs); and since the latter two imperatives are relativistic, then non-institutional categorical imperatives must be absolutist. This would allow him to read premise A as implying premise 1.

I hesitate to attribute this argument to Finlay, however, since its invalidity is so apparent. If HIs are relativistic in some manner R1, and ICIs are relativistic in manner R2, then all that can be concluded about NICIs is that they are not relativistic in either manner R1 or manner R2; it doesn't follow that they are absolutist. Of course, they will be absolutist *in a restricted manner*: namely, they will lack those relativistic elements of HIs and ICIs. By comparison,

---

[3] The terminology in this area—"relativist," "relational," "contextualist," "indexical," "perspectival"—has mushroomed of late, but it seems to me (and, I think, to Finlay) that what is at issue between us is sufficiently coarse-grained that, thankfully, these niceties need not be respected.

for example, if one is contrasting categorical imperatives (CIs) with hypothetical imperatives (for pedagogical purposes, say), and is pointing out that HIs depend on the addressee having certain ends, then it would be natural to say that CIs are, by contrast, "absolute." But that means absolute *with respect to whether they conduce the addressee's ends*; it doesn't mean that CIs are *absolutely* absolute. This is what Mackie is doing when he uses the word "absolute" in the passages which Finlay so gleefully quotes (2011: 536-7): Mackie is saying that CIs are, in contrast to HIs, absolute *in a certain respect* in which HIs are not—namely, in not requiring reference to the addressee's ends. There's an irony in the fact that Finlay offers these quotes as evidence that Mackie holds absolutism to be morality's poisoned presupposition, for Finlay agrees that moral imperatives are CIs and thus even he accepts that they are absolute in this sense.

That NICIs (and ipso facto CIs) are not necessarily (absolutely) absolute is evident if one reflects on how they can relativistic. Consider a crude form of relativism according to which when people say that something is morally right or wrong they really mean "...relative to my culture."[4] Suppose the mountain people favor bravado and martial virtues, while the valley people favor patience and peacefulness. Both peoples know this of the other and accept it as the proper state of things. Asked why things are this way, they all give a (crazy) realist answer: "It is the nature of the universe that the mountain people and the valley people should have these different moral dispositions. It's a metaphysical yin and yang thing." Moreover, they accept that different obligations attend these different natures: A mountain person has obligations that a valley person does not, and vice versa. These obligations are taken to be non-hypothetical and non-institutional. Whatever realist imagery you care to muster—about must-be-doneness, or the normative furniture of the universe, or whatever—these people imbue their norms with. They see the obligations as utterly authoritative, inescapable, reason-bringing, and binding. *Yet they also see them as relativistic*. Ask any one of them whether Fred has an obligation to cultivate stoic passivity and he or she will answer "It depends whether Fred is from the mountains or the valley." Ask a mountain person whether such passivity is a good thing and he or she will answer "No, not from my point of view." If a mountain person says "Daring is good" then he has said something true; if a valley person utters the same sentence then she says something false. There's a sense in which the mountain person can still claim that daring is good *period*. He means that *for him* there is an inescapable requirement to cultivate that virtue and there is no further discussion to be had. "But that doesn't further my ends" cuts no ice.

So Finlay's argument that (according to my own lights) NICIs must be absolutist fails. NICIs can be relativistic; they just cannot have the relativity inherent in HIs or ICIs. But in fact the argument doubly fails, since (in my opinion) HIs and ICIs do not even count as relativistic. Let us consider them in turn.

A hypothetical imperative is an "ought"-claim that applies to a person in virtue of that person's ends. "You ought to go to bed soon" is an HI if its utterance includes reference, either tacitly or explicitly, to one of the agent's ends—say, "...if you want to get a good night's sleep." Clearly, HIs involve a *relation*—between the action and the addressee's ends—but they are not thereby relativistic or contextualistic claims, for there is no parameter involved (either as part of the content of the sentence or against which the truth value of the sentence must be assessed) whose value shifts with frame of reference. In the useful terminology provided by C.L. Stevenson, HIs are "relation-designating" but not "relativistic"

---

[4] Call it "contextualism" if you wish.

