‘An Assumption of Extreme Significance’: Moore, Ross and Spencer on Ethics and Evolution

Hallvard Lillehammer, Department of Philosophy, Birkbeck, University of London, Malet Street, London WC1E 7HX, UK, <h.lillehammer@bbk.ac.uk>

Abstract

In recent years there has been a growing interest among mainstream Anglophone moral philosophers in the empirical study of human morality, including its evolution and historical development. This chapter compares these developments with an earlier point of contact between moral philosophy and the moral sciences in the early decades of the Twentieth century, as manifested in some of the less frequently discussed arguments of G. E. Moore and W. D. Ross. It is argued that a critical appreciation of Moore and Ross’s response to the emerging moral sciences of their day has significant implications for contemporary moral epistemology. The chapter also offers a novel interpretation of G. E. Moore’s ‘open question argument’.
‘An Assumption of Extreme Significance’: Moore, Ross and Spencer on Ethics and Evolution

1. Introduction

After something of a ‘hiatus’, in recent years there has been a growing interest among mainstream Anglophone moral philosophers in the empirical study of human morality and its significance for moral philosophy. In this chapter I raise three questions. First, what happened in the early years of the Twentieth Century that made prominent members of mainstream Anglophone moral philosophers turn away from an interest in the ‘moral sciences’ (as they were then called), in the face of the rapid emergence of the empirical disciplines of psychology, sociology, anthropology and evolutionary biology? It may be tempting to think that the moral philosophers of the time simply ignored these developments in adjacent disciplines, but a closer reading some of the best known works of the period shows that they did not. On the contrary, they responded to these developments, and turned away from them on the basis of philosophical arguments.

1 Parts of this material have been presented to the Amoral Sciences Club in Cambridge, the Royal Institute of Technology in Stockholm, the Metaphysical Society in Trinity College Dublin, the Phileas Society in Geneva, and Birkbeck, University of London. I am grateful to the audience on those occasions for comments and questions, and to Katerina Deligiorgi for written comments on a previous draft intended for the 2013 Morality and Explanation conference at Nottingham.
Second, why did the philosophers in question turn away from these developments? It may be tempting to think that the explanation is closely connected with the alleged diagnosis of the so-called ‘naturalistic fallacy’ in G. E. Moore’s seminal 1903 work *Principia Ethica*, but once more a closer reading of some of the best known works of the period shows that this hypothesis wrongly overplays the metaphysical aspect of the arguments in question at the expense of their epistemological aspect.\(^2\) Third, were the philosophers in question right to turn away from these developments in the way they did? It may be tempting to think that the philosophers in question turned away from these developments for reasons that more recent advances in philosophy and the moral sciences show to be either primitive or outdated, but I will suggest that their reasons for doing so display striking similarities with a common way of responding to analogous developments in contemporary moral philosophy. I will illustrate this point by briefly commenting on the ongoing debate about the normative significance of neuroscience (c.f. Berker 2009). In doing so, I will show that comparing the earlier point of contact between moral philosophy and the moral sciences with analogous points of contact in the present can serve to illuminate a number of issues in contemporary moral epistemology.

2. Moore and Ross on evolutionary ethics

I begin with the obvious, and therefore with the place of ‘the naturalistic fallacy’ in philosophical responses to developments in the moral sciences in the latter part of the

\(^2\) Of course, the metaphysical and epistemological aspects of these arguments are closely connected, insofar as the identification of moral qualities as non-natural went hand in hand with the identification of basic moral knowledge as *a priori*. More of this below. See also Hurka 2014; Skorupski, MS.
Nineteenth Century and the early parts of the Twentieth. On one common way of reading the history of moral philosophy, the obvious way to explain the early and frequent dismissal of the moral sciences as irrelevant to ‘ethics proper’ is to see it as an implication of the alleged diagnosis of the ‘naturalistic fallacy’ by Moore in his *Principia Ethica*. On this reading, the normative insignificance of the moral sciences boils down to a basic truth of logic and/or metaphysics, namely the existence of an irreducible gap between evaluative/normative as opposed to descriptive/natural properties or predicates: what is good as opposed to what just is; or what ought to be as opposed to what is the case. It is one thing to say what our moral sensibility is actually like, or to describe and explain its development, function, and variable historical expressions. It is quite another thing to say what, if anything is good about it, or how it ought to be expressed. The enormous influence of this line of argument cannot be denied, and is clearly present at some level in both the authors I go on to discuss in this chapter. Yet even so, the idea that it is their diagnosis of the naturalistic fallacy that explains their reactions to the moral sciences of their time is a truth with important qualifications. As I will show in what follows, the naturalistic fallacy plays at best an auxiliary role both in Moore’s response to the moral sciences in the guise of the evolutionary ethics of Herbert Spencer, and in W. D. Ross’s later response to the moral sciences in the guise of evolutionary ethics and the

