BELIEF PILLS AND THE POSSIBILITY OF MORAL EPISTEMOLOGY

NEIL SINCLAIR

1. Introduction

How should we respond to empirical discoveries about the evolutionary origins of our moral sensibilities? Broadly speaking, there are two types of answer. Optimists hold that substantive moral claims, or theories of the nature of moral practice, or our continued engagement in moral practice, are *vindicated* by an evolutionary account of morality (for the first, see Casebeer 2003; Mikhail 2011; for the second see Harms 2000; for the third see Campbell 1996). Pessimists take moral practice – or some of its features – to be *undermined* or *debunked* by evolutionary understanding. Debunking arguments begin from a premise concerning the evolution of some aspect of moral practice, and can be distinguished by their target conclusion. Some urge that substantive moral theories are false or unjustified (Greene 2008; Singer 2005). Others target metaethical theories. Street, for example, targets moral realism, according to which 'there are at least some evaluative facts…that hold independently of all our evaluative attitudes' (2006: 110). On Street's view, realists are 'unable to accommodate the fact that Darwinian forces have deeply influenced the content of human values' (2006: 109). Joyce (2014) calls this 'theory debunking'.

Two further types of debunking argument target the moral judgements of ordinary folk. The first claims that we can move from evolutionary premises to a conclusion regarding the *falsity* of such judgements. Joyce (2014) calls this 'truth debunking' (see Ruse 1986: 254). The second aims for a conclusion regarding the *epistemic status* of moral judgements: that they are *unjustified*, *not knowledge* or otherwise epistemically deficient. Joyce calls this

'justification debunking'. Such debunking itself comes in many flavours. For example, some arguments allow that some moral judgements might be initially justified and take evolutionary discoveries to undermine this justification. Others allow only that some moral judgements appear justified and urge that evolutionary considerations show these appearances to be deceptive (White 2010: 575). An orthogonal distinction concerns whether the supposed debunking proceeds from the claim that (some) moral judgements rest on ignorance or false belief about the nature of morality or simply from the claim that moral judgements fail to meet some necessary condition for justification (Joyce 2006: 180; Lillehammer 2003a: 571-3). In addition, truth and justification debunking arguments can be local or global, depending on whether they target some or all moral judgements.

Joyce is perhaps the foremost proponent of a global justification debunking argument for moral judgements. He writes:

> We have an empirically confirmed theory about where our moral judgments come from…This theory doesn't state or imply that they are true, it doesn't have as a background assumption that they are true, and, importantly, their truth is not surreptitiously buried in the theory by virtue of any form of moral naturalism. This amounts to the discovery that our moral beliefs are the products of a process that is entirely independent of their truth, which forces the recognition that we have no grounds one way or another for maintaining these beliefs. (2006: 211)

According to Joyce, this parallels the following argument:

> Suppose that there were a pill that makes you believe that Napoleon won Waterloo, and another that makes you believe that he lost…[I]magine that you are proceeding through life happily believing that Napoleon lost Waterloo…and then you discover that at some point in your past someone slipped you a

2

"Napoleon lost Waterloo" belief pill…Should this undermine your faith in your belief that Napoleon lost Waterloo? Of course it should. It doesn't show that your belief is *false*…but this knowledge is certainly sufficient to place this belief on the dubious list. (2006: 179)

In this paper I evaluate global justification debunking arguments such as Joyce's. I understand such arguments as aiming to establish their conclusion independently of truth-debunking arguments. So understood, I argue they are dialectically ineffective.

One point is worth mentioning at the outset. Many take justification debunking arguments in morality to deploy substantive metaethical assumptions of realism or cognitivism. According to the former, the epistemic credentials of moral judgements are threatened on the assumptions that (a) those judgements are cognitive and (b) that the realm of moral facts which such judgements seek to cognise exists independently of the minds and languages of human beings (Clarke-Doane 2012: 315-17; Kahane 2011: 112; Street 2006: 110.). According to the latter the epistemic credentials of moral judgements are threatened simply on the assumption of cognitivism (Handfield 2016; Joyce 2016: 126-9; Mason 2010: 775). These assumptions have done much to shape discussion of such arguments. For example, those holding the former view tend to present these arguments as parts of arguments against moral realism – the thought being that if a global moral scepticism results from an assumption of moral realism, we should reject that assumption (Street 2006: 109).

But are these assumptions necessary for debunking arguments? Arguably not. Such arguments target the idea that moral judgements are epistemically justified. The minimal assumption required, therefore, is simply that moral judgements are the sorts of thing that can be, or might have been, epistemically justified. But it is well-known that sophisticated versions of non-cognitivism hold that moral judgements can be so justified. Quasi-realists

hold that there are available notions of truth and justification that can warrantedly apply to moral judgements, even though such judgements are, at root, non-representational (Blackburn 1993). Since such views assume that there are conditions on epistemic justification which some moral judgements meet, they are potentially vulnerable to justification debunking. Hence the following discussion assumes only that moral judgements are the sorts of things that can, or might have been, epistemically justified (Lillehammer 2010: 361).

2. Evolutionary Explanations of Moral Judgements

Global justification debunking arguments such as Joyce's start from the premise that moral practice can be given an evolutionary explanation, that is, an explanation in terms of that practice's ability to increase the relative reproductive fitness of some ancestral population. But what, precisely, do such explanations *explain*? One possible explanandum is our possession of moral concepts – generating the 'capacity hypothesis'. Another is our tendency to deploy moral concepts in particular ways – generating the 'tendency hypothesis'. Joyce defends the former (2006: 108-42), Street the latter (2006: 114-15; see FitzPatrick 2015 for the distinction). But note that neither is a *complete* explanation of all moral judgements. The capacity hypothesis by itself does not explain any moral judgements (only our capacity to make them); the tendency hypothesis only explains (the tendency to make) *some* moral judgements. This incompleteness is one source of vulnerability in debunking arguments.

It is worth contrasting evolutionary explanations of *moral* practice with evolutionary explanations of other practices (Fraser 2014: 463; Joyce 2006: 182-4). Consider, for example, *dry-goods* practice: the practice of thinking and talking about medium-sized dry-goods, such as boulders. The capacity to make dry-goods judgements can be given an evolutionary explanation. But there is crucial asymmetry with the moral case. Whereas the capacity to

make dry-goods judgements plausibly evolved in part as a mechanism to track mind-independent facts about the location, number and movements of independently existing dry-goods, the capacity to make moral judgements evolved (according to Joyce) as a mechanism to promote prosocial behaviour and help sustain co-operative social frameworks. Thus whereas the capacity to make dry-goods judgements was selected partly on the basis of its possession enabling the possessor to track mind-independent facts, the capacity to make moral judgements was (according to Joyce) selected purely on the basis of its possession enabling (or making more likely) certain behaviours. Thus it seems that whereas the dry-goods faculty was selected partly on the basis of its upstream tracking abilities, the moral faculty was selected purely on the basis of its downstream motivational effects. (The same point applies, *mutatis mutandis*, for the tendency hypothesis.) This highlights the fact that what matters in assessing justification debunking arguments is not whether our moral faculties and judgements can be given *some* kind of evolutionary explanation but whether the *type* of explanation offered undermines justification (Joyce 2006: 212).

