

This is a preprint of an article whose final and definitive form will be published in the *Australasian Journal of Philosophy* in 2013; the *Australasian Journal of Philosophy* is available online at: <http://www.tandf.co.uk/journals/>.

## SCIENCE'S IMMUNITY TO MORAL REFUTATION

Alex Barber

### ABSTRACT

Our moral convictions cannot, on the face of it, count in evidence against scientific claims with which they happen to conflict. Moral anti-realists of whatever stripe can explain this easily: science is immune to moral refutation because moral discourse is defective as a trustworthy source of true and objective judgments. Moral realists, they can add, are unable to explain this immunity. After describing how anti-realists might implement this reasoning, the paper argues that the only plausible realist comeback turns on the practical nature of moral reasoning. This comeback, however, places significant constraints on the structure of evidence for moral judgments. These constraints cannot be met by coherentist defenders of reflective-equilibrium methodology or by anyone sympathetic to bottom-up, case-driven foundationalism, including those who claim we have perceptual or perception-like access to moral truths. Unfortunately for realists in these categories, alternative realism-friendly accounts of science's apparent moral immunity are unpromising. These neither explain nor explain away our unwillingness to infer an *is* from an *ought*, as Hume might have put it. Science's apparent immunity to moral refutation therefore poses a serious problem for any realist unhappy with top-down, theory-driven conceptions of the structure of moral evidence.

### KEYWORDS

moral realism, moral epistemology, moral immunity

Arguments against moral realism usually speak in favour of some particular variant of anti-realism.

In this paper I describe an argument for anti-realism that floats above distinctions between its sub-species. For reasons that will become clear, I call it the argument from moral immunity. Its import is that moral discourse is defective as a trustworthy source of true and objective judgments, but because the argument is neutral on the precise nature of the defect it can be adopted by non-cognitivists (who deny that our moral utterances express judgments), subjectivists (who deny them objectivity), extreme sceptics (who deny them trustworthiness), or error theorists (who deny them truth).<sup>1</sup>

---

<sup>1</sup> By *moral realism*, then, I mean the view that moral discourse is a trustworthy source of true and objective judgments. This is rough but serves for present purposes. It has an epistemic dimension and some readers may prefer to substitute 'non-sceptical moral realism' wherever I use 'moral realism'.

The argument centres on a phenomenon that it seems even the most hard-boiled moral realist must acknowledge: our strong resistance to a moral claim's being cited in evidence against a scientific claim with which it is inconsistent. Oddly, this phenomenon has not been pressed into service by anti-realists in the way I am about to set out. I describe the phenomenon in §1, show how readily anti-realists can accommodate it in §2, and consider how moral realists might approach it in §3—either by denying the existence of the phenomenon or by explaining it in a realism-friendly way. Being a moral realist by inclination, I do not set out to dismiss all realist responses, but several are surprisingly flimsy once one moves beyond their initial appeal, and the single plausible response has serious implications for the structure of moral knowledge. Specifically, I argue in §4 that the phenomenon forces realists towards a top-down, theory-driven foundationalist moral epistemology and away from bottom-up, case-driven moral reasoning, analogies between moral intuition and perception, and reflective-equilibrium coherentism.

### **1. Science's moral immunity**

The moral-immunity phenomenon has to do with how we may respond to scientific (or purportedly scientific) studies that challenge (or purportedly challenge) received moral wisdom. Here are sketches of five examples:

- (i) Various studies were presented by the authors of a notorious 1994 book, *The Bell Curve*, as evidence for a partially genetic intellectual hierarchy among racial groups [Herrnstein and Murray 1994].
- (ii) Freewill is an illusion according to neuroscientists who claim to be able to predict which of two options a subject will select, and to do so long before the subject herself is aware of deciding—up to ten seconds before in one study [Soon, Brass et al. 2008].

- (iii) Marine biologists claim to have discovered in fish the behavioural tropes and neural structures that, in humans, are strongly associated with extreme pain and distress (see Braithwaite [2010] for an overview).
- (iv) Psychologists and empirically minded philosophers (e.g. Harman [1999]; Doris [2002]) have questioned the robustness of character types presupposed in both folk and philosophical versions of virtue-theoretic moral assessment.
- (v) According to some physicists, our universe is but one among many physical universes, each isolated from but just as real as the one we inhabit; nothing we do makes any difference to the total good (e.g. the total amount of happiness) across reality's entirety.<sup>2</sup>

With due emphasis on the words 'alleged' and 'potentially', these alleged findings potentially conflict with prevailing moral assumptions about, respectively, *the just distribution of resources*, *moral responsibility*, *food ethics*, *character assessment*, and *acting to maximize value*. From each example we are able to generate an apparently inconsistent set. Logic then compels us to reject either the alleged scientific finding, the prevailing moral assumption, one of the background assumptions used to generate the apparent inconsistency, or the appearance of inconsistency itself, which could depend on a term's being used ambiguously. The phenomenon of moral immunity is this: while the scientific finding (plus the background assumptions and logic) may count as evidence against the moral assumption, it seems deeply misguided to regard the moral assumption (plus this same background) as evidence against the scientific finding.

---

<sup>2</sup> The 'many-worlds' interpretation of quantum mechanics has an analogue in philosophy, modal realism, the apparent moral implications of which are drawn out by Mark Heller [2003] and others.

To illustrate this we can focus in on two of the examples. In (iii) the scientific claim is that our consumption of fish causes them enormous amounts of pain, and the relevant prevailing moral assumption is that it is morally permissible for us to consume fish. The background needed to generate an apparent inconsistency between the two will include both moral and non-moral assumptions, e.g. that it is morally impermissible for us to cause unnecessary pain, that if our consumption of a sentient creature causes pain and we can survive happily without consuming it then the pain that our consumption causes is unnecessary, and that it is indeed possible for us to survive happily without consuming fish. When assessing the overall strength of evidence for and against the scientific claim, it seems a mistake to include, alongside the various neurological and behavioural studies, the fact that it belongs to this inconsistent set. That fact cannot, for example, be taken as a sign that there is some methodological flaw in the studies' data or in the interpretation of those data (though of course it may quite reasonably motivate a hunt for such flaws).<sup>3</sup>

(iv) is more complicated. The received moral assumption under threat here is that character-based commendations or condemnations of people or actions—as honest or cowardly, say—are, often anyway, correct. The main background assumption is that such commendations and condemnations are mistaken unless humans manifest the robust, stable and situation-independent behavioural patterns associated with the relevant character trait. The scientific studies that threaten to generate an inconsistent set purportedly show that humans fail to manifest these patterns. It would be odd, even for someone who attaches high antecedent plausibility to character-based assessments and to this background assumption, to treat this inconsistency as evidence against the studies (as opposed to a motive for scrutinising the studies very closely).

---

<sup>3</sup> Notice that the inconsistent sets can have both axiological and deontological elements. The set in (iii), for example, may contain the deontological requirement that we not cause negative value and the axiological claim that pain has negative value. So someone could account for what I am sweepingly calling science's 'moral' immunity by being realist about deontology but anti-realist about axiology—or vice versa [Tappolet 2000: 63-7].