---

3 As formulated by Moore, the naturalistic fallacy concerns the definition of good and has only indirect implications for claims about what is right, what ought to be, and therefore the so-called ‘is/ought’ gap targeted in what has come to be known as ‘Hume’s Law’. For more on this issue, see Moore 1993, 1912; Ross 1930. I say more about how to read Moore’s ‘open question argument’ in support of his non-naturalist position in the Appendix below.
French sociology. Moreover, this is the case even in passages of their work where Moore and Ross are explicitly concerned to dispute the attempt to draw substantially normative inferences on the basis of empirical claims about the nature of morality.

A clear case in point can be found near the start of W. D. Ross’s 1930 work *The Right and the Good*. As a prolegomenon to his account of our basic moral knowledge as consisting of a set of *prima facie* moral principles, Ross takes it upon himself to argue against the temptation to draw positive moral conclusions directly from empirical facts about the function of moral systems. His prime target in these passages is the emerging science of sociology, in the guise of the sociologists Emile Durkheim and Lucien Levy-Bruhl, both of whom who Ross appears to have read (or have read about) in the original French. Ross writes:

> It would be foolish to deny the value of such a study, or the interest of many of the...

---

4 Of course, both Moore and Ross endorse a form of ethical non-naturalism that entails the rejection of any strict identification between normative and descriptive properties. My point is that it is not this metaphysical thesis that is the primary driver of their responses to work in the moral sciences, but rather a distinct, and logically independent epistemological claim.

5 For one philosopher whose response to the emerging moral sciences did put the semantic/metaphysical features of moral thought at the forefront of discussion, see C. D. Broad’s review of Julian Huxley’s *Evolutionary Ethics*, published four decades after Moore’s *Principia* and in the decade following the appearance of Ross’s main ethical works (see e.g. Broad 1944, 366-7). In this review, Broad also explicitly connects the metaphysical and epistemological elements of the philosophical response I describe in this chapter (see e.g. Broad 1944, 361).

6 The claim that the French sociologists were guilty of committing the naturalistic fallacy (or that they somehow assumed that you can infer an ‘ought’ from an ‘is’) is implausible (see e.g. Levy-Bruhl 1905, 93; Durkheim 1933, 32; Durkheim 1993, 30). For a more detailed discussion of Durkheim and Levy-Bruhl in this context, see Lillehammer forthcoming.
facts it has brought to light with regard to the historical origin of many such [moral] beliefs and practices… What must be denied is the capacity of any such inquiry to take the place of moral philosophy. (Ross 1930, 12)

Some readers may see traces of ‘the naturalistic fallacy’ as a diagnostic tool in this passage. Yet when Ross proceeds to deny the normative significance of the empirical facts that sociology has brought to light, what he offers is not an argument against identifying moral properties with natural properties, or inferring an ‘ought’ from an ‘is’; but rather an argument aimed at a narrowly functional analysis of the correctness conditions of beliefs in general (whether moral or non-moral). He writes:

… the analogy which it [sociology] draws between a moral code and a natural system like the human body… is an entirely fallacious one… [B]eliefs have the characteristic which bodies have not, of being true or false, of resting on knowledge or being the products of wishes, hopes and fears… (Ross 1930, 13)

Ross’s point in this passage is that a purely functional analysis of the correctness conditions of moral (and other) beliefs fails to take account of the fact that beliefs have contents, the correctness conditions of which cannot be assumed without further argument to be a simple function of the social or psychological role that these beliefs contingently serve. Thus, it is not clear that at this point in Ross’s discussion of the normative significance of the moral sciences that ‘the naturalistic fallacy’ is playing any significant role at all.
More plausible traces of ‘the naturalistic fallacy’ as a philosophical diagnostic might be thought to be present in the following remarks from Ross’s Gifford lectures (published as *The Foundations of Ethics* nine years after *The Right and the Good*). The main target of Ross’s discussion of the moral sciences in this book is Herbert Spencer’s evolutionary ethics. In this later work, Ross writes:

> Now what we are considering at present is views as to the *meaning* of rightness, and it is surely obvious that the suggestion that ‘right’ *means* ‘comparatively evolved’ is not one that can be seriously entertained. (Ross 1939, 13)

So there is at least one type of naturalistic definition of at least one basic moral term that is definitely implausible. Still, it does not follow that no plausible naturalistic definition either of that, or of any other, basic moral term exists. Nor does this stronger claim play any explicit role in Ross’s argument, either at this, or at any other point in his response to Spencer.¹⁰ So even if Ross does implicitly rely on a diagnosis of the ‘naturalistic fallacy’ in his response to Spencer’s evolutionary ethics, that diagnosis is not at centre stage in

---

¹⁰ Where Moore’s favoured ethical primitive is ‘good’, Ross favoured ethical primitive is ‘right’, or ‘duty’.

¹⁰ On the next page, Ross continues: ‘But even if we admit that the characteristics of being highly evolved, of tending to promote the maximum good, and of being right tend to a large extent to go together, we must surely recognize a closer relation between the first [i.e. being highly evolved] and the third [i.e. being right]. It will not be because of the merely historical fact that they come later in the course of evolution, but because they share in a characteristic common to the later stages of evolution, the characteristic of being promotive of wider good, that acts will tend to be right’ (Ross 1939, 14). The claim now is one about the explanatory relation between the good, the right, and being more evolved. It is the claim that *if* acts that come later in evolution tend to be right, *then* that is because such acts tend to be promotive of the wider good. It is not the claim that ‘the wider good’ cannot be defined in terms of some other property, natural or non-natural.
that response.

Of course, Ross did not claim to have diagnosed ‘the naturalistic fallacy’, Moore did. So it is natural to think that it is to his work we should go in order to find this diagnosis being explicitly applied against any claim on behalf of the moral significance of the moral sciences. Yet what we find, both in *Principia Ethica* and in the later *Ethics*, is both quite different and independently instructive. Here are some of Moore’s remarks on Spencer’s evolutionary ethics:

> It might… be held that the direction in which living things have hitherto developed is, as a matter of fact, the direction of progress… It may be held that the more evolved, though not itself the better, is a *criterion*, because a concomitant, of the better… Finally… it may be held that, though Evolution gives us no help in discovering what results of our own efforts will be best it gives some help in discovering what it is *possible* to attain and what are the means of its attainment.

(Moore 1903, 107)

So far, Moore seems happy to play along. Yet he is not very impressed by the apparently alleged implications of these claims either for the nature of the good, or for what actually is good. He continues:

> In the mere fact, then, that these non-fallacious views of the relation of Evolution to Ethics would give so little importance to that relation, we have evidence that what
is typical in the coupling of the two names is the fallacious view to which I propose to restrict the name ‘Evolutionist Ethics’. This is the view that we ought to move in the direction of evolution simply *because* it is the direction of evolution. (Moore 1903, 107-8)

Note Moore’s use of the term ‘because’ in this passage. This is not plausibly read as stating any identity. So far, then, there is no mention of any ‘naturalistic fallacy’ being committed. True, ‘the naturalistic fallacy’ does appear in Moore’s diagnosis of what might have lead the evolutionary ethicist to make their claim about how moral claims are related to claims about evolution. Yet even here, Moore stops short of accusing Spencer of actually committing this fallacy. Instead, Moore suggests that Spencer is not really an ‘evolutionary ethicist’ in any proper sense at all. This is what he says:

> It is plain, then, that Spencer *identifies* the gaining of ethical sanction with being more evolved: this follows strictly from his words. But Mr Spencer’s language is extremely loose, and we shall presently see that he seems to regard the view it here implies as false. We cannot, therefore, take it as Mr Spencer’s definite view that ‘better’ means nothing but ‘more evolved’; or even that what is ‘more evolved’ is *therefore* ‘better’. But we are entitled to urge that he is influenced by these views, and therefore by the naturalistic fallacy. (Moore 1903, 100-101)

The closest Moore gets to using his diagnosis of ‘the naturalistic fallacy’ against Spencer in this passage is to accuse him of being ‘influenced’ by it. This simple fact is itself
headline news in the present context, and a good reason to consider the possibility that there is something else, of equal philosophical importance, going on. By looking first (in this section) at how Ross makes use of Moore’s argument against Spencer in his own response to evolutionary ethics a few decades later, and second (in the next section) at Ross’s response to French sociology, I shall now argue that there is.