(1963: 74).[5]

According to Finlay's positive view (to be outlined shortly), even categorical imperatives have the form of instrumentally relating an action to an end (just not the addressee's ends). When discussing CIs in general, Finlay leaves open the kinds of ends to which (he thinks) the prescribed action contributes (2009: 330ff); but in discussing *institutional* CIs, he suggests that the end in question is the body of norms that constitutes the institution: "S ought-relative-to-I to φ" (2011: 540)—or, at least, he attributes this view to Mackie and me. But this isn't my view. Consider an ICI: say, a road sign reading "No U-turn." The imperative is certainly not hypothetical, for it is applied to agents regardless of their interests. Even if someone had to do a U-turn here in order to save a life, and we accept that the person all-things-considered did the correct thing, the very fact that we can still describe the action as the acceptable "breaking of a rule" shows that the imperative continued to apply. The sign "No U-turn" can be reasonably rendered as the imperative "Do not perform a U-turn here," but in no sense hides the tacit structure: "Relative to the local road rules, do not perform a U-turn here." The local road rules issue or include the imperative, but are not part of its content. The fact that another set of road rules (another institution) may make contrary prescriptions doesn't affect this conclusion. In Britain the ICI is "Drive on the left"; in France it is "Drive on the right." One might then say "Relative to one set of rules we should drive on the left, but relative to another we should drive on the right." That's fine, but it doesn't mean that the British ICI contains a hidden indexical or relational element. Thus, I quite reject Finlay's interpretation of ICIs as "S ought-relative-to-I to φ."[6]

Finlay writes that "if my defence of moral indexicality is correct, then Joyce's and Mackie's error theories fail." If by this he means that showing moral imperatives to have *any* kind of indexicality (or relation-designating content) would refute my argument, then this, as I have shown, is indeed a misidentification of the poisoned presupposition. Contrary to Finlay's interpretation, I can happily allow that moral imperatives may often (and even conceivably always) have the form of relating an action to an end endorsed by the speaker, in exactly the way Finlay describes. But this doesn't address the basic skeptical challenge: *Can such a relational property satisfy the desideratum of moral authority?* If moral imperatives are of an action-end format (as Finlay thinks), then the challenge will focus on whether the end itself commands the requisite authority (which it then imparts to the imperative regarding the action). The fact that Finlay sees his goal in defeating the error theorist to be the demonstration that moral claims could be relational—not realizing that *even if successful this would leave my basic error theoretic worry unaddressed*—shows the extent to which we are talking past each other. Since I have tried to express this authority in terms of moral imperatives being non-institutional categorical imperatives—as in premise A above—then an

---

[5] Well, I used to think it was useful, but the introduction of new terminology has complicated matters. What Stevenson called "relativism" many would nowadays call "contextualism" *in contrast to* something else they call "relativism." An HI is relativistic in neither the old nor the new sense: There is no shifting parameter involved in "You ought to go to bed soon if you want to get a good night's sleep," any more than there is in "If you drop the vase it will likely shatter" (not counting the pronouns, of course).

[6] Suppose there are different normative frameworks giving contradictory advice about what an agent should do in a given situation. If Mary has to do a U-turn to save a life, then one might sensibly say that according to morality she ought to φ, but according to the local road rules she ought not φ. (Perhaps, in addition, according to etiquette she ought not, but according to prudence she ought, and so forth.) In such a situation, one might ask what Mary all things considered ought to do, and one item figuring in our deliberations might be rendered as "Mary ought-relative-to-I to φ." But it doesn't follow that this latter "ought" (with the hyphens) reflects the structure of the categorical "ought" that is issued by institution I.

argument against this premise would have to show that moral imperatives are either hypothetical imperatives or institutional categorical imperatives. Since Finlay agrees that morality is categorical, it is the non-institutionality that he must attack.

How plausible is Finlay's positive metaethical theory?

Let us remind ourselves in general terms of what is under dispute here. Mackie and I are worried that moral imperatives and values seem to be imbued with a kind of special practical authority of which we can make no sense. I see philosophers *try* to make sense of it, but in my opinion their efforts fall flat. Part of this authority can be captured by noting that morality is categorical rather than hypothetical, but that doesn't seem enough. The rules and values of morality seem to be treated as if they are not just human conventions or constructs forged of human attitudes and practices. The idea that if we all changed our minds then coins would be worthless seems acceptable in a way that the idea that if we all changed our minds then infants would be worthless does not. Calling moral imperatives "non-institutional" is an attempt to capture this aspect of morality.[7] And it seems a very important aspect of morality, so much so that there is room to doubt that the discourse and all of the practices it supports could well survive its extirpation. It is the importance of this special practical authority in how morality is *used* that leads me to sympathize with Mackie's suggestion that it "has been incorporated in the basic, conventional meanings of moral terms" (1977: 35).