Or would it? This case is less straightforward than (iii) because character traits seem to have both a moral and a non-moral dimension built into them, creating space for an objection. Hunches often provide highly defeasible evidence against the scientific studies with which they are inconsistent, and plenty of us have a hunch that the world is indeed populated by the honest and the dishonest, the brave and the cowardly, etc. Does this show that scientific findings aren't, after all, entirely immune to moral counterevidence? To see why not we need to distinguish our ordinary conceptions of the character traits from conceptions that have been stripped of ethical content. Suppose C is entirely neutral on the commendability or otherwise of the trait of, say, never knowingly misleading one's audience. C has a strong hunch that some people have this trait while others lack it. C's hunch is challenged by the studies in (iv), but equally, C's hunch is a challenge to those studies. D lacks C's neutrality. D has C's hunch plus a hunch that various people are commendable because of their possession of the trait. The immunity phenomenon as it applies to this example is best described as: D's extra hunch adds nothing to C's in the way of evidence for or against the studies in (iv). Some hunches sometimes have evidential import for science, but moral hunches *as such* never can.

Extrapolating from these and the other examples, we have an (apparent) asymmetry: scientific claims can figure as evidence within the moral domain yet moral claims as such cannot figure as evidence within the scientific domain. Moral realists and anti-realists alike recognize this (apparent) asymmetry, yet anti-realists seem far better placed to account for it (or its appearance). That, in essence, is the argument from moral immunity. But before turning in §2 and §3 to how those in each camp should interpret the immunity phenomenon, I want to bring the phenomenon itself into sharper focus with three important clarifications.

The first is implied by my use of parentheses in the preceding paragraph. I shall speak loosely of 'the moral immunity phenomenon' but it is open to question whether the phenomenon is genuine, i.e. whether science really is immune to moral refutation. At this point in the paper I am only

claiming that the appearance of immunity needs to be either explained or explained away. Later we'll see that both strategies are available, to realists and anti-realists alike.

Second, what (apparently, etc.) has immunity from moral considerations is the *strength of evidence for* rather than the *acceptability of* a scientific claim. Acceptability is the broader notion. It takes into account practical ramifications as well as strength of evidence. Moral considerations can legitimately enter into deliberations over acceptability because, for example, we formulate social policy on the back of accepted theories. After a risk analysis of the relative costs and benefits of error, we may decide to accept and act on a theory with relatively weak evidential support. Allowing moral and other pragmatic concerns to influence our assessment of acceptability is consistent with refusing to let them influence our assessment of strength of evidence.<sup>4</sup>

Finally, granting that science has moral immunity is compatible with a healthy scepticism towards scientific reports that trumpet a particular social outlook. Scientists are human and humans have all manner of conscious and unconscious biases. If a study appears to licence racism, this fact cannot stand as counterevidence in and of itself, but it does give us reason to suspect that we'll find problems with, say, the methodology used or the interpretation of data, since shoddy science with a racist agenda has a rich lineage. (On the *The Bell Curve* in particular, see Newby and Newby [1995].)

## **2. Anti-realist perspectives on moral immunity**

The general form of any anti-realist explanation of science's moral immunity is that moral discourse is defective as a trustworthy source of true and objective judgments.<sup>5</sup> Realists, in contrast, face a difficulty: the more success they have in establishing the credentials of moral judgments, the harder

---

<sup>4</sup> This is what Philip Kitcher has in mind when he writes, of sociobiology, that 'the question of what counts as sufficient evidence... is not independent of political considerations' [1985, p. 9].

<sup>5</sup> The qualification ('as a... etc.') is important. Many anti-realists do not see moral discourse as defective *simpliciter*.

they will find it to explain why moral judgments cannot contribute to settling scientific disputes. That's the topic for §3. In this section I describe how each of the four main types of anti-realism identified in the introduction might be used to account for the immunity phenomenon, then address a simple objection that applies to all four accounts. This fourfold classification is deliberately course-grained. My aim here is simply to generate a presumption that anti-realists have plenty of options; the main goal of the paper is to show how narrow the options are for realists.

*Moral non-cognitivism:* Non-cognitivists ('expressivists') hold that our moral utterances express, not genuine judgments of belief, but mental states for which truth is an inappropriate measure of evaluation. Premises that are not even truth-apt, let alone true, cannot legitimately be used to infer anything that is truth-apt; hence we can never reasonably infer from our moral convictions to the falsity of a scientific theory with which they seem to conflict.

This is too simplistic as it stands. Many moral expressivists are also epistemic expressivists, holding that when we give voice to norms in either domain we are expressing non-representational mental states (see e.g. Gibbard [2003], and Chrisman [2012] for a critical overview). This raises the question of whether moral expressivists can recognize the immunity phenomenon as genuine. They could do so by denying that moral expressivism needs to be matched by expressivism in the epistemic domain, or by embracing the distinction drawn earlier between the *acceptability* of a theory (to be understood expressivistically) and the *strength of evidence* for a theory (to be understood non-expressivistically). Since it is ostensibly about evidence not acceptability, the immunity phenomenon could still be endorsed as genuine. (For sobering thoughts on the prospects for a non-cognitivist understanding of 'evidence', see Schroeder [2010: 183–4].) Other expressivists, though, may prefer to reject that earlier distinction. For them, to say that science is immune to moral refutation is not so much to describe a phenomenon in need of explanation as to commit to never letting ethical considerations enter into scientific deliberation. This rejection of immunity as a genuine phenomenon would hardly

help moral realists, however, since they are unlikely to be epistemic expressivists and so would have to find their own way of explaining—or explaining away—science’s apparent moral immunity.<sup>6</sup>

*Moral subjectivism:* Moral subjectivists allow that moral utterances express genuine judgments of belief but deny that these judgments are objective. Their truth conditions are tied too closely to the inclination of the judger to make the judgment. If this is right, the feature of our moral convictions that explains why we cannot draw on them to overturn scientific claims is poverty of content. Their content is so proximal that they cannot be used to draw objective scientific—distal—conclusions.

The loose definition offered here sidesteps the vexed topic of how best to understand the subjective/objective distinction. It should, though, be compatible with views that emphasize, for example, mind dependence or relativity to social context (e.g. Prinz [2009] or Harman [1996]).

*Moral scepticism:* Scepticism deserves to be classified as a form of moral anti-realism only as it tends towards extreme and generalized pessimism about the quality of evidence available in the moral sphere. Sceptics of this calibre can readily account for its being wrong to pitch moral considerations against scientific evidence, since an argument is only as secure as its premises.

Moral sceptics of this strength tend, in practice, to be outright error theorists. (Walter Sinnott-Armstrong, for example, though not a moral nihilist, also describes himself as only a moderate moral sceptic [2006: 58, 250].) I have kept the questions of evidence and truth apart, though, so as to make

---

<sup>6</sup> Another complication is that non-cognitivists are often drawn to deflationary notions of truth and/or belief. I leave open here the question of whether these moves are compatible with explaining the immunity phenomenon.



room for the realist strategy, discussed under (c) in §3 below, of treating moral judgments as a weak rather than non-existent source of evidence in the scientific domain.

*Error theory:* Error theorists have the simplest explanation of all for why we should not call on our cherished moral beliefs to settle a scientific question: they aren't true (see e.g. Mackie [1977] and Joyce [2001]).