There is noticeable overlap between Moore and Ross with respect to which aspects of Spencer’s work they address in their respective responses to his defense of evolutionary ethics. For example, at one point Moore directly quotes a number of passages from Spencer’s *The Data of Ethics*, where Spencer writes:

Yes, there is one postulate in which pessimists and optimists agree. Both their arguments assume it to be self-evident that life is good or bad, according as it does, or does not, bring a surplus of agreeable feeling. (Spencer 1894, Sect. 10)

And a bit later:

No school can avoid taking for the ultimate moral aim a desirable state of feeling called by whatever name — gratification, enjoyment, happiness. Pleasure, somewhere, at some time, to some being or beings, is an inexpugnable element of the conception. (Spencer 1894, Sect. 16)

Moore’s response is as follows:
Mr Spencer himself tells us his ‘proof’ is that ‘reversing the application of the words’ good and bad… ‘creates absurdities’ (Sect. 16)… So… he is… a naturalistic Hedonist. (Moore 1903, 104-105)

Moore’s point here is twofold. First, Spencer is not, at bottom, an ‘evolutionary ethicist’ in Moore’s sense, as much as a (naturalistic) hedonist about the good, who thinks that evolution proceeds by producing more pleasure, and therefore more good on the whole. Second, the fundamental case for Spencer’s hedonism is not that it somehow accords with the direction of evolution, but rather that its denial ‘creates absurdities’. This is most plausibly read as an *a priori* claim about which ideas can be coherently affirmed together, and is therefore one that can be made from the comfort of the philosopher’s armchair.10

The crucial premise in this argument is a commitment to a distinctive view about the foundations of moral knowledge.11 Moore formulates this view as follows:

---

10 On exactly these grounds (i.e. the non-creation of absurdities), Moore rejects hedonism, not because he can prove its falsehood (he doesn’t think there can be any proofs at this fundamental level), but because ‘it contradicts other propositions which appear to be equally true’ (Moore 1903, 145). Thus, if Spencer is ever guilty of committing ‘the naturalistic fallacy’, it is here: according to Moore he wrongly supposes that the question ‘But is it good?’ is ‘closed’ once it is agreed that the ‘it’ in question involves pleasure. 11 According to Moore, moral ‘intuition’ can furnish a reason for *holding* a proposition to be true, at least when that proposition is self evident, i.e. when ‘there are no reasons which prove its truth’ (Moore 1903, 144). For critical discussion of Moore’s conception of self evidence, see Audi 1996. On Audi’s less restrictive (and in my view less implausible) definition, a self evident truth is a claim such that a) adequately understanding it entails being justified in believing it, and b) believing it on the basis of adequately understanding it entails knowing it. For further discussion of the notion of self evidence and how it relates to that with which we are acquainted, see the Appendix below. See also Skorupski MS.
We cannot tell what is possible, by way of proof, in favour of one judgement that ‘This or that is good’... until we have recognised what the nature of such propositions must always be. In fact, it follows from the meaning of good... that such propositions are all of them... ‘synthetic’: they all must rest in the end on some proposition which must be simply accepted or rejected, which cannot be logically deduced from any other proposition. This result... may otherwise be expressed by saying that the fundamental principles of Ethics must be self evident... The expression ‘self evident’ means properly that the proposition so called is evident or true, by itself alone; that it is not an inference from some proposition other than itself. (Moore 1903, 143)

Ross’s later discussion in *The Foundations of Ethics* makes exactly the same move as Moore’s in this respect. Ross claims that Spencer’s ‘fundamental ethical theory’ is that ‘it is conduciveness to pleasure that is... the real ground of rightness’; that he believes that ‘life always contains a surplus of pleasure over pain; and that conduciveness to life and conduciveness to pleasure always go together’; that therefore ‘[h]is fundamental theory turns out to be universalistic Hedonism, or Utilitarianism’; and that consequently evolutionary ethics ‘need not be examined as a separate form of theory regarding the ground of rightness’ (Ross 1939, 59). Once more, then, Spencer is at bottom really a

---

12 The intimate connection between Moore’s non-naturalism and his *aprioristic* moral epistemology might be thought to be evident in at least one of the sentences in this passage (i.e. the one that refers back to his argument about the meaning of ‘good’). The fact that these two views are intimately connected does not undermine the claim that they are logically independent.
hedonist, and is not (in Moore’s terms) a genuine evolutionary ethicist at all.¹³