3. A Generic Debunking Argument

Gathering the foregoing threads together, consider the following argument:

(1)   All (actual, human) moral judgements can be given plausible evolutionary explanations.

(2)   The process referred to in these explanations does not track moral truth.

Hence

(C)   All (actual, human) moral judgements are unjustified.

This generic argument is not formally valid. Turning it into a valid argument is the task for the next three sections.

Consider (2) – the 'epistemic premise' (Kahane 2011: 106). What does 'does not track moral truth' mean? Synonyms include: 'an off-track process' (*ibid.*), 'tracking failure' (Lillehammer 2010: 365), 'truth-mooting genealogy' (Mason 2010: 773). Different answers to this question generate different debunking arguments. But there seems to be considerable agreement on the following, minimal, account: the process referred to in these explanations does not track moral truth insofar as these explanations do not invoke or assume the truth of the judgements concerned (Fraser 2014: 459; Joyce 2006: 183-4; Street 2006: 155).

(1) is the 'causal premise'. Neither the capacity nor tendency hypothesis provides a complete explanation of all moral judgements. Hence (1) does not follow from either. Given the capacity hypothesis particular moral judgements will be partly the result of the evolutionary history of the moral capacity, partly the result of more proximal environmental stimuli. There is thus the potential for debunking arguments based on this hypothesis to fall at the first hurdle (Toner 2010). I will not explore the details of this issue here. Instead I note that debunkers usually overcome this problem by assuming that whatever non-evolutionary factors are involved in the explanation of specific moral judgements, these are also truth-mooting with respect to moral truth (Kahane 2011: 106; Mason 2010: 774). Joyce, for example, responds to an analogous worry about his belief-pill argument by modifying the scenario to that of a 'concept pill' – one which 'makes you form beliefs about Napoleon in general'. He suggests a case in which, knowing one has been slipped such a pill, one's Napoleon beliefs are determined 'randomly' by 'certain environmental triggers' (presumably not involving Napoleon himself). In such a case, Joyce suggests, the resulting beliefs are still undermined (2006: 181).

To incorporate this into the argument, call an 'evolutionary+' explanation one that involves the capacity or tendency hypothesis together with additional morally truth-mooting factors. Together with the minimal account of 'does not track moral truth', this generates the following:

(1*)   All (actual, human) moral judgements can be given plausible evolutionary+ explanations.

(2*)   These explanations do not invoke or assume the truth of such judgements.

Hence

(3)   All (actual, human) moral judgements can be plausibly explained without invoking or assuming their truth.

Hence

(C)   All (actual, human) moral judgements are unjustified.

(3) may also be supported by non-evolutionary premises, concerning, for example, sociological explanations of moral judgements. For now though, my concern is validating the move from (3) to (C).

4. Redundancy

What additional premises might bridge this gap? Consider:

(3)   All (actual, human) moral judgements can be plausibly explained without invoking or assuming their truth.

(4a)   If a judgement can be plausibly explained without invoking or assuming its truth then it is not justified.

Hence

(C)   All (actual, human) moral judgements are unjustified.[1]

This is valid and plausible. After all, what reason could there be to believe in a set of truths if reference to those truths was not required to explain our judgements about them? (Lillehammer 2003a: 577; Nichols 2014: 729.) But despite this, *Redundancy* is unsuccessful. There are at least two problems with (4a).

First, the possibility of reduction. Suppose that property A (such as *being water*) is reducible to property B (such as *being H_2O*). This entails, among other things, that A is identical to B. Suppose a particular judgement concerning A (e.g. Jo's belief that the water in the pipes has frozen) can be explained by citing an instance of B (e.g. 'Jo believes that the water in the pipes has frozen because the H_2O in the pipes is frozen'). Then it seems that the judgement can be explained without invoking or assuming its truth, but can still be justified. The general point is that an explanation of a judgement in terms of a level of properties not invoked or assumed by the judgement is consistent with (but does not require) those properties being the reductive basis of the properties ascribed by the judgement, and hence consistent with the judgement being responsive to the properties it ascribes (Lillehammer 2003a: 577; Quinn 1986: 537-49).

Of course, this problem with (4a) may not be a problem if reduction is unavailable. Joyce (2006: 190-209) argues that this is so in the moral case. But on reflection the same problem arises with weaker metaphysical relations. Suppose that Jo's judgement that x is C

---

[1] Throughout I understand debunking arguments to utilise externalist conditions on justification, such as the antecedent of (4a). But my criticisms of debunking arguments apply equally to internalist versions.

is explained by x's being D. This explanation does not invoke or assume the truth that x is C. But it is perfectly consistent with this explanation that C supervenes on D and hence that Jo's judgement is counterfactually dependent on x's possession of C. Such sensitivity is in turn consistent with (and suggestive of) the judgement being justified. Again the point is general: an explanation of a judgement in terms of one level of properties is consistent with (but does not require) those properties being metaphysically related to the property ascribed in the judgement in such a way as to render the judgement justified.

Another reply to the reduction problem urges that the relevant judgements are explained *neither* by the properties referred to by the judgement itself, *nor* by any plausible reductive or supervenience base for those properties. But though this response delivers a more plausible version of (4a) it renders (3) implausible. To secure validity, (3) would need to be modified to the claim that all moral judgements can be plausibly explained without invoking or assuming their truth, or any reductive or subvening base for that truth. But such a claim is unsupported by our evolutionary hypotheses. Because such hypotheses are relevantly *incomplete*, it is consistent with them to suppose that some moral judgements are explained in part by invoking a plausible reductive or subvening base for the properties they refer to.

The second problem with (4a) is the existence of counterexamples. I judge that the sun will rise tomorrow and that all men are mortal, and both judgements are justified. But in neither case does the relevant truth play a role in explaining my judgement (Dancy 1985: 34).

5. Further Arguments

Redundancy employs the minimal sense of tracking failure noted above. In light of its failure, the would-be debunker may argue that there is a further – robust – sense in which moral

judgements are 'off-track'. This sense may be supported either by the minimal sense itself or by the considerations which support it. This section outlines one argument of this type and considers prospects for others.

5.1. Sensitivity

The thought behind the sensitivity argument is that given evolutionary explanations such as the capacity hypothesis, we would make the same moral judgements that we have now, even if the moral truths had been different (Clarke-Doane 2012; Joyce 2016: 129; Lillehammer 2010: 365; White 2010: 581; Wielenberg 2010: 16-19).[2] More formally:

(3)   All (actual, human) moral judgements can be plausibly explained without invoking or assuming their truth.