For my money, whether anything other than a NICI can do the work of morality turns in large part on whether it can accommodate this special kind of practical authority. Finlay undertakes to show that something can: namely, ICIs, which, as we have seen, he treats as necessarily relativistic/relational (incorrectly, in my opinion). Indeed, he has a novel story to tell about how relativistic language can generate a sense of authority, via speakers leaving a relativistic element tacit, thus creating an expectation that the audience must accept this element.

Finlay's view is as follows: When we morally denounce a criminal, we relativize his actions to some end, such as *promoting general human well-being*. This may not be *the criminal's* end (unlike with a hypothetical imperative); it is an end the speaker endorses. (Since Finlay is attempting to show that ICIs can do the work, I assume that this privileging of the connection between action and end in some manner reflects a normative institution.) By omitting mention of this end, the speaker succeeds (via rhetoric) in *condemning* the criminal in a way that would have been impossible had the speaker stated the whole proposition explicitly (Finlay 2011: 543). As far as I can see (and I have to use these qualifications since Finlay is by his own admission "coy about the particular details" of his theory, yet gets terribly huffy when misunderstood), despite the clarifications now offered, the theory still holds that the authority of morality comes from speakers *leaving out* relativistic components, implying that the crucial sense of authority would be lost if the components were explicitly stated.

I'll say this much for Finlay's theory: It is trying to fix the right problem. The perennial general problem with interpretations of moral imperatives as relative to some end stems from a Humean thesis: that taking an interest in that end (whatever it may be) seems rationally

---

[7] Given that this "special practical authority" seems so mysterious to me, I find it rather hard to give a crisp articulation of it. In a sense, it's like asking an atheist to explain what is meant by *sacred* or *the Holy Ghost*. The atheist might reasonably object "*I* don't know!—but I've heard enough from your vague remarks to have grave doubts about its existence."

optional (but taking an interest in moral imperatives does not seem rationally optional).[8] So Finlay labors to show how a relativistic language can impart a sense of requirement to the *end*. (See also Finlay 2009: 331ff.) His account is certainly intriguing, but I nevertheless continue to find it implausible as an attempt to capture the sense of authority with which morality is imbued, or even as an explanation of condemnation. While it may be true that a speaker who leaves a component tacit creates an air that all parties to the conversation are on the same page with respect to that component, how that translates into a rhetorical *demand* escapes me. If John is 6 foot tall, but I hear a bunch of people describing him as "short," then I might be confused. When it is discovered that they were tacitly relativizing the conversation to *professional basketball players*, whereas I was relativizing it to *adult men*, then I'm sure we'll all have a good chortle. But their assumption that I was using the same contrast class as them when describing John's height hardly creates a *demand* on their part that I adopt that frame of reference. Even if there is some pressure on me to conform to their contrast class in order to participate sensibly in the conversation, this is just with respect to this conversation; there is no demand to privilege this class in other contexts or to prefer it generally.

This problem arises not only regarding conversations involving parties who endorse divergent ends. If everyone in the conversation *is* on the same page with respect to the relevant end, then clearly that end can be left tacit. But even here it is doubtful that this succeeds in capturing the sense that endorsing the end is *required*—a sense that is, arguably, characteristic of moral discourse. If we all know that we're relativizing our discussion of John's height to *professional basketball players*, then we needn't mention it; but it doesn't follow that we thereby think that we *must* choose that comparison class. My complaint here focuses on a tiny but crucial moment in Finlay's presentation of his thesis: "...it is a rhetorical way of expressing the expectation (demand) that the audience subscribes..." (2008: 357); "...the speaker expects (i.e., requires) his audience to care about that end..." (2009: 334). What bothers me are the little parenthetical implications, for expectations are not demands or requirements. This may seem like nit-picking till it is recalled that accounting for the sense of moral *demands* and *requirements* is the whole name of the game.