Although he follows Moore very closely, Ross begins his quotations from Spencer a bit earlier than Moore. The first passage quoted by Ross from Spencer is this:

> Is there any postulate involved in these judgements of our conduct? Is there any assumption made in calling good the acts conducive to life, in self or others, and bad those which directly or indirectly tend towards death, special or general? Yes; an assumption of extreme significance has been made – an assumption underlying all moral estimates… (Spencer 1894; quoted in Ross 1939, 26-7)

The assumption in question (and the assumption mentioned in the title of this chapter) is that something (namely pleasure) is inevitably connected with things that are ‘conducive to life’, and that this thing (i.e. pleasure) is what good is. This is the claim that allegedly stands as the foundation of our moral knowledge. It is a claim the truth or falsity of which, according to Ross, is assumed on both sides of the debate to be self evident a priori. No development in the moral sciences therefore has the potential to change our fundamental entitlement to this claim one way or another. With respect to this fundamental truth of morality, the moral science, be it evolutionary or sociological, is ‘simply beside the mark’ (Ross 1930, 15).

¹³ Moore’s contemporary, and fellow intuitionist, Hastings Rashdall takes a similar swipe at Spencer in his 1907 work The Theory of Good and Evil, where he claims that ‘Morality essentially consists in the promotion of a good or ideal of life, the nature of which is discerned by our rational judgements of value’, at least some of which are self evident a priori. (Rashdall 1907, 401. See also Marett 1902, 246).
3. Ross on debunking explanations

This epistemological reading of the Moore and Ross’s response to the moral sciences of their time is reinforced by Ross’ discussion, in *The Right and the Good*, of what is nowadays called a ‘debunking argument’ against our possession of moral knowledge. This discussion includes the following oft-cited passage, with which many contemporary moral philosophers will be familiar. Ross says:

… the nature of the self-evident is not to be evident to every mind however undeveloped, but to be apprehended directly by minds that have reached a certain level of maturity… (Ross, 1930, 12)

This quotation has sometimes been used to illustrate the fact that self evident truths need not be obvious, or be immediately apprehended as such (c.f. Audi 1996). Yet in the context of the discussion in *The Right and the Good*, Ross’s aim can be shown to be much more ambitious. Here is the quotation again, accompanied by its surrounding argument:

… if human consciousness is continuous, by descent, with a lower consciousness which had no notion of right at all, that need not make us doubt that the notion is an ultimate and irreducible one, or that the rightness (*prima facie*) of certain types of act is self-evident; for the nature of the self-evident is not to be evident to every
mind however undeveloped, but to be apprehended directly by minds that have reached a certain level of maturity [my italics], and for minds to reach the necessary degree of maturity the development that takes place from generation to generation is as much needed as that which takes place from infancy to adult life. (Ross, 1930, 12)

So why does Ross, the moral philosopher, think it is necessary to consider the possibility that ‘human consciousness is continuous, by descent, with a lower consciousness which had no notion of right at all’? What concerns him in this passage, close to the beginning of his best known ethical work, is that all our claims to moral knowledge are based on contingent and shaky foundations. He is concerned to rebut the challenge that knowledge of the causes and functions of our moral beliefs will effectively debunk those beliefs, in virtue of the contents of those beliefs being at best accidentally related to their justificatory, or veridical, grounds.

Like many of his philosophical predecessors who were also concerned about the ‘origin’ of our moral faculty, Ross is very clear that he accepts the potential both for debunking, and for vindicating, explanations of moral beliefs, whether in moral or non-moral terms. He writes:

An inquiry into the origin of a judgement may have the effect of establishing its validity… If we find that the pre-existing judgements were really an instance of knowing, and that the inferring was also really knowing… our inquiry into the
One potential case is human knowledge with a divine source (c.f. Lillehammer 2010). Another is justified reliance on testimony with a human source. A third (and possibly the one that Ross has in mind) is inferential knowledge based on previously grasped self-evident truths. Yet Ross’ immediate concern in this passage are the prospects for a debunking explanation of our moral beliefs based on specific empirical claims made by the moral sciences of his day (c.f. Levy-Bruhl 1905, 70; 166; Durkheim 1993, 33) The general phenomenon is one that will be familiar to contemporary readers as an epistemic defeater in the form of ‘tracking failure’, or responsiveness to (ethically or epistemologically) ‘irrelevant factors’ (cf. Lillehammer 2010; Berker 2009). Ross writes:

… if anyone can show that A holds actions of type B to be wrong simply because (for instance) he knows such actions to be forbidden by the society he lives in, he shows that A has no real reason for believing that such actions have the specific quality of wrongness, since between being forbidden by the community and being wrong there is no necessary connexion… He does not, indeed, show the belief to be untrue, but he shows that A has no sufficient reason for holding it true; and in this sense he undermines its validity. (Ross 1930, 14)\textsuperscript{18}

\textsuperscript{18} In his discussion of Spencer’s view in \textit{The Foundations of Ethics}, Ross also considers a metaphysical debunking argument against the very existence of moral obligation, along the following lines: ‘May the upshot of the evolutionary account be… that there is no such thing as obligatoriness; that there is \textit{nothing} in reality answering to the meaning which we have in mind when we use the word obligatory, the only distinction that remains being that between less and more evolved acts?’ (Ross 1939, 15). Ross’s claim
In other words, regardless of the truth-value of a given moral belief, our warrant for that belief can be undermined by showing that our possession of that belief has the wrong kind of relation to what makes it either true, or justified.  

Ross denies that this possibility suffices to undermine our claim to have at least some moral knowledge. His position, in a nutshell, is this: basic moral knowledge (in the form of *prima facie* principles) is *a priori* accessible as self-evident to minds at a certain stage of development, and therefore available regardless of any *a posteriori* knowledge those minds may or may not have of their own developmental history, the social function of their moral sensibility, or other facts about the causes of their moral beliefs. Furthermore, the human mind has reached a stage of development at which basic moral knowledge is, in fact, *a priori* accessible as self-evident. In fact, we actually demonstrate that we have access to such knowledge whenever we exercise the relevant capacity (e.g. when we correctly rule out certain moral claims as absurd or implausible). Ross writes:

... the human mind... is competent to see that the moral code of one race and age is in certain respects inferior to that of another. It has, in fact, an *a priori* insight into certain broad principles of morality, as it can distinguish between a more or less

---

that we have *a priori* access to basic moral knowledge in the form of self evident moral principles provides him with a response to this form of debunking argument also.

19 It might plausibly be objected that Ross is overstating the case when he says that justification requires there to be a ‘necessary’, as opposed to a ‘reliable’ connection between the causal and justificatory grounds of beliefs. It might also be questioned whether reliable tracking is everywhere a necessary condition for knowledge or justified belief. I pass over these complications in what follows. For discussion, see e.g. Enoch 2012).
adequate recognition of these principles… [T]here is a system of moral truth, as objective as all truth must be, …and from the point of view of this, the genuine ethical problem, the sociological inquiry is simply beside the mark. (Ross 1930, 14-15)\textsuperscript{20}

So what ultimately protects our basic moral beliefs from potential debunking is that at least some of our moral beliefs are responsive to self-evident principles that we know \textit{a priori}. Not only do we not need the \textit{a posteriori} discoveries of moral sciences like sociology to access these truths, no such discoveries are able to undermine our actual entitlement to them.\textsuperscript{21}

It is worth briefly stating how modest Ross’s claim actually is before concluding the historical part of this chapter. First, and as Ross would be the first to admit, from the fact that some of our basic moral beliefs are immune to debunking explanations, it does not follow that all, most, or even many of them are. For example, Ross makes no analogous claim on behalf of moral beliefs inferred from basic principles via \textit{a posteriori} linking

\textsuperscript{20} At this point, Ross and the French sociologists definitely part company (See e.g. Levy-Bruhl 1905, 122; 204; Durkheim 1993, 130).

\textsuperscript{21} In his 1907 discussion of Spencer, Rashdall considers and rejects a potential debunking argument for moral and other beliefs, claiming that ‘Spencer’s theory involves us in a hopeless scepticism, as does every theory which attempts to account by experience for the principles of thought, which are implied in every step of the process by which experience is turned into knowledge’ (Rashdall 1907, 371). Rashdall’s response is based on an analogy between moral norms and other ‘laws of thought’, such as the law of contradiction; an argument he attributes to Cook Wilson. He writes that ‘If these laws are really no necessities of thought but mere inherited results of accidental experiences, it is possible that they are untrustworthy’ (Rashdall 1907, 370). Discussion of this argument and the various comparisons on which it depends would take me to far afield here (c.f. Ross 1939, 15-19).
premises. Nor (as we have seen) does he claim that all thinking creatures have infallible insight into which moral principles are, in fact, self-evident. In both cases, someone’s firm conviction about the privileged status of one moral belief or another could in principle be the target of a successful debunking argument. Second, Ross is obviously aware of the epistemological challenge that has traditionally been thought to arise from the fact that there is widespread disagreement on moral questions both ‘in different societies’ and ‘within the same society’; including ‘a real difference of opinion as to the comparative worth of different goods’ (Ross 1939, 17; 19). In response, Ross claims that ‘on examination’ such disagreement (at least among ‘sufficiently mature’ minds) can be shown to depend ‘not on disagreement about fundamental moral principles, but partly on differences in the circumstances of different societies, and partly on different views which people hold, not on moral questions but on questions of fact’ (Ross 1939, 18). The fact of widespread moral disagreement, Ross thinks, ‘should weaken perhaps our confidence in our own opinions, but not weaken our confidence that there is some opinion that would be true’ (Ross 1939, 19).