Hence

(4b)   All (actual, human) moral judgements are *insensitive* to moral truth.

(5b)   If a judgement is not sensitive to the truth of its content then it is not justified.

Hence

(C)   All (actual, human) moral judgements are unjustified.

To say that a judgement is *sensitive* to the truth of its content is to say that were that content false, the judgement would not be made. (Strictly speaking, sensitivity needs to be relativized to a method. See Nozick 1981: 179.)

One problem here is that (4b) is unsupported. It follows neither directly from (3) nor from the evolutionary premises (1*) and (2*) and it is independently implausible (Clarke-

---

[2] Bedke (2014) discusses a related condition – that of 'obviousness' – but his resulting debunking argument targets only specific types of cognitivism.

Doane 2012; Wielenberg 2010: 454-6). All three can be demonstrated with the same example. Consider my judgement that Jones' torturing is wrong, which I possess partly because Jones' torturing causes unwanted agony. Grant that we can explain this judgement without invoking or assuming its truth. For example we can say that I make this judgement partly because evolution has given me a moral capacity and partly because of the proximal fact that Jones' torturing causes unwanted agony. Still, my judgement can be sensitive to moral truth so long as the wrongness of Jones' act supervenes on (or is reducible to) its causing unwanted agony. And that seems to be the case here. Were Jones' act not to be wrong, it wouldn't be an act of causing unwanted agony, and hence I wouldn't judge it wrong. In this case, my moral judgement is sensitive to moral truth even though it can be explained without invoking that truth, and even though it can be given an evolutionary+ explanation which does not invoke or assume that truth. Thus (4b) follows neither from (3) nor from (1*) and (2*), nor is it independently plausible.

There is something unsettling about this reply. It relies on two substantial moral assumptions: that Jones' torturing is wrong and that its wrongness supervenes on its causing unwanted agony. But, in the current context, this is unproblematic. For the burden of proof here is with the would-be debunker. It is sufficient, to resist the argument, to show that (4b) does not follow from the premises offered in its support. To show this, all that is necessary is to show that there is a possible situation where (3) holds and (4b) does not. My judgement concerning Jones is such a case. There is a difference, then, between assuming a moral truth for the sake of explaining moral judgements, and assuming such a truth for the sake of testing whether an evolutionary+ explanation of those judgements shows them to be insensitive. The former may be question-begging, but the latter is not (see FitzPatrick 2015: 897).

A deeper problem for *Sensitivity* concerns (5b). Sensitivity does not seem necessary for justification, since false judgements can be justified, but are insensitive (Dancy 1985: 39;

11

White 2010: 580). Further, *Sensitivity* inherits well-known problems about the assessment of the counterfactual conditionals it employs (Joyce 2016: 130-2).

5.2. Other Arguments

Sensitivity holds that moral judgements exemplify a robust sense of 'tracking failure'. It claims that this robust sense is supported by the minimal sense and that the absence of such tracking failure is a necessary condition for justification. The argument fails because the robust sense of tracking failure does not follow from the minimal sense and because the putative necessary condition is implausible.

There are several further robust senses of 'tracking failure' which can be used to construct similar arguments. For instance, one might hold that minimal tracking failure shows moral judgements to be *not safe*, *contingent*, *probably false,* or *unreliable*. But all of these arguments are vulnerable in a similar way. First these more robust senses of tracking failure don't follow from the minimal sense. Second, the putative necessary conditions are contentious (Ichikawa & Steup 2012). Most of these arguments have received extensive discussion elsewhere. (For arguments from safety see Clarke-Doane 2012; Handfield 2016; Joyce 2016; for contingency see Lillehammer 2010; Setiya 2012: 91-2; Wielenberg 2010; for 'probably false' see Brosnan 2011; Street 2006; for unreliability see Clarke-Doane 2016; Wielenberg 2010). So here I will examine a different type of argument, suggested by Joyce's pill cases.

6. Lack of Method

One natural response to my dismissal of *Sensitivity* is as follows. It *might be* that some moral judgements are sensitive and it *might* be that one can provide particular examples of moral

judgements which satisfy this condition. But this is not to provide any general account of the mechanisms by which humans form moral judgements such that those judgements might end up being in good epistemic standing. In particular, it is not to provide any general account of how the mechanisms for forming moral judgements which we actually employ might constitute ways of getting into contact with moral truth. According to this argument, evolutionary explanations of our moral judgements debunk their justification insofar as they emphasise the *absence* of a positive epistemological story for those judgements (Joyce 2006: 135; Setiya 2012: 96-115).

This argument might be expressed thus:

(3)   All (actual, human) moral judgements can be plausibly explained without invoking or assuming their truth.

(4c)   There is no plausible general account of how moral truths could help explain (actual, human) moral judgements.

(5c)   If there is no plausible general account of how a set of truths could help explain judgements concerning them and we can plausibly explain those judgements without invoking or assuming their truth, then those judgements are unjustified.

Hence

(C)   All (actual, human) moral judgements are unjustified.

The antecedent of (5c) is not met in the case of dry-goods judgements, since even in a time before human beings studied their own perceptual systems, they did not have an explanation of their dry-goods judgements that did not invoke their truth. But this condition *is* seemingly met in the case of moral judgements. For not only do we lack general account of the

mechanism where by our moral judgements might hook up with moral truth, we also possess an (evolutionary+) account of those judgements that nowhere invokes or assumes their truth.

How might the opponent respond? I think they can reasonably reject (4c). In the next subsection I consider one unsuccessful way of doing this. In the following subsection, I outline a better way.

6.1. Pure Epistemology

The field of moral epistemology is not empty. There are many extant accounts of the putative mechanisms by which our moral beliefs are epistemically justified. Many of these accounts are motivated independently of the results of debates about the ontological or substantive nature of moral truth. As such they might be labelled 'pure epistemologies' (Joyce 2006: 211). According to the current reply, the availability of these epistemologies allows reasonable rejection of (4c).

Three popular 'pure' moral epistemologies are conservativism, coherentism and foundationalism. According to conservativism moral judgements are *prima facie* justified in virtue of being held (Joyce 2006: 216). According to coherentism moral judgements are justified insofar as they are part of a coherent system of judgements (Brink 1989: 100-143). According to foundationalism moral judgements are justified insofar as they are self-evident (Ross 1930) or suitably related to foundational 'seeming states' (Huemer 2005).

Unfortunately for the anti-debunker, none of these epistemologies provide, by themselves, sufficient reason to reject (4c). For none goes any way to explaining why the judgements formed as a result of their method might be explained by moral truth. For example, the conservativist position, by itself, gives no reason to think that any of the judgements we happen to find ourselves with are explained by moral truth. Likewise, it is a

well-worn criticism of coherentist views that a perfectly coherent set of beliefs can be false. The problem is that these epistemologies are *too pure* to provide any positive reason for thinking that the judgements they produce are connected to moral truth. They *presuppose* this claim, rather than independently supporting it (Joyce 2016: 139-43).