I think, in fact, that Finlay focuses rather too closely on *conversations* in his discussions of this point. In his 2008 paper, he tries to downplay the significance of a theory's need to deal with moral condemnation of people with utterly alien moral outlooks, on the grounds that such disagreement is much rarer than one might assume: "Try to recall the last time you engaged in moral discourse with someone like Charles Manson or a neo-Nazi" (356). But I think this misses the point. Encounters with alien moral outlooks are ubiquitous, even if conversations with actual proponents of such outlooks are rare. Movies and novels are full of nihilistic or sociopathic baddies; the evil step-parent is a stock character of our children's fairy tales; and even small homogenous societies have their myths and religions that are full of destructive characters standing outside the accepted moral order. (One is tempted to suspect that part of the reason we so constantly engage imaginatively with such characters is precisely that we are unlikely to encounter many in person.) And just as we encounter moral monsters usually in a mental sphere rather than over coffee, so too our moral condemnation frequently occurs in contexts other than conversation. It is a challenge (to say the least) for Finlay's proposal to explain someone deliberating in her own mind over a moral dilemma or

---

[8] Not everyone buys this Humean thesis, of course, and rejecting it is another way of opposing certain error theoretic arguments. Finlay and I both endorse the Humean thesis, however (though no doubt with various qualifications).

struggling with temptation toward immorality. Let's say she is tempted to gratify an immediate selfish desire, yet she also strongly feels that she has a moral duty to refrain from performing the action—a duty she finds quite unappealing. She doesn't think of the duty in instrumental terms; she thinks of it as something she simply *must* do. Whence this sense of authority with which she imbues the moral option? One hypothesis is that she simply sees the imperative as non-institutional and categorical (though obviously not under that description). By contrast, the proposal that the imperative really is instrumental but she has supressed awareness of the end in question because it is so obvious, or because by doing so she creates a sense of demand via some kind of internal rhetoric involving ellipsis, seems quite unappealing.

I doubt, moreover, that Finlay's account does justice to the emotional profile of moral judgment. First, for the sake of comparison, consider the proposal that has sometimes been made that moral imperatives are hypothetical imperatives. According to this proposal, moral assessment is essentially *advice*, and moral wrongdoing is a species of self-harm. But this doesn't seem to mesh with how we *feel* about a moral wrongdoer: The emotions that go along with seeing someone harm himself (e.g., pity) appear quite distinct from the blame that we direct at the moral criminal. This seems sufficent ground for rejecting the proposal. A similar objection may be lodged against Finlay's theory. According to this theory, when a speaker, Mary, morally condemns someone, what she asserts is that the criminal has failed to further an end which she, Mary, endorses. But this too seems to fail to mesh with how we feel about a moral wrongdoer. When someone fails to act so as to further an end endorsed by Mary, or acts in a way that detracts from such an end, then Mary will probably *dislike* this—perhaps intensely—but something else is needed in order to account for the emotion of *blame* that goes along with a moral denunciation. Consider the difference between two cases. First, picture someone purposely frustrating an end which you endorse in circumstances where one would ordinarily think that she is perfectly within her rights to do so: Say, she outcompetes you fair and square for an important job and thus ruins your livelihood. Second, picture someone purposely frustrating an end which you endorse in circumstances where one would ordinarily think that she has acted inappropriately: Say, she outcompetes you for an important job by lying outrageously on her CV. Your emotional responses will be different to the two situations: for the first, disappointment and dislike; for the second, moral denunciation and blame. The question is whether Finlay's theory can accommodate this difference.

Certainly he can try. He might claim that in the second case your rival has failed to follow a further instrumental imperative: "In order to promote general human well-being [say], one ought not lie on one's CV." But rather than being an explanation of why *blame* is appropriate in the second case but not the first, this seems rather just to push the problem back, for now we can ask why, when someone fails to promote general human well-being, we do not merely feel *dislike* (because an end we endorse has gone unfulfilled); how does *blame* emerge? The problem here is possibly just a corollary of the aforementioned Humean-inspired worry: The natural response to this question is that the emotion of blame (rather than mere dislike) makes sense only when one judges that the end that the agent has purposely ignored *is* one of his/her ends (whether s/he likes it or not) or is an end that the agent *ought* to adopt. But it is (inter alia) its incapacity to satisfy these intuitive ideas in a convincing manner that renders Finlay's theory doubtful as an account of moral judgment.[9]

---

[9] Finlay is unlikely to be impressed by these last considerations, for he brackets off so-called *second-order* issues pertaining to blame from the first-order issues (pertaining to right and wrong, good and bad) on which our

I won't criticize Finlay's positive account further on this occasion; it deserves a much more careful treatment that I can offer here. I merely signal some serious misgivings about whether he has succeeded in accounting for the putative authority of moral language. If he has not, then support grows for competing hypotheses that explain this phenomenon—the conspicuous one being the claim that conceptually morality involves NICIs.

Does Finlay succeed in refuting the moral error theory?