I have just outlined four broad ways in which anti-realists could explain or explain away science's apparent moral immunity. A single assumption unifies this otherwise motley group: moral discourse is defective as a trustworthy source of true and objective judgments, disqualifying it as a source of premises with which to challenge an alleged scientific finding. Anti-realists hoping to use the argument from moral immunity against realists do, though, face an interesting puzzle. Pretty much everyone is sympathetic to the thought that science is (genuinely or apparently) immune to moral refutation, but since not everyone is a moral anti-realist it looks as though something other than anti-realist commitments must explain this sympathy. What is this something? And might it not generate an alternative explanation of moral immunity, one that even a realist could embrace?

The simplest response is to defend an 'implicit belief' account of our recognition of moral immunity. Anti-realists can say that only at an explicit, conscious level do any of us hold moral realist views. Below the radar of conscious reflection we are all moral anti-realists. This unconscious commitment reveals itself in our refusal to draw directly on our moral convictions when arguing against scientific evidence, no matter how feeble we regard the latter. The response is less ad hoc than it may seem. First, comparable appeals to the notion of implicit ('tacit', 'unconscious', 'sub-personal') mental states are now extremely widespread in the cognitive sciences. Second, many moral anti-realists are already committed to our minds being somewhat opaque to us. Non-cognitivists, for example, accept that we don't consciously appreciate the attitude-type expressed by our moral utterances. If we did, there'd be no debate. What's new on the present proposal is just that

at an unconscious level we *do* appreciate the anti-realist character of our moral outlook—we do appreciate how non-truth-apt our moral attitudes are, for example, though each variant of anti-realism will have its own way of spelling out the suggestion.

### 3. Realist perspectives on moral immunity

Moral realists in the present context share a characteristic vulnerability. Traditionally they have sought to defend the status of our moral pronouncements—by shrinking an apparent evidence deficit, by overcoming concerns about a lack of objectivity, and so on.<sup>7</sup> Suppose we grant success to them in this. Suppose we allow that moral judgments are every bit as respectable as judgments in other domains, including science. In doing so we seem to place realists in the embarrassing position of being able to deploy moral considerations *against* these other judgments whenever there is a clash.<sup>8</sup>

How can they defuse this embarrassment? We can divide their options into attempts to explain moral immunity away by denying that it is a genuine phenomenon, and attempts to explain the phenomenon in ways that are compatible with realism. In what follows I consider three instances of the explaining-away strategy ((a)–(c)) and two of the explaining strategy ((d) and (e)). Only (e) stands up to scrutiny. I discuss the failed attempts because (e) has burdens—set out in §4—and my overall conclusion is that a moral realist has no choice but to take on those burdens. This conclusion depends on ruling out other stances a realist might take towards moral immunity.

What follows is written in the spirit of an invitation to those moral realists who don't like the conclusion to supply something better. There is no space here to anticipate all conceivable realist

---

<sup>7</sup> Geoffrey Sayre-McCord, for example, opens his [1988] with, 'Among the most enduring and compelling worries about moral theory is that it is disastrously isolated from confirmation,' before trying to supply the missing justification.

<sup>8</sup> Moral realists do sometimes worry about having 'proved too much', but their concern is only that other manifestly non-realist discourses—humour, etiquette—could be vindicated as realist by the same pattern of reasoning. See e.g. Huemer [2005: 211–4]; Sturgeon [2006: 105–6].

responses, and personal judgment is a factor in how much space to give even to those I do consider.<sup>9</sup> My sense is that at least some of them could be developed so as to converge with (e). In that case, however, they would share (e)'s burdensome consequences.

I start, then, with three attempts to show that moral immunity is but an illusion, making its compatibility with moral realism moot.

*(a) The moral immunity phenomenon presupposes that you can derive an 'ought' from an 'is'; but as Hume pointed out, you can't.*

The studies sketched in §1 were presented as having potential moral import. Citing Hume on the impossibility of deriving an *ought* from an *is*, a moral realist could insist that the studies must all have been misleadingly described, and hence that we never really have to choose between an existing moral view and a scientific finding. If this is right, the argument from moral immunity is presupposing a phenomenon that doesn't really exist.

Looking at Hume's dictum in context helps us to see where this response goes wrong. He was making a purely formal point: that you need a bridging premise to get from a set of premises that makes no mention of morality to a conclusion that does, and the bridging premise would itself need to include a moral component.<sup>10</sup> Hume makes this formal point while drawing out the implications of his foundationalist empiricism, which had led him to wonder where the evidence for moral claims

---

<sup>9</sup> My experience has been that realists vary considerably in how they respond to the argument from immunity. Strategies dismissed by some as straw targets are heralded by others as clearly along the right track. This is true of the approaches I consider in the main text as well as others I have not included, such as that of trying to explain science's immunity to moral refutation in terms of our resistance to the *ad misericordiam* fallacy.

<sup>10</sup> See final paragraph of 3.1.1 of his *Treatise*. I say 'purely formal' because the point is general. Setting aside trivial cases, the same would hold of an inference from premises that make no mention of water to a conclusion that does.

could come from. Because his contemporaries had failed to spot the need for well-supported bridging premises, they were in effect cheating.

Moral realists are well advised to steer clear of Hume's astringent epistemology, not least because of the subjectivist direction in which it took Hume himself. But the simple formal point—the catchy *is/ought* slogan divorced from Hume's broader empiricism—won't help them either because the immunity phenomenon is undisturbed by it. The phenomenon was characterized in §1 in terms of how we may respond to inconsistent sets. To be genuinely inconsistent these sets will include bridging assumptions of precisely the kind that Hume's contemporaries supposedly overlooked. Once an inconsistency is established, though, the immunity phenomenon is readily described. To reverse the slogan, we are resistant to inferring an *is* from an *ought*. The salient *is* will be the alleged scientific finding while the *ought* will really be a conjunction of all the *oughts* in the set together with any other background assumptions.

*(b) We should sometimes question scientific judgments on moral grounds.*

Not all inferences from moral premises to scientific or other non-moral conclusions are as immediately unappealing as those envisaged in §1. Nicholas Sturgeon notes how readily we might try to explain a growth in opposition to slavery by noting its increasingly oppressive character, or the degradation and death of millions by reference to Hitler's depravity [1988: 232, 245]. These explanations embody inferences from premises involving the moral concepts *oppression* and *depravity*. Do they show, then, that the immunity phenomenon is not genuine and hence that realists are under no pressure to explain it?

Before accepting this we should recall that, as with Hume's dictum, Sturgeon's examples have their own context. He is responding to Gilbert Harman's [1977] claim that purported moral facts are explanatorily inert (unless given a reductionist translation), depriving (non-reductionist) ethical naturalists of inference to the best explanation as a route out of moral scepticism. The Harman-

Sturgeon exchange is not my direct concern here, and its connection to the moral immunity phenomenon is subtle. One obvious contrast is that it is concerned with the availability of evidence running from non-moral to moral, by inference to the best explanation, whereas the present paper is asking about evidence in the other direction.<sup>11</sup> Still, good explanations often supply good predictions, and Sturgeon's examples can be adapted to generate evidence running from moral to non-moral.<sup>12</sup> More significantly, Sturgeon's examples are offered as part of an existence proof: he wants to show that ethical theory is sensitive to observational evidence by finding suitable examples of moral explanations. Even if (*pace* e.g. Miller [2003: 138–77]) we side with Sturgeon and endorse *these* (and relevantly similar) moral explanations of non-moral phenomena, thereby embracing the possibility of *some* moral to non-moral inferences, other moral to non-moral inferences—in particular those in §1—lack even *prima facie* plausibility. So perhaps moral reasoning sometimes has and sometimes lacks evidential force in the scientific domain. How does this possibility affect the topic of science's alleged moral immunity?