6.2. An Alternative Reply

One diagnosis of this failure is that it is difficult to provide an account of the mechanisms whereby some moral judgements are explained by moral truth without saying anything about the nature of that truth. Just as it seems impossible to give an account of how we know the meaning of words without assuming that they have meaning, it seems impossible to give an account of how we can know moral truths without assuming that there are such truths to be known. The pertinent question, then, is this: *What, if anything, is it legitimate to assume about the nature of moral truth in order to construct a general account of how some of our moral judgements might be explained by it*.[3] My answer to this question will be: enough to make rejecting (4c) reasonable.

6.2.1. General Principle

Consider a set of judgements concerning a putative subject matter – S-judgements – a set of truths which constitute that subject matter – S-truths – and a general account of how such truths could help explain some of those judgements – an S-epistemology. The following principle appears undeniable: In order to be justified in claiming that there is no plausible S-epistemology we need to make non-trivial assumptions about the nature of S-truth. This can

---

[3] This is an instance of a general question: What, if anything, is it legitimate to assume about the nature of moral truth in order to answer *any* sceptical challenge? Lillehammer (2010: 374) and Vavova (2014) note that it is impossible for the debunker to establish that our moral belief-forming mechanisms are *off* track with respect to moral truth without making assumptions about that truth. However, the 'Lack of Method' argument does not seek to show that our moral belief-forming mechanisms are *off* track, only that we lack an account of how they might be *on* track. Therefore it remains, at this point, an open question whether assumptions about moral truth are legitimate. Fitzpatrick (2015) assumes that they are.

be seen using a variant of the concept-pill scenario. Suppose we are told that X-judgements are (partly) the result of being slipped a pill which gives us X-concepts. Does this show that there is no plausible X-epistemology? Hardly. For unless we are told something about the nature of X-truth, we cannot rule out that it features in the explanation of some of our X-judgements. X-truth might, for example, be truth concerning the composition of the pills we ingested. Returning to Joyce's pill cases, it is only because we make tacit non-trivial assumptions about the nature of Napoleon-truth – viz. that Napoleon is a physical being who exists independently of our own minds and who has nothing to do with the pills we were slipped – that we take the scenarios described to show that there is no plausible Napoleon-epistemology.

Having established the general principle, it follows that in order to assess whether there is a plausible moral epistemology– i.e. whether (4c) is true – we need to make non-trivial assumptions about *moral* truth.

6.2.2. Formal vs. Substantive Assumptions

But *which* non-trivial assumptions are legitimate? Distinguish two types of answer.

According to the *formal view*, in assessing whether there is a plausible S-epistemology we are permitted only to make non-trivial assumptions about S-truth which do not (by themselves) entail any substantive S-claims. For example, a non-trivial yet formal assumption about Napoleon-truth is that it is truth concerning a physical being. A non-trivial yet formal assumption about moral truth is that it is mind-independent in Street's sense (§1). Two further non-trivial formal assumptions about moral truth are suggested by Joyce (2006: 57-64): that it concerns requirements that are inescapable (i.e. do not depend on the agents to whom they apply having relevant motives) and that have rational authority (i.e. necessarily entail practical reasons). These assumptions are formal insofar as they do not entail that the

Napoleon-related or moral properties are ever instantiated: they merely tell you what those properties would be like, were they to be instantiated.

According to the *substantive view,* in assessing whether there is a plausible S-epistemology we are permitted to make non-trivial assumptions about S-truth which entail substantive S-claims. In other words, according to the substantive view, it is permissible, in constructing an S-epistemology, to assume the approximate truth of some of our positive S-judgements.

Which view should we prefer? The formal view seems *prima facie* legitimate, at least in some cases. Suppose a friend claims that they can communicate with ghosts – non-physical beings not bound by physical laws. In assessing whether there is a plausible general account of how ghost-truths could explain some of our friend's ghost-judgements we need to make *some* assumptions about ghost-truth (e.g. that it is truth concerning non-physical beings) but we surely grant our friend too much if we permit them to assume, for the task of constructing their ghost-epistemology, that auntie's ghost lives in the attic.

Yet in other cases the formal view seems inappropriate. Consider dry-goods judgements. Suppose that in assessing whether there is a plausible dry-goods epistemology we are permitted only to make formal assumptions about dry-goods truth – for example that it is truth concerning mind-independent material entities. On such a slender basis, constructing an epistemology is impossible. Unless we begin with the substantive assumption that we ourselves are dry-goods, possessed with sense organs which provide some type of connection to other dry-goods, and inhabiting a dry-goods world somewhat like our initial dry-goods judgements take it to be, we cannot begin to investigate in detail how some of our dry-goods judgements might be explained by dry-goods truth. But once we make such assumptions the beginning of a dry-goods epistemology (involving, for example, details about the functioning

of our sense organs) is relatively easy to construct. This partial epistemology can in turn help illuminate how some (perhaps most) of our dry-goods are well-explained by dry-goods truth, while others (including some of those on which we initially relied) are *not* so explained, and should be disowned. From here, further processes of refinement, such as eliminating inconsistencies and arbitrariness, adding further beliefs that increase overall coherence and weeding-out judgements based on unjustified assumptions, can help create a more complete theory of the dry-goods world (perhaps even including details about the evolutionary history of certain humanoid dry-goods and their perceptual capacities). This in turn can help us provide a more complete epistemology, and clearer ideas of when our dry-goods judgements are in error. Thus a gradual process of refinement emerges, with initial judgements helping to construct a tentative epistemology, which may in turn (together with processes of reasoning and reflection) help us refine our theory of the dry-goods world, which in turn helps refine our epistemology, and so on. The success of such an account is neither trivial nor inevitable (Shogenji 2000), but seems impossible to get off the ground without assuming, at the outset, the defeasible justifiedness of some substantive dry-goods judgements (Field 1989: 25-30, see Boyd 1988 §3.2).

Blackburn makes a similar point for the case of physics. He asks what kind of theory might vindicate our confidence that, if we exercise our sensory and cognitive faculties well and end up believing that *p*, then *p*. He continues:

> If *p* is a thesis from basic physical theory, only the theory itself…Any attempt at a
>
> background, an underwriting of the conditional from outside the theory, is certain
>
> to be bogus. (1993: 167)

If so, in assessing whether there is a plausible general account of how S-truths could explain some of our S-judgements, it is sometimes legitimate to make substantive assumptions about the nature of S-truth. But: Is it legitimate in the *moral* case?