The arguments discussed thus far pertain to whether the error theorist has succeeded in showing that morality possesses the commitments that seem so problematic. I agree that no error theorist has yet succeeded in making this case persuasively, in part because there is no agreed upon method or type of evidence for deciding such a claim (not just with respect to morality, but conceptual commitments in general). In other words, confirming premise A of the earlier argument is no easy matter. Yet refuting it is no easy matter either. But it is this latter ambition that Finlay sets himself. Not content with showing that I have yet to prove my case (a charge I fully accept, though naturally battle against), he deploys an argument to show that "error theory is actually false" (2008: 360).

To be honest, I really had trouble following Finlay's 2008 argument that was designed to make good on this bold claim. Now he's tried to clarify his argument in the 2011 rejoinder, but I remain both slightly confused and not in the least persuaded that his arguments show that error theory "is actually false."

The common ground between us is that regarding those topics for which we humans have uncovered significant errors in our thinking, there is a division to be observed. On the one hand there are those subjects for which we realize the error of our ways, correct our beliefs on the topic, and carry on with the same range of concepts as before. On the other hand, there are those subjects for which we realize the error of our ways but for which correcting our beliefs involves dropping concepts from our conceptual scheme. The number of items in both classes is large.

Pretty much every physical substance known to the folk is something about which false beliefs were once widespread: water, wood, gold, etc. We had massively false beliefs about virtually every astronomical object, property, and relation. We thought the Earth was unmoving, which gave us an erroneous idea of absolute motion. We had false beliefs about the biological world, both in the particular (e.g., thinking that gorillas are violent brutes) and in the general (e.g., in our ignorance of Darwinian selection). And so the list goes on. Yet, importantly, these mistakes didn't suffice for an error theory for the items in question; we still believe in water, wood, gold, the Sun, stars, motion, gorillas, and so forth.

In contrast, there is a large domain of things in which pretty much everyone sensible has stopped believing: witches, phlogiston, tapu, animism, totemism, superstitions, bodily humors (melancholia, etc.), geocentricity, extispicy, vitalism, homeopathy, astrology, alchemy, telekinesis, unicorns, and so on. Other things are widely disbelieved but still attract

present disagreement is focused. Indeed, he writes: "I confess to harbouring some sympathy for an error theory about judgements of blameworthiness, according to which they assume a form of responsibility for action that nobody actually has. ... But that is a separate question from whether that act is morally right or wrong" (2008: 359). I, however, am very far from accepting that these are separate questions. In fact, I have even more confidence in the claim that moral imperatives conceptually imply notions of blameworthiness than I have in the claim that they conceptually incorporate an insupportable kind of practical authority. Generally speaking, if someone is willing to grant me an error theory for blameworthiness and praiseworthiness, then I'll consider my moral skepticism pretty much home and hosed.

significant numbers of followers: gods, souls, sacredness, karma, sin, houris, transubstantiation, heaven and hell, reincarnation, etc.[10] (It will be noticed that there is a bit of an ontological hotchpotch here; some of the items are objects, some are properties or relations, and some are theories. I think that they could all be redefined as *properties* if one wished.)

Let us call the former category "revisable discourses" and the latter "erroneous discourses." (Note that the former are in a sense *erroneous*; it's just that the error is not so great as to render the discourse unrevisable.) The matter upon which Finlay and I disagree is into which category morality falls. One strategy which seems doomed to failure is to attempt to uncover some characteristic that all members of one category have and none of the other category have, and then proceed to show that morality has or does not have this characteristic. The reason this will fail, in my opinion, is that I doubt that there is anything that all members of either category have in common (and which no members of the other category have) beyond the fact that they satisfy the criteria for membership as described above. In other words, the only substantive thing that, say, reincarnation and geocentricity discourses have in common is that are bankrupt. There is not some other, independently specifiable, hidden common feature which renders these discourses erroneous. (This is not to deny that it might be useful to draw comparisons and note similarities for illustrative purposes.)