It would seem that *either* there is something about the §1 cases (and others like them) that accounts for their status as real or apparent exceptions to the broadly true rule that moral reasoning has evidential potency in the scientific domain; *or* there is something about Sturgeon's cases (and others like them) that accounts for *their* status as real or apparent exceptions to the broadly true rule that moral reasoning is evidentially irrelevant in the scientific domain. A realist pursuing strategy (b) will be sympathetic to the first disjunct, but will still need to explain the apparent exceptionality of the other cases, possibly by drawing on (c)–(e) in the hope that these don't undermine the Sturgeon-cases. Strategy (b) alone will not suffice. But the second disjunct is more tempting in any case, in so

---

<sup>11</sup> Harman does talk of morality's 'apparent immunity from observational testing' [1977: vii], but this is immunity in the non-moral-to-moral direction.

<sup>12</sup> Suppose I wish to know whether a regime in 4500BCE survived to 4400BCE. The later historical record is inconclusive. I uncover credible evidence that in 4500BCE the regime was oppressive in a way that can (assuming Sturgeon is right and Harman wrong) explain regime collapse. This gives me defeasible evidence that it did not survive.

far as Sturgeon-type cases have something in common that could account for their status as apparent exceptions: by rephrasing them in non-moral language we can bracket<sup>13</sup> our moral convictions without apparent loss of explanatory force.<sup>14</sup>

In the context of his debate with Harman, Sturgeon is unimpressed by such rephrasing. When Harman suggests that ‘the psychology or moral sensibility’ of the individuals involved suffices for good explanation [1977: 6], Sturgeon replies that we have no motive for falling back on these non-moral features rather than sticking with the morality-involving explanations he prefers—or rather, no motive that is specific to moral explanations and so wouldn’t also jeopardize, say, microphysical explanations [1988: 250–3]. Whether or not this is a fair point in the context of his dispute with Harman, in the present context we have a morality-specific motive for taking the availability of such rephrasing seriously, one that Sturgeon does not address. If science seems partially but not wholly immune to moral refutation, that appearance needs explaining; no parallel need exists for in the microphysical case. In the moral case, moral-to-non-moral evidence that seems legitimate (see e.g. note 12) is arguably special because it can be repackaged in non-moral language without loss of force, i.e. made to work even if one brackets one’s moral stand. This is not true of the cases in §1 (with the possible exception of (iv), for reasons discussed already). Still (and as with (iv)), Sturgeon-type cases do at least suggest we need to qualify the immunity phenomenon as the unavailability to science of evidence with a *non-bracketable* moral ingredient.

---

<sup>13</sup> ‘Bracket’, like ‘be agnostic about’ in the next note, should not be equated with ‘treat as false’. Comments by Sturgeon [1988: 250] could encourage such a reading. Beyond making this point I leave it open how bracketability in my sense intersects with the Harman-Sturgeon debate.

<sup>14</sup> We could explain increased resistance to slavery, for example, as arising from the fact that people *came to regard* the institution as increasingly oppressive, *took* this increase to undermine its legitimacy, and were spurred to action *by this assessment*. Here, the moral dimension of the explanation is placed within psychological operators and there is no need for us to take a stand on whether the institution really is oppressive, illegitimate, or deserving of resistance. We can be agnostic about these questions without harming the explanation. An alternative or complementary route to rephrasing would involve teasing apart the moral and non-moral dimension of the concept of oppression—cf. the treatment of honesty in §1.

(c) *Science is not immune to moral refutation but the evidence we have for our moral judgments is generally weak when set against the evidence of our eyes and ears.*

Only extreme moral sceptics can reasonably be classified as anti-realists (see §2). Weak or moderate sceptics—still realists, potentially—can try to explain moral immunity away by noting the relative frailty (as opposed to non-existence) of the evidence underpinning moral judgments when it is set against the robust perceptual evidence underpinning science.<sup>15</sup>

This move founders on the observation that direct moral evidence counts for *nothing at all* in the scientific sphere. Suppose two scientific theories are fairly evenly matched on the usual empirical measures. ((v) in §1 is a possible case in point. Physicists have struggled to find empirical data to discriminate between the many-worlds view and its competitors.) Even then, a moral judgment's consistency with just one of the two won't shift the evidential balance. It wouldn't even matter that this moral judgment seems relatively secure. This shows that a mismatch in strength of evidence between the scientific and the moral domains can explain all cases of moral immunity only if evidence in the moral domain is vanishingly weak—an anti-realist outcome.

A realist could respond that intuition is on her side here. Suppose you think the empirical data support Theory X over its competitors by the slimmest of margins. An infallible oracle then informs you that if X is correct then it is permissible to torture irritating neighbours. Should this testimony not require us to assign a lower probability to X? My own sense is that it should not, but contrary intuitions call for comment.<sup>16</sup>

One consideration is that our moral values seems to affect our intuitions about non-moral questions, without our realising it, even in scenarios where this influence is manifestly illicit (see

---

<sup>15</sup> Realists frequently stress the defeasibility of the evidence for moral claims. See e.g. Sayre-McCord [1996: 138]; Huemer [2005: 105–6]; McMahan [forthcoming].

<sup>16</sup> Several people, including an anonymous referee for this journal, have offered variants of this oracle example to suggest that science may not be immune in the way I suppose.

Knobe [2010] and in particular the studies sketched on pp. 317–20). Whatever cognitive mechanism is at work in these settings—and it seems to be unconscious save at the output-level of intuitions about cases—could also explain why the oracle’s testimony translate into a downgrading of the overall state of evidence for X.

Knobe’s work is controversial, so a different possibility is that, because oracles (in fiction) are respected for their global wisdom, the scenario above is easily conflated with a distinct case in which the oracle also informs us, again in a trustworthy fashion, that gratuitous torture is wrong. We’d be right to respond to this fictional event by rejecting X, but that hardly shows we could ever *in reality* use moral reasoning as a source of evidence in science. If we stipulate that the original oracle has no opinion either way on the wrongness of torture—if we assume only that it is exceptionally good at drawing out the implications of scientific theories—then we get to something closer to what we have in reality: evidence of inconsistencies between scientific theories and moral opinions. But our reaction to this explicitly pared-down oracle, I suggest, is more likely to be one of re-evaluating the state of evidence for our moral theory—all the more so if the oracle tells us *why* and not only *that* X is incompatible with our existing moral outlook. Given more information about the nature of the conflict, X could turn out to *undermine* rather than simply *outweigh* the evidence coming from gratuitous torture’s wrongness. Whether science undermines or merely tends to outweigh the evidence we have for contrary moral judgments is not something we can settle using under-described fictional examples.

I turn now from attempts to explain moral immunity away to attempts to explain it that are compatible with moral realism.