6.2.3. Epistemology for Mind-Dependent Conceptions of Moral Truth

Suppose it is not. On this *formal view*, in assessing whether there is a plausible moral epistemology we are not permitted to make substantive moral assumptions. Does it follow that there is no plausible moral epistemology? Not necessarily. We saw in the case of ghost-judgements and dry-goods judgements that it was impossible to construct a plausible epistemology while only making formal assumptions about the relevant truths. In both cases, the formal assumptions included the assumption that the relevant truth is (or would be) mind-independent. But suppose that the assumptions in the moral case include the thought that moral truth is mind-*dependent*.[4] This is to say that moral truth concerns the instantiation of properties which satisfy the following schema:

Object, x, possesses property M iff. x tends to elicit R from P in C.

Where R is a reaction (e.g. an emotion), P a set of persons and C a set of circumstances (such as conditions of imaginative acquaintance). To capture genuine mind-dependence, the conditions on the right-hand side must also be specifiable independently of the extension of M (Johnston 1989).

There are many possible versions moral mind-dependence. But the present concern is just whether such an account can help construct a moral epistemology. And it seems that it can. As Lewis notes:

---

[4] Another possibility is to deploy the formal assumption that moral terms are rigid designators, and therefore that moral truth is whatever natural truth is revealed by a privileged subset of our uses of those terms (see Boyd 1988). On this view, moral epistemology is straightforward: some instances of moral properties are causally responsible for some of our moral judgements (Setiya 2012: 119-120; Lillehammer 2016: 117).

In general, to find out whether something is disposed to give response R in conditions C, you can put it in C and find out whether you get R. (1989: 116)

Hence, assuming that moral properties are mind-dependent, moral judgements formed as a result of people placing themselves in (or approximating) circumstances C and seeing whether, in those circumstances, they have reaction R, will be partly explicable in terms of moral truth. (Perhaps this is just the method we employ when using thought experiments in ethics; note also that it is consistent with the possibility of error.) Hence, even if we are not permitted to make substantive assumptions about moral truth in order to construct a moral epistemology, the purely formal assumption that moral truth is mind-dependent is enough.[5]

In response, debunkers might reject the assumption that moral truth is mind-dependent. Joyce argues that construing moral truth as mind-dependent fails to accommodate the fact that moral properties, when instantiated, provide requirements which are inescapable and rationally authoritative (2006: 190-209). But two points are worth mentioning. First, if one accepts Joyce's approach, then one must hold that (4c) is only reasonable to those who have independent reasons to reject moral mind-dependence. These reasons will be metaethical. Hence, '[this] debunking argument has teeth only if certain metaethical arguments succeed' (Joyce 2016: 136).

Second, one might think that the mere fact that the debunking argument *would* succeed were one to reject moral mind-dependence casts doubt on that rejection (compare Setiya 2012: 115; Bedke 2014: 123).[6] That is, one might *tollens* the following *ponens*:

(i)   If moral truth is not mind-dependent, then (by the *Lack of Method* argument) all (actual, human) moral judgements are unjustified.

---

[5] For a different attempt to construct a moral epistemology from purely formal assumptions, see Behrends 2013.
[6] Given footnote 5, the following argument over-simplifies a little.

(ii)   Moral truth is not mind-dependent.

Hence

(iii)   All (actual, human) moral judgements are unjustified.

Dialectical stalemate looms. Joyce accepts (i), (ii) and (iii). The alternative accepts (i), not-(iii) and not-(ii). Joyce may insist that his position is preferable, since (ii) is independently more plausible than not-(iii). But in fact the very same types of reasons which Joyce cites in support of (ii) support the denial of (iii). One type of argument Joyce gives is a version of the argument from moral appearances (Glassen 1959; Joyce 2001: 12-16; see Sinclair 2012b). According to this argument the way in which ordinary moralisers use moral language reveals the linguistic conventions which govern moral discourse and (thus) delineates our conception of moral truth. Further, when we look at this usage, we see that ordinary moralisers deploy moral language in a way that confirms the assumption that moral truth is (conceived as) mind-independent. Unfortunately, in the current context a similar argument supports the view that some moral judgements are justified. For when we look at the way people use moral language and concepts, we see that this use reveals a conception of moral truth such that some of our current judgements are justified (if approximate) intimations of it. One potential lesson here is that actual usage reveals that everyday morality exemplifies no coherent conception of moral truth (Lillehammer 2003b). But it doesn't follow that the best conception of moral truth is the Joycean version (which eschews epistemological optimism in favour of mind-independence) rather than the alternative (which eschews mind-independence in favour of epistemological optimism).

     The second type of argument Joyce gives for (ii) is that views which take moral truth to be mind-dependent cannot accommodate the inescapability and rational authority of moral truth (2006: 190-199). These are 'non-negotiable' features of moral practice: ideas about the

nature of moral truth such that, if they were not assented to by a population of speakers, that population would not be interpreted as engaging in *moral* practice at all (2006: 199-209). Here again the problem is that if the argument goes through in the case of the assumptions of inescapability and rational authority, it also goes through in the case of the assumption that some of our moral judgements are justified. For a population of speakers and thinkers otherwise like us, but who took none of their moral judgements to be justified, could scarcely be considered to be engaging in moral practice.

Even if the arguments of the last two paragraphs are unsuccessful, the previous point stands. Given our current assumptions about what it is permissible to assume when constructing a moral epistemology, it is reasonable to accept (4c) only if one has sufficient independent reason to reject the claim that moral truth is mind-dependent. The set of philosophers who deny that there is such reason is not empty. It includes, for example, Brower (1993), Lewis (1989), Railton (1986), Smith (1989) and Street (2006).

6.2.4 Epistemology for Mind-Independent Conceptions of Moral Truth

The previous subsection considered the formal view applied to the moral case. On the alternative, substantive, view in assessing whether there is a plausible general account of how moral truth could explain some of our moral judgements, it is legitimate to make substantive assumptions about moral truth.

There are at least two versions of this approach. According to the first – *reductivist* – version the relevant assumptions include claims regarding the identity of moral with natural properties. For example, one might hold that the property of moral rightness is identical with the property of being maximally pleasure-producing (it does not follow that the two phrases are synonymous). According to the second – *non-reductivist* – version the relevant assumptions include both formal claims about the nature of moral properties (e.g. that they

are mind-independent) and substantive claims about the distribution of their instances. According to both versions moral truth is relevantly mind-independent. But non-reductive views are compatible with taking the relevant substantive assumptions to be *partial* accounts of moral truth – that is, assumptions which do not specify necessary and sufficient conditions for an action to be right or a state of affairs good (for example), but which nevertheless include substantive claims.