4. Back to the present

So what (if anything) has this got to do with us? Ross’s response to debunking explanations of moral beliefs is a response to an argument of the following form:

The psychological states that give rise to moral response M respond to environmental factors of type N; there is no reliable connection between
environmental factors of type N and the appropriateness conditions of M; So, the psychological states that give rise to moral response M respond to environmental factors that are not reliably connected with the appropriateness conditions for M; So, moral response M fails to have any genuine normative force.\textsuperscript{23}

This is an argument that will be familiar from discussions of debunking explanations of moral beliefs and other ethically significant attitudes in contemporary moral philosophy (c.f. Joyce 2005; Singer 2005; Street 2006; Berker 2009). The response given by Ross to instances of this argument form arising from the moral sciences of his time is of basically the same kind as the response given by a number of contemporary philosophers to instances of this argument form arising from the moral sciences of our own time. One example will suffice to make the point. In his 2009 paper ‘The Normative Insignificance of Neuroscience’, Selim Berker aims to defuse a number of debunking arguments targeted at a limited range of commonsense moral intuitions (more specifically: deontological commonsense moral intuitions), and put forward in recent work by Joshua Greene, Peter Singer and others (e.g. Singer 2005; Greene 2007). The gist of these arguments is that if we reflect rationally on the causes of our moral intuitions we will be forced to abandon a significant range of moral beliefs at the heart of common sense morality (more specifically: a range of beliefs traditionally associated with deontological, as opposed to consequentialist moral theories).\textsuperscript{24}

\textsuperscript{23} The term ‘appropriateness condition’ as used here is meant to be neutral between questions of truth and questions of justification. Nothing substantial in what follows hinges on this piece of terminological convenience.

\textsuperscript{24} The scope of the debunking argument considered by Ross is the set of our moral beliefs as a whole. The scope of the debunking argument considered by Berker is some subset of our moral beliefs, or moral intuitions, of a broadly deontological nature. The two
arguments considered, and rejected, by Berker is an instance of the argument form stated in the paragraph above. The particular instance of this argument form considered and rejected by Berker proceeds as follows:

The emotional processing that gives rise to deontological intuitions responds to factors that make a dilemma personal rather than impersonal; the factors that make a dilemma personal rather than impersonal are morally irrelevant; So, the emotional processing that gives rise to deontological intuitions responds to factors that are morally irrelevant; So, deontological intuitions, unlike consequentialist intuitions, do not have any genuine normative force. (Berker 2009, 321)

Berker calls this argument ‘the argument from morally irrelevant factors’. The key premise of the argument, he argues, is the second (in italics above). About this premise, he makes the following comment:

---

arguments are nevertheless importantly connected, for at least three reasons. First, if Ross’s response to the arguments he considers is successful, then this may provide a vindication of some of the deontological intuitions targeted in the argument considered by Berker. (Ross argues that some deontological prima facie principles are a priori self evident.) Second, if there is no response to the arguments considered by Ross, then it is up for grabs whether any of the intuitions considered by Berker and Greene are safe from debunking (this is one way of reading Singer’s challenge in Singer 2005 and de Lazari-Radek and Singer 2012). Third, on at least one way of reading the argument considered by Berker, the reason why consequentialist intuitions are safe from debunking is that they can be vindicated on Rossian terms (this is arguably the most plausible way of reading Singer’s response to his own challenge in the above cited works). Another way in which the arguments considered by Ross and Berker differ is with respect to the centrality of moral principles in the beliefs or intuitions that are said to be evaluable from the armchair. Ross thinks it is (prima facie) moral principles that are a priori self evident. Berker makes no such claim.
The first thing to note is that… [this] premise… appeals to a substantive normative
intuition, which presumably one must arrive at from the armchair, rather than
directly read off from any experimental results; this is why the argument does not
derive an ‘ought’ from an ‘is’. I believe that this feature is a virtue of the argument;
however, it is also its ultimate undoing. (Berker 2009, 322)