*(d) Moral immunity is just a symptom of moral supervenience*

Most realists assume that moral properties supervene on physical properties. Can they use this supervenience relation to explain science’s immunity to moral counterevidence? Perhaps there is a



connection, but if so it is not immediately obvious. It is one thing to say that there can be no difference in A-properties without a change in B-properties; it is quite another to say that evidence drawn from domain A can have no bearing on the plausibility of claims made in domain B. Moral supervenience is a thesis of the first kind, a metaphysical one, while moral immunity is a thesis of the second kind, an epistemic one. That both theses are plausible in the present case does not make the first explanatory of the second, if only because some instances of the metaphysical thesis are true without that truth carrying over to the parallel epistemic thesis. The mental supervenes on the physical, for example, but no one insists on the immunity of neurophysiological claims to psychological evidence. Quite the opposite, in fact: someone's feeling a sharp pain is treated as evidence that their C-fibres are firing. In contrast, a strong intuition that it is okay to eat fish counts not in the least against the scientific evidence (such as it is) that fish can suffer extreme pain.

So far in this section I have been advising against various accounts of moral immunity that could attract a realist first confronted with the argument from moral immunity. I end with a more promising rejoinder. Indeed, I hold that it is the *only* potentially viable response. This is important because it has consequences (described in §4) that many realists will be unwilling to embrace. It is entirely possible that this final response is but a charitable development of one or more of the responses just dismissed;<sup>17</sup> even so, it is only in this final formulation that the potentially unwelcome consequences become manifest.

*(e) Moral immunity reflects the fact that moral judgments are practical in nature*

Moral reasoning is a form of practical reasoning: it aims to tell us how we ought to act given our circumstances. Telling us what those circumstances are is the function of theoretical (including

---

<sup>17</sup> This is especially true of (d). Frank Jackson, for example, explains ethical supervenience in a way that is strongly suggestive of a fundamental distinction in function between practical and theoretical judgments [2000: 125].

scientific) reasoning. Noting this practical/theoretical distinction does not immediately deal with our puzzle. It fails to tell us *why* I shouldn't take a highly plausible judgment concerning how I ought to act as a sign that circumstances must be such and such and not so and so (given that, were they so and so, it wouldn't be true that I ought to act in this way). So to develop the suspicion that the practical dimension of moral reasoning is key to the immunity phenomenon, let's temporarily rid ourselves of the complexities associated with moral cases and focus instead on a simple case of prudential reasoning.

You want pizza. You believe there is a pizzeria in your nearest town. At this point a trip to the town is looking highly prudent (we'll assume it is healthy pizza). Then, though not vividly, you recall someone telling you something about the pizzeria having closed down. It would be perverse of you to reason that, since travelling to the town is prudentially reasonable, the testimony must have been flawed or else garbled in your memory. The prudential reasonableness of the trip is not just *very weak* as evidence against the faintly remembered testimony; rather, it is no evidence whatsoever.<sup>18</sup>

What explains the evidential impotence in this case? Answers parallel to (a)-(d) are available but unappealing. By far the most attractive answer seems to be that, in a stand-off between a pre-existing prudential judgment and a novel judgment about my circumstances, the prudential judgment is vulnerable because it *lacks any independent support*. In the pizza case, various judgments collide in your mind: your original prudential judgment, T; a new claim about the pizzeria having closed, C; and a collection of relevant background assumptions that include the fundamentals of prudential reasoning. While you arguably have independent sources of evidence for C (in the form of the faintly remembered testimony) and the background assumptions (perhaps an *a priori* source in the case of the fundamentals of prudential reasoning), T has nothing going for it at all aside from being entailed by not-C and the background assumptions. Using T in alliance with the background assumptions as

---

<sup>18</sup> This complete absence of evidential import in the prudential case lends analogical support to the claim, defended under (c) above, that moral reasoning is non-existent and not merely weak as a source of evidence in the scientific domain.

evidence against C would be utterly circular. T can add nothing to the existing evidence for not-C because T is devoid of support that is independent of your assumption that not-C.

This highly intuitive analysis of the prudential case paves the way for a structurally identical explanation of science's moral immunity. That phenomenon was described in §1 as a puzzle over why membership of an inconsistent set fails to constitute evidence against an alleged scientific finding when that set includes, *inter alia*, previously secure moral judgments. What explains this failure is the innocent fact that the set will always contain a moral judgment lacking in independent support. This is clearest in the fish-consumption case. Using the moral permissibility of fish consumption as part of a case against the marine biologists' studies would involve circular reasoning. The evidence for this moral assumption is exhausted by its derivability from, among other things, the claim that fish consumption does not cause significant pain. One claim cannot be used to defend another if all its own plausibility depends essentially on the other claim's being correct.

Notice, too, that immunity would vanish in both prudential and moral cases if the relevant practical judgments were suddenly to acquire independent support. Imagine God appeared and informed you, convincingly, that you should go to the nearby town. Even if he didn't tell you why, his testimony would transform your original prudential judgment into evidence that the pizzeria is still in business. If he also told you that it is morally permissible to consume fish without saying why, the scientific studies would immediately lose their immunity to moral refutation. This observation adds ballast to the thought that missing independent support is what really accounts for the immunity phenomenon.

To summarize, a survey of options suggests that the only plausible realist response to the argument from moral immunity emerges from an analogy with prudential cases. Whenever we have a moral argument against a scientific finding, one of the moral premises will depend for all its plausibility on the argument's conclusion, making the argument essentially circular. In the next section I explore this response further, arguing that it affects different forms of moral realism

differently. Specifically, it calls into question a popular model of moral theorizing, the science-analogy model, and so has knock-on effects for related conceptions of the structure of evidence.

#### **4. Implications of moral immunity for case-driven moral reasoning**

This section revolves around two claims. The first is that the realist response just outlined carries with it an implicit commitment to what I call a *strategic division* of moral judgments into two categories. The second is that this strategic division is incompatible with the science-analogy model of moral reasoning. After defending these claims in turn I draw out their ramifications for realists. It turns out that realists must reject coherentism and bottom-up models of moral reasoning and commit instead to a top-down, derivational model.

Implicit in reply (e) is a strategic division of our moral judgments into two categories, call them *M*- and *m*-class judgments. Within *M* we might tentatively include the judgments that pain and distress have negative value, that promise-breaking is wrong, or that ought implies can; and within *m* the judgments that I may consume fish, that I ought to phone my brother, or that I needn't run for the US Presidency. The distinction between *M*- and *m*-class judgments is built on two contrasts, the first concerning the kinds of evidential support they have and the second concerning the kinds of evidential support they lack. Our *M*-class judgments have an *a priori* source of justification whereas our *m*-class judgments derive support from the *a posteriori* assumptions that, together with relevant *M*-class judgments, entail them (e.g. that fish consumption does not cause extreme pain, that I promised my brother I would phone, or that only the US-born can serve as US President). Just as essential for the explanation of moral immunity are the kinds of justification that judgments in each category lack: *m*-class judgments lack support, *a priori* or otherwise, that is independent of their

derivational justification, while *M*-class judgments cannot draw additional support from our *a posteriori* assumptions about the world.<sup>19</sup>