If it is legitimate to make substantive moral assumptions in the process of assessing whether there is a plausible general account of how moral truth could explain some of our moral judgements then, it seems, it is relatively easy to construct a moral epistemology. For example, assuming utilitarianism, all judgements of rightness formed after a process which is sensitive to facts about which actions maximise pleasure would be explicable in terms of moral truth (see Setiya 2012: 112). Or consider the view that 'what's morally good has to do with behaviors that promote rather than hinder wellbeing' (Brosnan 2011: 62; see Raz 1986: 194) or which answer to fundamental human needs (Boyd 1988 §4.3). Again, judgements of goodness formed by a process sensitive to facts about which actions promote wellbeing or fundamental needs will be good candidates for judgements that are explicable by moral truth. Or consider the view that what's just for a society has to do with which social arrangements best satisfy or respect the deeply held concerns of its members. Then judgements of justice formed after a process that is sensitive to such facts are good candidates for being explicable by moral truth. In these latter cases, anti-debunkers can even co-opt elements of Joyce's evolutionary account into their epistemology, since if, as Joyce's theory claims, one role of moral judgements is to facilitate productive schemes of social co-operation, we can expect such judgements to be somewhat sensitive to facts about human needs and concerns, and about which schemes of co-operation best promote these (see Boyd 1998 §4.4; FitzPatrick 2015: 894; Sinclair 2012a). Insofar as these latter cases also begin with substantive

assumptions about moral truth that are *partial*, they are compatible with a process of refinement that mirrors that described in the dry-goods case (§6.2.2), whereby initial assumptions about moral truth help generate a tentative moral epistemology, which in turn helps refine our sense of which of the initial assumptions are, after all, trustworthy, which in turn helps further our theory of moral truth, which allows us in turn to refine the epistemology, and so on, until ultimately a complete conception of moral truth and the nature of our contact with it is reached. This, of course, is just the method of broad reflective equilibrium in ethics (Boyd 1988 §4.2; Brink 1989: 122-33).[7]

It seems, therefore, that constructing an epistemology for mind-independent conceptions of moral truth is perfectly possible, *provided we are permitted substantial assumptions about where that truth lies*. The debunker will reply that it is *not* legitimate to rely on such assumptions. In fact there are two distinct replies here, worth distinguishing. (a) It is illegitimate, in the moral case, to begin the task of epistemology by making *any* substantive assumptions. (b) It is illegitimate to begin this task with any of the particular substantive assumptions mentioned above. I'll take these in turn.

Why might it be illegitimate, in the moral case, to make *any* substantive assumptions when constructing an epistemology? I argued above (§6.2.2) that it is not *universally*

---

[7] Note that the accounts outlined here are not identical with so-called 'third-factor' views. According to Enoch (2010), for example, the substantive assumption that survival is good explains why there is a pre-established harmony between some evolutionarily-selected moral beliefs and moral truth, but particular moral truths are not involved in the explanation of particular beliefs (for variations and criticism see Bedke 2014; Setiya 2012: 111-16). The accounts suggested here are distinct insofar as, if successful, particular moral beliefs can be explained in terms of (and sensitive to) their content. This is one place where the relevant incompleteness of the evolutionary hypotheses is key (see FitzPatrick 2015 for one version of this view). Note also that this method moves beyond simply assuming particular moral truths for the sake of explaining our beliefs in them (see Vavova 2014: 86). Rather it begins with defeasible assumptions about the nature of that truth and uses them to attempt to construct a more detailed yet general moral epistemology (Bedke (2014: 113) calls this 'probationary epistemology'). As in the perceptual case, the assumptions with which the process begins (e.g. that the world is roughly like initial perception presents it as being, that goodness has to do with fundamental needs) will not generally be explanans in the resulting explanations (e.g. the table in front of me explains my belief in it, the goodness of this action explains why I believe it to be good) but rather parts of the general picture within which such explanations are justified.

illegitimate. What particular reason might there be for thinking that the formal approach is the right one to take in the moral case? I consider four putative arguments.[8]

First, one might appeal to the analogy of the poisoned well, or the principle of junk-in/junk-out (Street 2006: 123-4, 140). The thought is that epistemology cannot begin by making assumptions whose epistemic credentials are themselves in doubt. But sawn of rhetorical imagery, this argument is no more than the re-assertion of the applicability of the formal view to the moral case. Furthermore, the current context is one of attempting to provide independent support for a premise in a global justification debunking argument. This is not a context where the epistemic credentials of all moral judgements are as yet in doubt.

Second, one might think that the level of disagreement in moral judgements counts against assuming even the approximate truth of any (positive) moral judgement when constructing a moral epistemology. However, on this view, the strength of the debunking argument depends entirely on premises about the scope and nature of moral disagreement. These premises are both highly contentious and have been taken to feature in distinct

---

[8] Four further arguments can be quickly dismissed. The first holds that the fact that moral judgements can be explained without citing moral truth shows that they are subject to *distorting* influences (Fraser 2014: 469; Street 2006: 121). This argument fails since the distortion metaphor only makes sense given substantive assumptions about moral truth (Lillehammer 2010: 375; White 2010: 590). The second suggests that were we 'allowed to rely on some substantive moral claims in assessing the reliability of moral faculties, then we would be effectively immunizing a class of moral beliefs from genealogical doubts' (Shafer-Landau 2012: 21, see Vavova 2014: 81). But according to the methodology outlined above, no judgement is completely immune from doubt – rather, certain judgements are presumed to be justified for the purposes constructing an epistemology. Both the failure to construct such an epistemology, and its details, may be sufficient to undermine this presumptive justification (this applies in both the perceptual and the moral cases; see Bedke 2014: 113-114, Wielenberg 2016: 93). The third holds that any substantive assumptions we make are illegitimate because contingent, i.e. we could have easily begun with different assumptions, making any belief formed using the methods of the resulting epistemology at best accidently true. But this point is too general since it would equally rule out, for example, relying on substantive assumptions in the dry-goods case – we are equally lucky that we are not brains-in-vats. Given these assumptions, and the fact that it is possible to construct an epistemology on the basis of them, it is no accident that any belief well-formed by the methods of that epistemology is true. Neither will it be an accident that such methods are themselves reliable (contra Setiya 2012: 112-14) since the process of reflective equilibrium, if successful, will have refined them in light of their earlier reliability. (For discussion of contingency see Boyd 1988 §3.3, 4.4; Lillehammer 2010: 375-77. An alternative response would be to begin with necessary substantive assumptions – see Cuneo & Shafer-Landau 2014, discussed in Vavova 2015). Fourth it might be illegitimate to make substantive moral assumptions because, unlike in the case of dry-goods judgements, we have no way of testing or validating (e.g. in terms of predictive power) the theories of the moral domain that may result (see Boyd 1988 §4.2). But this is question-begging in the current context, since it amounts to the demand that moral theory be assimilated to scientific theory.

arguments for moral scepticism (see e.g. Decker & Groll 2013). On this view the evolutionary debunking argument is at best redundant, at worst unsound.