Finlay's argument focuses on a particular revisable discourse: motion. Once upon a time humans took various observed phenomena for evidence that absolute motion was occurring. But these data were also evidence that relative motion was occurring. Given this, it would be "grossly uncharitable," Finlay thinks, to accuse past speakers of having error theoretic commitments regarding motion; rather, we interpret them as referring to relative motion all along (even, possibly, in the face of their explicit intentions to the contrary). Finlay simply wants to take this model and apply it to absolute morality versus relative morality. In this he closely follows Gilbert Harman's lead, who writes:

> Before Einstein, judgments about mass were not intended as relative judgments. But it would be mean-spirited to invoke an "error theory" and conclude that these pre-Einsteinian judgments were all false! Better to suppose that such a judgment was true to the extent that an object had the relevant mass in relation to a spatio-temporal framework that was conspicuous to the person making the judgment, for example, a framework in which that person was at rest. (Harman & Thomson 1994: 4)

In my first critique (2011*a*) I misunderstood Finlay as claiming that there exists some relational property to which every moral judgment makes reference. The reason I did so is that the reversal of the quantifiers (that every moral judgment makes reference to some relational property) seems inadequate to his purposes. Bear in mind two things: (i) that Finlay is not requiring that the speaker is *aware* of the relational nature of his/her moral evaluations, and (ii) that it is important that the relativist theory allows for a distinction between true and false moral judgments. I don't recall Finlay explicitly saying anything about (ii), but as a kind of champion of common sense, he will, I assume, find this desideratum important. (I reckon a theory according to which every moral judgment anyone has ever made counts as true

---

[10] I don't want to get involved in debates as to which are considered erroneous by "nearly everyone" and which by just a substantial subset of persons. I myself think that they should all get lumped together as completely crazy ideas, but I'm trying to show a little tolerance.

relative to some R—and no other kind of truth is possible—is as "grossly uncharitable" to common sense as an error theory.) But these qualifications seem to raise problems for Finlay. Consider Jack the Ripper making a calm and sincere judgment that his actions are morally good. I assume that Finlay will want this judgment to come out false. But how can it? There will certainly exist *some* end that Jack's murderous actions promise to satisfy, such that if one identifies the property of conducing to that end with *moral goodness* then Jack's judgment is true. Of course, *we* won't relative *our* moral judgments to *that* end when we evaluate Jack's actions; we will relativize our judgments to an end according to which his action is morally forbidden. But Jack just chooses other ends, according to which his judgment is true. (And even if he doesn't consciously "choose" them, such action-end relational properties are still *there*.) And someone else can truly claim (privileging their preferred relational property) that Jack's actions are permissible but not required. Someone else can truly assert that his actions are noble and supererogatory. And so on. (I once admonished Jesse Prinz's relativism for having this same counter-intuitive result: "any given action could instantiate the full complement of moral properties: Hitler's Final Solution could be good, bad, permissible, evil, blameworthy, praiseworthy, obligatory, unfair, reasonable, and supererogatory—*all at the same time*" (Joyce 2009).) I should add that although I have chosen the examples of Jack the Ripper and Hitler in order to make the point palpable, the same problem goes for far more mundane moral misdemeanors. Basically, *any* wrongdoing (relative to end E1) will be rightdoing (relative to E2), morally neutral (relative to E3), and so on.

I'm not claiming that Finlay is stuck with this undesirable result; I'm just explaining why in 2011 I had assumed that he did not intend merely to relativize the truth of each moral judgment to *some* end. I suspect he will seek to place some restriction on the relational properties in question (such that Jack's judgment comes out false), but since Finlay has yet to do so (as far as I can see), the adequacy of his positive theory remains difficult to assess.

Putting all that aside, the reason that I am slightly confused by Finlay's argument is that I fail to see anything strikingly novel in it. The absolutist about morality has always known that every absolutist claim has a relativistic surrogate. Whenever someone asserts "X is wrong (period)!" there is always lurking in the wings the wimpified version: "X is wrong (relative to standards S)." The absolutist who worries that the former is indefensible is usually perfectly aware that the latter may well be true. The route of resisting the error theory by denying that morality is non-negotiably committed to the supposedly problematic component is a well-trodden one, and the surrogate that is offered as a successful successor is generally some kind of relativistic or relational or subjectivist one. Proponents of this strategy include Harman 1975, Railton 1986, Lewis 1989, Johnston 1989, Dreier 2005, Copp 2005, and Prinz 2007. In a collection of essays I co-edited on Mackie's metaethics (Joyce & Kirchin 2010), the strategy is advocated by Simon Kirchin and Caroline West. Far from being a radical new feint against moral skepticism, Finlay's general position is virtually orthodoxy.