This strategic division allows a moral realist to have her cake and eat it too. She can hold that moral judgments are, broadly speaking, epistemically justified; but they are never a source of evidence in the scientific domain despite sometimes coming into logical conflict with alleged scientific findings. To illustrate how, let's accept that *Ought implies can* belongs to the *M*-class while *I needn't run for the US Presidency* is in the *m*-class. Both are therefore epistemically justified. The first, like all *M*-class judgments, has an *a priori* source of justification; the second, an *m*-class judgment, is derived from the first plus relevant *a posteriori* knowledge about eligibility requirements and my having been born in the UK. Despite being justified, these two judgments cannot, individually or as a pair, be marshalled into a decent case against a scientific or other non-moral finding. My assumption that ought implies can will only ever logically conflict with scientific findings if it is conjoined with *m*-class judgments. (This is the significance of stipulating that *M*-class judgments cannot have additional support from an *a posteriori* source. If they did, their *a priori* support would itself, potentially, have traction in the scientific domain.) On the other hand, if my not being obliged to run for President is an essential premise in a case against a purported scientific or other non-moral finding, that case will always be circular. It will be circular because an *m*-class judgment will conflict with such a finding only if that finding is inconsistent with one of the *a posteriori* assumptions on which the moral judgment depends for all its support. Suppose the purported finding is that I was born in Oregon. Even if my not being obliged to run is inconsistent with this finding—so, birthplace aside, I'd be the dream candidate—it still fails to threaten the

---

<sup>19</sup> This *M*-/*m*- contrast is comparable to Amartya Sen's between basic and non-basic value judgments [1970: 59], G. A. Cohen's between fact-insensitive and fact-sensitive principles [2008: 231–4], and the distinction implicit in Nick Zangwill's *Because Constraint* between those moral judgments for which it would be inappropriate to ask after the natural facts on which the judgment's truth turns and those where this would be an appropriate question. Cohen use his contrast to argue for the thesis that ultimate principles are fact-insensitive and Zangwill uses his constraint to argue for the *No Independent Access Thesis* [2006: 276]. Both theses resonate with the conclusion of this section.

finding. That's because its plausibility depends essentially on the newly unavailable assumption that I was born in the UK. Of course, if it had an additional, non-derivational source of support, the circularity would vanish; but it doesn't. (This is the significance of stipulating that *m*-class judgments lack independent support.)

Assuming this is a fair elucidation of the only plausible realist response to the immunity phenomenon, moral realists seem committed to making this strategic division. Making it is, however, incompatible with a way of moral theorizing popular with, among others, moral realists: the science-analogy model.

What is this model? Theory confirmation in science (let's assume) involves the empirical testing of hypotheses. The hypothesis is used to generate conditionals of the form, 'If circumstances are such and such, then phenomenon so and so will be observed'—if C then O. Einstein's theory of general relativity, for example, entails that light will be deflected as it passes a large mass. C might be 'The eclipsed sun partially occludes a constellation of stars', with O being 'The visible parts of this constellation appear slightly out of position'. Evidence for or against the theory is achieved by bringing C about (or, as in this example, waiting for it to happen) and using perception to check whether O also occurs. O's being perceived supports the conditional and hence the hypothesis. The moral analogue of this involves generating and testing conditionals in order to confirm or overturn a moral generalisation. These conditionals might have the form, 'If circumstances are such and such, my obligations will be so and so'—if C, then I should A (or should not A, or am permitted to A, etc.). A moral theorist then envisages herself in C, recalls having been in C, or even puts herself in C; she then attempts to sense, feel, or intuit the moral facts of the case—she 'consults her moral intuitions'—and finds herself thinking that she should A or else that she needn't or mustn't. Intuitive judgments are the analogue of perceptual judgments, and depending on whether they accord with the conditional, they either do or do not support the moral theory that entails the conditional.

To see how this familiar model clashes with the strategic division, notice that in science perceptual data has to supply O with support that is independent of O's derivability from the

conjunction of C plus the hypothesis being tested. The moral methodology appears to require the same. To support a moral theory, my intuitive responses to scenarios must do more than concur with it; they must have independent support, support that goes beyond their derivability from the theory plus the real or imagined circumstances. In a paradigmatic use of this methodology, his Experience Machine example, Robert Nozick warns of this need for our intuitions about scenarios to be independent of the theory they are testing:

[The] example must be looked at on its own; to answer the question by filtering it through a fixed view . . . would lose [us] the opportunity to test that view *independently*. One way to determine if a view is inadequate is to check its consequences in particular cases . . . but if someone always decided what the result should be in any case by *applying* the given view itself, this would preclude discovering it did not correctly fit the case.

[Nozick 1989: 105. First emphasis added]

Herein lies the problem. Such independent support is explicitly ruled out by the strategic division. Contingently situated judgments with the form 'I should A' belong to the *m*-class of moral judgments. The only support they can have consists in their derivability from judgments in the *M*-class plus relevant non-moral judgments. If we supposed them to have independent support, realists would need some other way of explaining science's moral immunity.

Things are hardly that simple. Realists who like the science-analogy model have two potentially strong rejoinders available. Both turn on differences they may be willing to accept between scientific and moral reasoning. The first difference is that scenarios tend to be fictional in ethics but not in science. The second is that, while perception in science is meant to supply us with evidence for the scientific theory rather help us identify our prior scientific commitments, consulting our moral intuitions could be construed as an aid to identifying our broader moral commitments rather than as a source of evidence for those commitments. Discussing these two differences,

especially the second, will help us to appreciate which kinds of realism are called into question by the moral immunity phenomenon and which kinds are not.

The fictionality of many scenarios does not eliminate the conflict between case-based moral reasoning and the strategic division. To see why, let's distinguish between  $m_R$ - and  $m_I$ -class judgments, according to whether the situation prompting the judgment is real or imaginary. This allows us to clarify the first rejoinder but also to see why it fails. Suppose a science-analogy model purist insists on using only  $m_R$ -class judgments for moral reasoning. Intuitions generated by real-life morally charged situations—hostage takings, marital betrayals, etc.—can be used to support moral hypotheses, says this purist, but intuitions arising from fictional scenarios cannot. This view is straightforwardly incompatible with making the strategic division to explain science's moral immunity. Few supporters of the science-analogy model are purists in this fashion. They can therefore maintain that our  $m_I$ -class judgments have independent support (this being good enough for case-driven moral reasoning) while our  $m_R$ -class judgments lack independent support (this being good enough for a version of the strategic division that can explain science's moral immunity).

This first rejoinder trips up on the observation that if our  $m_I$ -class judgments have independent support, so too must our  $m_R$ -class judgments. Overstressing the fictionality of many cases used in moral theorizing means we miss something important: that our moral judgment develops as we live through or seek out and learn from morally charged real events—visiting an abattoir, say, or giving up a child for adoption. Even fictional cases work only in so far as they produce in us the reaction that would have been produced had the events occurred 'for real'. The same cognitive structures are at work in each case—running offline or on, so to speak—and they deliver the same normative output. In view of this, a realist supporter of the science-analogy model cannot save the strategic division by drawing a contrast between  $m_I$ - and  $m_R$ -class judgments and assuming that only the former have independent support. And since they cannot make the strategic division, it is a consequence of their view—though very much an unintended one—that moral judgments can be cited in evidence for non-moral, and in particular scientific, judgments.