Third, one might hold that insofar as one of the tasks of metaethics is to *locate* moral practice within the world-view which begins with dry-goods judgements and ends with a fully developed natural science, we should not begin this task by assuming the actual instantiation of any moral properties (by contrast, there is no 'location' problem for science; Blackburn 1993: 166). However, on this view the debunking argument is no threat to those who deny the naturalistic methodological assumption that the task of metaethics includes a *placing* of moral properties within the natural realm (McDowell 1985). Furthermore, even methodological naturalists may question whether their naturalism precludes them from assuming the approximate truth of some moral judgements when constructing a moral epistemology. After all, one notable feature of the natural phenomenon of moral practice is that practically all its practitioners are convinced of the truth of *some* moral judgements; a plausible explanation of this is that some of those judgements are (approximately) true.

Finally, the debunker might argue that reflection on the nature of moral properties supports the error-theoretic view that such properties *cannot* be instantiated and hence no moral judgement can be relied on when constructing a moral epistemology. However, this debunking argument is entirely dependent on a distinct truth-debunking argument (see §1). Furthermore, one might question the latter: it seems that the very same sorts of considerations which support the error-theoretic claim that moral judgements are cognitive also support that claim that some of those judgements are approximately true (Kirchin 2010).

Hence it seems that there is no reason as yet to think that it is illegitimate to begin the task of constructing a moral epistemology by making substantive moral assumptions. Still, the debunker might reply– this is point (b) above – that none of the particular substantive

assumptions with which anti-debunkers begin this task are plausible. This seems to be Joyce's view:

> I'll go out on a limb and assert…that we do not have a believable account of how moral facts could help explain the mechanisms and forces that give rise to moral judgments. …. Some people think that they already have believable accounts, but they're mistaken…Utilitarians, for example, believe that they can explain how moral facts relate to moral judgments, but anyone who isn't a utilitarian…thinks that they're mistaken. Kantians…believe that they can explain how moral facts relate to moral judgments, but anyone who isn't a Kantian…thinks that they're mistaken. And so on. It doesn't really add anything shocking to this…picture to accept the view that all people who think that they can explain how moral facts relate to moral judgments are mistaken. (2016: 135)

Note the assumption of this passage: that in order to construct an account of how moral facts could explain some of our moral judgements, we need a *complete* account of moral truth, such as utilitarianism. Joyce may be right that this is one way of proceeding, but it is not the only way. Given the analogy with dry-goods judgements, all that is required to get the epistemological story off the ground is the assumption of the approximate truth of *some* judgements: the content of these judgements needn't constitute a complete account of moral truth.

The upshot is as follows. We have seen no particular reason to consider it illegitimate, for the task of constructing a moral epistemology, to begin by assuming the approximate truth of some (positive) moral judgements. Given this, it is possible to begin to construct a plausible account of how moral truth might be involved in the explanation of some of our moral judgements, and thus reject (4c).

6.2.5 Summary and 0bjection

Here, then, is my response to the *Lack of Method* argument. Whether there is a plausible general account of how moral truths could help explain some of our moral judgements depends on what it is permissible to assume about moral truth in order to construct such an account. If we make the formal assumption that moral truth is mind-dependent then such an account is available and (4c) can be rejected.[9] Distinct moral epistemologies are available if we assume that moral truth is mind-*in*dependent, but add to this substantive assumptions about the distribution of moral properties. Some such assumptions are complete accounts of moral truth, and I agree with Joyce that many such assumptions are implausible. But other assumptions are merely partial and, crucially, these are enough to begin to construct a moral epistemology, and hence reject (4c). Further, we have seen no particular reason to think that, in the moral case, making such substantive assumptions is illegitimate. (At least, no reason that is compatible with evolutionary justification debunking arguments maintaining independent force.)

One final reply from the would-be debunker runs as follows. 'I accept that there are many types of metaethical view which can reject (4c). In particular (4c) can be rejected by those who accept a mind-dependent conception of moral truth, by those who take moral properties to be reducible to particular natural properties, and by those who believe that moral truth is mind-independent, non-reducible, but supervenient on particular aspects of natural truth. But all these views are implausible for independent reasons. Both mind-dependent and reductionist views fail to capture the non-negotiable elements of inescapability and rational authority, and the non-reducible view is ruled out by Ockham's razor' (see Joyce 2006: 190-211). One problem here is the possibility of a quasi-realist view which provides expressivist-friendly interpretations of the claims that moral truth is mind-independent and non-reducible,

---

[9] Other formal assumptions may do the same job – see footnote 5.

and which doesn't violate Ockham's razor. But this aside, even if the above reply is accepted the evolutionary debunking argument would be redundant. For if we have good reason to think that moral truth is neither reducible to any natural truth nor existent but irreducible, then we have a much more direct argument for thinking that none of our moral judgements are justified. For if we know that moral truth does not exist, we know that none of our (positive) judgements about it are justified.

7. Conclusion

Here is a tentative hypothesis to diagnose the fact that some find justification debunking arguments compelling, while others are unmoved. The former consider it illegitimate, for the task of answering sceptical challenges, to assume even the approximate truth of any substantive moral judgement. Such people find premises like (4c) plausible. Conversely, the latter consider it *legitimate*, for the same task, to assume the approximate truth of some moral judgements. One way of understanding the latter group is that they hold that, as in the case of physics, the epistemic validation of our moral judgements can only come from a position *within* the relevant practice. If this is right, progress in debunking debates will only be made by addressing the underlying question of what it is legitimate to assume (about relevant truth) in answering sceptical challenges. This will involve, in turn, a clear grasp of the intended import, independence and generality of each sceptical challenge. More particularly, in the case of evolutionary debunking arguments of morality the challenge for the justification debunker will be to give independent arguments for their constraints on constructing moral epistemologies which do not also render their debunking argument superfluous.[10]

References

Blackburn, S. 1993. *Essays in Quasi-Realism*. Oxford: Oxford University Press.

Bedke, M. 2014. 'No Coincidence?' *Oxford Studies in Metaethics 9*: 102-125.

Behrends, J. 'Meta-normative Realism, Evolution and our Reasons to Survive' *Pacific Philosophical Quarterly* 94: 486-502.

Boyd, R. 1988. 'How to be Moral Realist' in G. Sayre-McCord, ed., *Essays on Moral Realism*. Ithaca: Cornell University Press. 181-228.

Brink, D. 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.

Brosnan, K. 2011. 'Do the Evolutionary Origins of Our Moral Beliefs Undermine Moral Knowledge?' *Biology and Philosophy* 26: 51-64.

Brower, B. 1993. 'Dispositional Ethical Realism' *Ethics* 103: 221-49.

Campbell, R. 1996. 'Can Biology Make Ethics Objective?' *Biology and Philosophy* 11: 21-31.

Casebeer, W.D. 2003. *Natural Ethical Facts: Evolution, Connectionism, and Moral Cognition*. Cambridge, Ma.: MIT Press.

Clarke-Doane, J. 2012. 'Morality and Mathematics: The Evolutionary Challenge' *Ethics* 122(2): 313-40.