Given this impressive array of opponents, why would an absolutist stick to his guns rather than just caving in to a relativistic or relational view? The answer is that the absolutist reckons that the absolutist conceptual scheme bears some crucial characteristic(s) that the latter lacks, such that the latter, though true, simply cannot pull the practical weight that the former can. For the sake of a simple example, consider the crude relativism according to which "X is morally right" means "X is approved of by my culture." Someone with skeptical worries about moral rightness will in all likelihood accept that most of those actions called "morally right" do indeed instantiate the property of being approved of by the speaker's culture. Yet despite the availability of this relativistic property which promises to yield true

claims, most people will balk at this proposed crude relativism. Why? Because, in broad terms, we want to do things with our concept of *moral rightness* for which the relativistic substitute just doesn't seem to provide license.[11]

Instead of the motion case, consider the witch case. Another point of agreement between Finlay and myself is acceptance that there may be some property that all and only those women who have been accused of being witches actually instantiate. (I don't think that anything like this is always true of erroneous discourses, but it may well be for some.) We needn't obsess too much about the details; the point is just to allow for the sake of argument that there may be such a property. Back in *Myth of Morality* I suggested that all and only women accused of witchcraft might "play a certain disruptive role in the patriarchal society" (96); Finlay prefers: "a complicated disjunctive property including, for example, the property of being a woman whose enemies have suffered illness and misfortune" (2008: 366). Finlay recognizes that this threatens to upset his central argument, since the question arises as to why we cannot revise our witch discourse such that "witch-hood" refers to this actually instantiated property, so that all those witch judgments turn out to be true. If this is possible, then Finlay's attempt to show that morality has a characteristic that puts it with motion (in the revisable pile) and not with witches (in the erroneous pile) falls flat.

I found Finlay's response to this in 2008 rather unclear, and in 2011 he has added a new element:

> Moral judgments therefore differ from witch judgments..., and are akin to motion judgments, in that while the actually instantiated properties of relative value and motion form unified kinds with significance for people, the actually instantiated property in the case of witches is gruesomely disjunctive. It isn't plausible that ordinary people's witch judgments aim at tracking such an uninteresting, gerrymandered property, and we can safely presume that they rather infer from the concrete evidence the interesting but uninstantiated property of being a witch. (546)

This is more than a clarification of the 2008 argument; it's a new consideration that distinguishes "unified kinds" from "gruesomely disjunctive" properties. I would like the reader to notice that this is really the heart of Finlay's whole argument purporting to demonstrate that the error theory is false. This is where the blow is supposed to fall.

So is the error theorist laid flat? Not at all. Notice, first of all, that there is nothing essentially disjunctive about the witch case, for there is a very straightforward relational property that all and only those women referred to as "witches" instantiated: they all had the property of *being accused of practicing witchcraft*. It might be objected that this is still a "gerrymandered" property (even though it is not disjunctive), but in response I would point out that it is certainly a property in which real people might be interested. Consider: "All

---

[11] The absolutist need not always accept that the relativistic surrogate yields truths. Even if Finlay is correct that moral claims are of an action-end instrumental format, this wouldn't automatically assure truth, for there may be something systematically wrong with the end. Mackie, for example, allows that "X is good" claims are always of the form "X satisfies the requirements of the kind in question" (1977: 55), and adds that when it is a *moral* assessment then the requirements are supposedly "simply there, in the nature of things" (59). But Mackie goes on to deny that there *are* any such universe-given requirements, and thus (even while allowing that moral judgments have Finlay's action-end relational form) he holds that all moral judgments about goodness are simply untrue. Another example would be any theory that identifies moral properties with what conduces to the will or commands of a (non-existent) divine being. This too would produce an error theory even while allowing that moral judgments have an action-end format. For discussion of how metaethical attempts to avoid the moral error theory sometimes stumble right into it, see Joyce 2011*b*.

women who have been accused of witchcraft shall be tried in the coming months." Or imagine a woman who has been charged with witchcraft: She is likely to be far more interested in the fact that she has *been accused* than in the question of whether she is actually a witch.

But irrespective of what we end up saying about the witch case, the real question is whether what Finlay says of this one example goes for other erroneous discourses. Even if in certain respects morality is more like motion than like witches, in other respects it may be more like other examples of erroneous discourse than like motion. Finlay, I think, would do well to ponder the phlogiston case. Here there is a perfectly unified kind available—namely, *oxygen*—which might have been used as the referent of a revised discourse (but wasn't). There is nothing gerrymandered, ad hoc, gruesomely disjunctive, or uninteresting about oxygen. Why, then, did scientists of the 18th century not just *revise* their mistaken phlogiston theory to the successful oxygen theory—correcting some erroneous beliefs about how the substance behaves in combustion? Let me repeat the answer I gave over a decade ago:

> the whole point of talking about phlogiston was to make reference to a substance that is *released* during combustion. To use the word 'phlogiston' to refer to oxygen—a substance that is *consumed* during combustion—is to undermine the very heart of phlogiston discourse. (2001: x)

And this is more or less what the moral absolutist can say about Finlay's relativistic surrogate: that the absolutist concept *is the whole point of the discourse*, that it furnishes us with something authoritative that the wimpified relativistic surrogate cannot. (Importantly, this arguably marks a disanalogy with the motion case.)