Or rather, this is a consequence of their view so long as they do not avail themselves of the second rejoinder: that our responses to morally charged cases are not *meant* to support the moral conditionals they are consistent with, or the broader moral beliefs from which those conditionals are derived. Our responses could be construed instead as a sign that we endorse these conditionals or that we hold these broader moral beliefs. Any evidence that we are *right* to endorse or believe would come, not from what we sense, feel, or intuit about real or imagined cases, but from some other source. Suppose you don't know how you feel about the acts/omissions contrast, for example. In so far as you have an underlying stance on it at all, you can come to appreciate that stance by toying with different formulations and deriving the implications of each for various (real or imaginary) scenarios before reflecting on your response to those scenarios as expressed in *m*-class judgments. At no point need you treat this response as evidence *for the stance* as opposed to support for the thesis that this is a *correct formulation* of it.<sup>20</sup>

My argument does not threaten this position, but it is an exaggeration to say that all who appeal to cases in moral reasoning do so hoping only to find out what we believe, not whether we are right to do so. Notice, too, how odd it is to suppose that we have beliefs that are justified *a priori* despite our being so ignorant of their content that we need to consider real and imaginary scenarios in order to derive *a posteriori* evidence for their having that content. Odd, but not incoherent, not without possible parallels,<sup>21</sup> and not susceptible to refutation by anything I have said here. My argument is directed instead at realists who believe the evidence we have for situated moral judgments goes beyond top-down derivation from an *a priori M*-class. That includes two overlapping categories of

---

<sup>20</sup> Jeff McMahan [forthcoming] takes this approach, which is compatible with *some* upwards support: thinking about cases could reveal that our underlying moral beliefs are at least consistent. Were an inconsistency within them to emerge, however, we'd have no bottom-up reason to resolve it one way rather than another.

<sup>21</sup> Mathematical knowledge offer a possible analogue. McMahan [forthcoming] draws a comparison with our unconscious knowledge of syntax.

moral philosopher: intuitionists (in a specific sense) and some users of reflective-equilibrium methodology.

Moral intuitionists, in one sense of the phrase, see moral intuition as like, or even as an instance of, perception through the external senses. They either hold that we can literally see moral facts, follow Hutcheson in hypothesising a distinct moral sense, or treat emotion as a form of moral perception. Intuitionists in this sense have an obvious difficulty here. Whatever else moral perception is, to be worthy of the label it should be an independent source of evidence for what is perceived.<sup>22</sup> Moral intuitionists in a different sense, meaning those who follow W. D. Ross in treating intuition as an intellectual rather than a sensory or quasi-sensory capacity, slot more easily into the category of those I just said I was not arguing against. They hold that our judgments about cases are but symptoms of a commitment to intuitive general principles. The present objection merely forces them to be explicit about the lack of independent support provided to these general commitments by reflection on particular cases.<sup>23</sup>

‘Reflective equilibrium’ can pick out either a method of moral reasoning or a theory of justification often affiliated to that method. The method involves reconciling considered judgments about real or imagined cases with more general judgments, giving up judgments of either kind in order to maximize coherence. Though compatible with a foundationalist account of moral evidence, this method is more naturally associated with a coherentist account in which any support attaching to individual members of a set of judgments is enhanced by that set’s coherence [Daniels 1979; Sayre-McCord 1996]. This coherentism is incompatible with making the strategic division. Coherentists

---

<sup>22</sup> I am treating this as a litmus test for the interest of the potentially nebulous or metaphorical claim that we can visually perceive the wrongness in an action (e.g. McDowell [1985]; McGrath [2004]; Chappell [2008]; Audi [2010]). Independent support is also the apparent motivation for those realists who regard emotion as a perceptual or quasi-perceptual moral guide (e.g. Little [1995]; Döring [2003]; Roeser [2011]; and Tappolet [2000], though see note 3).

<sup>23</sup> They are not always clear on this. Michael Huemer, for example, a neo-Rossian, classifies what I call  $m_R$ -judgments as non-intuitive [2005: 102], but a page later describes  $m_I$ -judgments as intuitive. Ross himself says that reflection on cases ‘is the only guide we have to our duty’ [1930: 42], but this is compatible with his adopting the position I have said I am not criticising here.

therefore face the charge that an alliance between the undifferentiated totality of our moral judgments could be used to do science. The strategic division is also unavailable to those who combine reflective-equilibrium methodology with a kind of ‘bottom-up’ foundationalism in which the foundation set consists of moral judgments about real or imaginary cases rather than judgments of general moral principle. Because the foundation provides bottom-up support to general principles, general principles and particular judgments could, as with coherentism, act in unison to refute scientific studies that conflict with a reasoned moral outlook.<sup>24</sup> The only users of reflective-equilibrium methodology with nothing to worry about hold that a foundational category (the *M*-class) of moral principles is not in the least reliant for support on the (*m*-class) judgments we make in response to real or imagined scenarios, and conceive of the method merely as a useful way to ascertain which principles are in fact foundational.

A realist unhappy at being boxed in in this way could respond that the independent support possessed by our particular, situated moral judgments is defeasible rather than non-existent. Immunity, on this picture, arises because this independent support is *undermined* (and not merely outweighed<sup>25</sup>) in any conflict with science; but so long as it remains undefeated this support can be drawn on in the usual way for the purposes of moral reasoning. A parallel can be drawn with perceptual illusion. Consider a sighted but naïve person who, for the first time, sees a stick immersed in water and takes it to be bent. His perceptual justification is independent in the sense of not being

---

<sup>24</sup> Uniting the views I am arguing against here is a commitment to particular moral judgments having independent (i.e. non-derivational) support. Whether this independent support is taken to carry the rest (‘bottom-up foundationalism’) or is merely added to the mix (‘coherentism’) is unimportant. The challenge facing anyone with this commitment is to say why this independent support can lend credence to our moral principles without being useable in the non-moral domain. Some possible attempts to meet the challenge are dealt with elsewhere in the paper, e.g. that cases are often fictional (see earlier in this section); that cases have independent support that is weak relative to, and hence outweighed by, evidence in the scientific domain (see (c) in §3 and note 18); or that this independent support is undermined and not merely outweighed by empirical discoveries (see the remaining two paragraphs in this section).

<sup>25</sup> For reasons to think that scientific evidence does not merely outweigh moral evidence, see (c) in §3 and note 18.

derived: he just (thinks he) sees a bend in the stick. This justification is undermined by an explanation of the stick's appearance in terms of refraction. Saying it is undermined is not the same as saying that his judgment was never justified, but nor is it the same as allowing that the appearance of a bend somehow still weighs, just a little, against a physicist's account of what is going. It doesn't count at all because the physicist's account incorporates an explanation of the appearance of a bend.

Notice that even top-down theorists allow that the justification possessed by situated moral judgments is undermined by contrary scientific findings. According to them, though, the undermined justification is derivational—rather than independent as the current response requires. The key question, then, is whether, when a scientific finding undermines a particular moral judgment, it is undermining an independent justification (as with perceptual illusion) or just a derivational justification. To answer this, imagine a world in which our situated normative judgments are very much independent: they are prompted by a distinctive Nietzschean whiff. We can literally smell the moral facts.<sup>26</sup> In this world it *would* be reasonable for us to draw on moral evidence in the scientific domain. For a contrary scientific finding to undermine (rather than outweigh) this independent olfactory evidence would take rather more than saying 'You were wrong about the non-moral facts on which the truth of your moral judgment depended'. It would require an explanation of how our misleading moral sense of smell arose. The fact that nothing like this is ever needed in reality—the fact that *all* we feel we need to say is that the judger has made a flawed non-moral assumption—is a sign, not only that there is no Nietzschean whiff (which we knew already), but also that it has no surrogate in the form of defeasible but independent evidence for situated moral judgments. The evidence we possess for particular moral judgments can be undermined in the simple way that only a top-down theorist predicts.