Clarke-Doane, J. 2016. 'Debunking and Dispensability' in Leibowitz and Sinclair 2016.

Cuneo, T & Shafer-Landau, R. 'The Moral Fixed Points' *Philosophical Studies* 171 (3): 399-443.

Dancy, J. 1985. *Introduction to Contemporary Epistemology*. Oxford: Basil Blackwell.

Decker, J. & Groll, D. 2013. 'The (In)Significance of Moral Disagreement for Moral Knowledge' In R. Shafer-Landau, ed., *Oxford Studies in Metaethics, Volume 8*. Oxford: Oxford University Press.

Enoch, D. 2010. 'The Epistemological Challenge to Metanormative Realism: How Best to Understand it, and How to Cope with it' *Philosophical Studies* 148: 412-438.

Field, H. 1989. *Realism, Mathematics and Modality.* Oxford: Blackwell.

FitzPatrick, W. 2015. 'Debunking Evolutionary Debunking of Ethical Realism' *Philosophical Studies* 172: 883-904.

Fraser, B. 2014. 'Evolutionary Debunking Arguments and the Reliability of Moral Cognition' *Philosophical Studies* 168: 457-73.

Glassen, P. 1959. 'The Cognitivity of Moral Judgments' *Mind* 68(269): 57-72.

Greene, J. 2008. 'The Secret Joke of Kant's Soul' in W. Sinnott-Armstrong, ed., *Moral Psychology: The Neuroscience of Morality*, vol. 3: 35-79.

Handfield, T. 2016. 'Genealogical Explanations of Change and Morals' in Leibowitz and Sinclair 2016.

Harms, W.F. 2000. 'Adaption and Moral Realism' *Biology and Philosophy* 15: 699-712.

Huemer, M. 2005. *Ethical Intuitionism*. Palgrave Macmillan.

Ichikawa, J. & Steup, M. 2012, 'The Analysis of Knowledge' *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2014/entries/knowledge-analysis/>.

Johnston, M. 1989. 'Dispositional Theories of Value' *Proceedings of the Aristotelian Society*, Supplementary Volume 63: 139-74.

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

Joyce, R. 2006. *The Evolution of Morality*. Cambridge, Ma.: MIT Press.

Joyce, R. 2014. 'The Evolutionary Debunking of Morality' in J. Feinberg and R Shafer-Landau, eds., *Reason and Responsibility: Readings in Some Basic Problems of Philosophy*, 15th edition.

Joyce, R. 2016. 'Reply: Confessions of a Modest Debunker' in Leibowitz and Sinclair 2016.

Kahane, G. 2011. 'Evolutionary Debunking Arguments' *Noûs* 45(1): 103-25.

Kirchin, S. 2010. 'A Tension in the Moral Error Theory' In R. Joyce &S. Kirchin, eds., *A World Without Values: Essays on John Mackie's Moral Error Theory*. London: Springer.

Leibowitz, U. & Sinclair N. 2016. *Explanation in Ethics and Mathematics: Debunking and Dispensability*. Oxford: Oxford University Press.

Lewis, D. 1989. 'Dispositional Theories of Value' *Proceedings of the Aristotelian Society*, Supplementary Volume 63: 113-37.

Lillehammer, H. 2003a. 'Debunking Morality: Evolutionary Naturalism and Moral Error Theory', *Biology and Philosophy* 18: 567-81.

Lillehammer, H. 2003b. 'The Idea of a Normative Reason'. In P. Schaber & R. Huntelmann (eds.), *Grundlagen der Ethik*. 41-65.

Lillehammer, H. 2010. 'Methods of Ethics and the Descent of Man: Darwin and Sidgwick on Ethics and Evolution' *Biology and Philosophy* 25: 361-78.

Lillehammer, H. 2016. '"An Assumption of Extreme Significance": Moore, Ross and Spencer on Ethics and Evolution' in Leibowitz and Sinclair 2016.

Mason, K. 2010. 'Debunking Arguments and the Genealogy of Religion and Morality' *Philosophy Compass* 5(9): 770-78.

McDowell, J. 1985. 'Values and Secondary Properties' in T. Honderich, ed., *Morality and Objectivity*. London: Routledge and Kegan Paul. 110-29.

Mikhail, J. 2011. *Elements of Moral Cognition*. Oxford: Oxford University Press.

Nichols, S. 2014. 'Process Debunking and Ethics' *Ethics* 124: 727-749.

Nozick, R. 1981. *Philosophical Explanations*. Oxford: Oxford University Press.

Quinn, W. 1986. 'Truth and Explanation in Ethics' *Ethics* 96(3): 524-44.

Railton, P. 1986. 'Moral Realism' *Philosophical Review* 95: 163-207.

Raz, J. 1986. *The Morality of Freedom.* Oxford: Clarendon Press.

Ross, W.D. 1930. *The Right and the Good.* Oxford: Oxford University Press.

Ruse, M. 1986. *Taking Darwin Seriously*. Oxford: Basil Blackwell.

Setiya, K. 2012. *Knowing Right From Wrong*. Oxford: Oxford University Press.

Singer, P. 2005. 'Ethics and Intuitions' *The Journal of Ethics* 9: 331-52.

Sinclair, N. 2012a. 'Metaethics, Teleosemantics and the Function of Moral Judgements' *Biology and Philosophy* 27(5): 639-662.

Sinclair, N. 2012b. 'Moral Realism, Face-Values and Presumptions' *Analytic Philosophy* 53(2): 158-179.

Shafer-Landau, R. 2012. 'Evolutionary Debunking, Moral Realism and Moral Knowledge' *Journal of Ethics and Social Philosophy* 7(1): 311-329.

Shogenji, T. 2000. 'Self-Dependent Justification without Circularity' *British Journal for the Philosophy of Science* 51: 287-298.

Smith, M. 1989. 'Dispositional Theories of Value' *Proceedings of the Aristotelian Society, Supplementary Volume* 63: 89-111.

Street, S. 2006. 'A Darwinian Dilemma for Realist Theories of Value' *Philosophical Studies* 127: 109-66.

Toner, C. 2011. 'Evolution, Naturalism, and the Worthwhile: A Critique of Richard Joyce's Evolutionary Debunking of Morality' *Metaphilosophy* 42(4):520-46.

Vavova, K. 2014. 'Debunking Evolutionary Debunking' *Oxford Studies in Metaethics 9*: 76-101.

Vavova, K. 2015. 'Evolutionary Debunking of Moral Realism' *Philosophy Compass* 10(2): 104-116.

White, R. 2010. 'You Just Believe that Because…' *Philosophical Perspectives* 24: 573-615.

Wielenberg, E. 2010. 'On the Evolutionary Debunking of Morality' *Ethics* 120(3): 441-64.

Wielenberg, E. 2016. 'Evolutionary Debunking Arguments in Religion and Morality' in Leibowitz and Sinclair 2016.