Now, these are not the terms with which *I* would present an error theoretic argument (confused comments of my past self notwithstanding); if I've made anything clear I hope it's my determination to eschew any mention of "absolutism" in making my case. What I am tempted to say, rather, is that seeing moral imperatives and values as transcendent of any human institution—as items not of our creation—furnishes us with authoritative possibilities that any wimpified version which acknowledges its institutional origins cannot. These authoritative possibilities, moreover, are the heart of the discourse: A system of categorical imperatives derived from a human institution does not adequately bind someone who doesn't care about that institution (i.e., that end), or who cares only occasionally and/or weakly. Of course, those who do care can complain about transgressors, punish them, and pull linguistic tricks to obscure the institutional foundations. The worry is that this makes morality out to be, at bottom, a species of whining combined with bullying along with a touch of rhetorical obscurantism. And moral norms, conceptually, are more than that.

This claim, obviously, is something about which Finlay and I can have a good old argument. I do not suggest that I have ever succeeded in demonstrating its truth. But Finlay's arguments ostensibly undermining the case I have made largely miss the mark, since he persists in misidentifying the thesis that I (and Mackie) have sought to establish. The only time that he engages with the real error theoretic worry is when he tries to accommodate authority within his relational framework (the bit about leaving the relativist component tacit), but I have expressed my doubts about this. So nothing has been settled, and in fact Finlay's discussion has not obviously advanced the debate one inch from the point at which we all already knew it stood. Thus his bold claim to have demonstrated that "the error theory is actually false" has proven to be vastly overblown.

If we are still sticking with the Lieutenant Columbo metaphor, then this is the point at which the LAPD enter and a handcuffed Professor Finlay is led away.

Victoria University of Wellington

REFERENCES:

Copp, D. 2005. "A skeptical challenge to moral non-naturalism and a defense of constructivist naturalism." *Philosophical Studies* 126: 269-283.

Dreier, J. 2005. "Moral relativism and moral nihilism." In D. Copp (ed.), *Handbook of Ethical Theory*. Oxford University Press: 240-264.

Finlay, S. 2008. "The error in the error theory." *Australasian Journal of Philosophy* 86: 347-369.

Finlay, S. 2009. "Oughts and ends." *Philosophical Studies* 143: 315-340.

Finlay, S. 2011. "Errors upon errors: A reply to Joyce." *Australasian Journal of Philosophy* 89: 535-547.

Harman, G. 1975. "Moral relativism defended." *Philosophical Review* 84: 3-22.

Harman, G & Thomson J. J. 1996. *Moral Relativism and Moral Objectivity*. Cambridge, MA.: Blackwell.

Johnston, M. 1989. "Dispositional theories of value. Part III." *Proceedings of the Aristotelian Society*, suppl. vol. 63: 139-174.

Joyce, R. 2001. *The Myth of Morality*. Cambridge University Press.

Joyce, R. 2009. "Review of Jesse Prinz's *The Emotional Construction of Morals*." *Mind* 118: 508-518.

Joyce, R. 2011*a*. "The error in 'The error in the error theory.'" *Australasian Journal of Philosophy* 89: 519-534.

Joyce, R. 2011*b*. "The accidental error theorist." In R. Shafer-Landau (ed.), *Oxford Studies in Metaethics, Vol. 6* (Oxford University Press): 153-180.

Joyce, R. & Kirchin, S. 2010. *A World without Values: Essays on John Mackie's Moral Error Theory*. Springer.

Lewis, D. 1989. "Dispositional theories of value. Part II." *Proceedings of the Aristotelian Society,* supplementary vol. 63: 113-137.

Prinz, J. 2007. *The Emotional Construction of Morals*. Oxford: Oxford University Press.

Railton, P. 1986. "Moral realism." *Philosophical Review* 95: 163-207.

Stevenson, C.L. 1963. *Facts and Values*. New Haven CT: Yale University Press.