---

<sup>26</sup> Cf. 'Bad air! Bad air! That something failed comes near to me, that I have to smell the bowels of a failed soul!' [Nietzsche 1887: 25].

## 5. Conclusion

Moral anti-realists can readily account for science's apparent immunity to moral refutation. Options for moral realists are more limited. They must deny that particular moral judgments about either fictional scenarios or real events have independent evidential support, i.e. support going beyond their derivability from general principles and the non-moral facts of the case. A top-down, theory-first foundationalist moral epistemology, in which *a priori* moral principles support but are not supported by judgments about particular cases, offers the only likely route to a realist treatment of science's apparent moral immunity.<sup>27</sup>

*The Open University, UK*

## REFERENCES

- Audi, R. 2010. Moral Perception and Moral Knowledge, *Proceedings of the Aristotelian Society: Supplementary Volume* 84/1: 79–97.
- Braithwaite, V. 2010. *Do Fish Feel Pain?*, Oxford: Oxford University Press.
- Chappell, T. 2008. Moral perception, *Philosophy* 83/4: 421–37.
- Chrisman, M. 2012. Epistemic Expressivism, *Philosophy Compass* 7/2: 118–26.
- Cohen, G. A. 2008. *Rescuing Justice and Equality*, Cambridge, Mass.: Harvard University Press.
- Daniels, N. 1979. Wide Reflective Equilibrium and Theory Acceptance in Ethics, *Journal of Philosophy* 76/5: 256–82.
- Döring, S. A. 2003. Explaining Action by Emotion, *Philosophical Quarterly* 53/211: 214–30.
- Doris, J. M. 2002. *Lack of Character: Personality and Moral Behavior*, Cambridge, UK: Cambridge University Press.
- Gibbard, A. 2003. *Thinking How to Live*, Cambridge, Mass.: Harvard University Press.

---

<sup>27</sup> For feedback on earlier versions I'm grateful to Chris Belshaw, Bill Cameron, Timothy Chappell, Chris Hookway, Adam Morton, Lindsey Porter, Carolyn Price, and audiences at the *Canadian Philosophical Association* in Waterloo and the *British Society for Ethical Theory* in Stirling. Anonymous comments from two referees and an Associate Editor for this journal were challenging and exceptionally helpful.

- Harman, G. 1977. *The Nature of Morality*, Oxford: Oxford University Press.
- Harman, G. 1996. Moral relativism, in *Moral Relativism and Moral Objectivity*, ed. G. Harman and J. J. Thomson, Oxford: Blackwell: 1–64.
- Harman, G. 1999. Moral philosophy meets social psychology: Virtue ethics and the fundamental attribution error, *Proceedings of the Aristotelian Society* 99/3: 315–31.
- Heller, M. 2003. The Immorality of Modal Realism, or: How I Learned to Stop Worrying and Let the Children Drown, *Philosophical Studies* 114/1–2: 1–22.
- Herrnstein, R. J. and C. Murray 1994. *The Bell Curve*, New York: Free Press.
- Huemer, M. 2005. *Ethical intuitionism*, New York: Palgrave Macmillan.
- Jackson, F. 2000. *From metaphysics to ethics: a defence of conceptual analysis*, Oxford: Oxford University Press.
- Joyce, R. 2001. *The Myth of Morality*, Cambridge: Cambridge University Press.
- Kitcher, P. 1985. *Vaulting Ambition: Sociobiology and the Quest for Human Nature*, Cambridge, Mass.: MIT Press.
- Knobe, J. 2010. Person as scientist, person as moralist, *Behavioral and Brain Sciences* 33/4: 315–29.
- Little, M. O. 1995. Seeing and Caring: The Role of Affect in Feminist Moral Epistemology, *Hypatia* 10/3: 117–37.
- Mackie, J. L. 1977. *Ethics: Inventing Right and Wrong*, Harmondsworth: Penguin.
- McDowell, J. 1985. Values and Secondary Qualities, in *Morality and Objectivity*, ed. T. Honderich, London: Routledge and Kegan Paul: 110–29.
- McGrath, S. 2004. Moral Knowledge by Perception, *Philosophical Perspectives* 18: 209–28.
- McMahan, J. forthcoming. Moral intuition, in *The Blackwell Guide to Ethical Theory, Second Edition*, ed. H. LaFollette and I. Persson, New York: Wiley-Blackwell.
- Miller, A. 2003. *An introduction to contemporary metaethics*, Cambridge, UK: Polity Press.
- Newby, R. G. and D. E. Newby 1995. The Bell Curve: Another Chapter in the Continuing Political Economy of Racism, *American Behavioral Scientist* 39/1: 12–24.
- Nietzsche, F. 1887 (2007). *On the Genealogy of Morality*, C. Diethe, trans., and K. Ansell-Pearson, ed. Cambridge, UK: Cambridge University Press.
- Nozick, R. 1989. *The Examined Life*, New York: Simon and Schuster.
- Prinz, J. 2009. *The Emotional Construction of Morals*, Oxford: Oxford University Press.
- Roeser, S. 2011. *Moral Emotions and Intuitions*, Basingstoke: Palgrave Macmillan.
- Ross, W. D. 1930. *The Right and the Good*, Oxford: Clarendon.

- Sayre-McCord, G. 1988. Moral Theory and Explanatory Impotence, *Midwest Studies in Philosophy* 12: 433–57.
- Sayre-McCord, G. 1996. Coherentist epistemology and moral theory, in *Moral knowledge?*, ed. W. Sinnott-Armstrong and M. Timmons, Oxford: Oxford University Press: 137–89.
- Schroeder, M. 2010. *Noncognitivism in ethics*, London: Routledge.
- Sen, A. 1970. *Collective Choice and Social Welfare*, London: Oliver and Boyd.
- Sinnott-Armstrong, W. 2006. *Moral Skepticisms*, Oxford: Oxford University Press.
- Soon, C. S., M. Brass, et al. 2008. Unconscious determinants of free decisions in the human brain, *Nature Neuroscience* 11/5: 543–45.
- Sturgeon, N. L. 1988. Moral Explanations, in *Essays on Moral Realism*, ed. G. Sayre-McCord, Ithaca: Cornell University Press: 229–55.
- Sturgeon, N. L. 2006. Ethical Naturalism, in *The Oxford Handbook of Ethical Theory*, ed. D. Copp, New York: Oxford University Press: 91–121.
- Tappolet, C. 2000. *Émotions et valeurs*, Paris: Presses Universitaires de France.
- Zangwill, N. 2006. Moral Epistemology and the Because Constraint, in *Contemporary Debates in Moral Theory*, ed. J. Dreier, Malden, Mass.: Blackwell: 263–81.