It is its ultimate undoing because, according to Berker, ‘the appeal to neuroscience is a
red herring’. All the real work in the argument is done by the ‘identification, from the
armchair’, first of ‘the distinction between dilemmas-eliciting-deontological-reactions
and dilemmas-eliciting-consequentialist-reactions with the distinction between personal
and impersonal moral dilemmas’, and second, of ‘a substantial intuition about what sorts
of factors out there in the world are and are not morally relevant’; the mere distinction
between something being ‘personal’ or ‘impersonal’ in the relevant sense not being one
of them (Berker 2009, 326). The ‘basic problem’ with this argument, according to
Berker, is that ‘once we rest our normative weight on an evaluation of the moral salience
of the factors to which our deontological and consequentialist judgements are responding,
we end up factoring out… any contribution that the psychological processes underlying
those judgements might make to our evaluation of the judgements in question.’ (Berker
2009, 326). To put the point in Ross’s words, from the point of view of the substantially
moral question, the neuroscientific inquiry is ‘simply beside the mark’ (c.f. Ross 1930,
15).

25 As is clear from this passage, Berker would appear to inherit from his intuitionist
ancestors not only the basic structure of their response to the debunking challenge, but
also their close association between what is accessible from the armchair (or a priori) and
what is irreducibly normative, and therefore allegedly non-natural.
In sum, both Ross and Berker maintain that the crucial premise in the debunking arguments they respectively consider is the second (in italics above). Both Ross and Berker claim that the defence of this premise relies on normative intuitions that are available from the armchair. Both Ross and Berker therefore claim that the empirical evidence cited in the debunking argument is a red herring: all the real argumentative work is done by means of ‘armchair theorizing’ (in Ross’s case, *a priori*) about whether the features of the world to which our moral sensibility is responding are, or are not, morally ‘relevant’. It follows that the response to the moral sciences given by armchair moral philosophers like Berker and others at the start of the Twenty-first Century is for practical purposes the same that that given by the classical intuitionists at the start of the Twentieth Century. So what (if anything) has changed?

5. *Armchair or garden chair?*

To begin answering this question, it is necessary to ask if there is any plausible combination of metaphysical and epistemological claims that could secure the ‘armchair’ response to the debunking challenge, short of a version of classical intuitionism that grounds our moral knowledge in our access to *a priori* self evident truths or principles. The range of options is arguably quite restricted. For example, any option that falls short of entailing that our entitlement to moral claims the truth of which is accessible from the armchair is asymmetrically privileged with respect to our entitlement to moral claims the truth of which is not thus accessible will be consistent with the possibility that reflection
on facts not accessible from the armchair (such as the actual causes of moral intuitions or beliefs) has the power to debunk at least some of the moral claims the truth of which is in principle accessible from the armchair (including, potentially, some deontological intuitions). To claim that such empirically informed reflection would not, in fact, undermine some particular intuition or belief is to take an inductive (if normatively laden) bet on the rational outcome of such a process, the *a posteriori* inputs to which cannot, by hypothesis, be assumed to be ‘simply off the mark’. In saying this, I do not want to suggest that a classical intuitionist conception of moral knowledge as based on our *a priori* grasp of self-evident truths or principles is the only option available to those who would hope to dismiss the normative significance of neuroscience, sociology, evolutionary psychology, or any other part of the moral sciences. I do, however, suspect that some very influential parts of the history of Twentieth Century moral philosophy can be shown to manifest a notable tendency in exactly that direction (See e.g. Audi 1996). But so what? Why should contemporary armchair philosophers worry about taking on the commitments of classical intuitionism?

My guess is that many contemporary philosophers who are impressed by recent progress in the human sciences would want to avoid a commitment to the strongly foundationalist intuitionism espoused by Ross and Moore. After all, there are well known objections to historical attempts to secure our claim to moral knowledge by grounding it in our grasp of *a priori* self-evident truths or principles. One of these objections (and one of which the classical intuitionists were obviously aware) is that even if there are some *a priori* self evident moral truths (e.g. ‘You should promote the good, whether in yourself or others’),
these truths could be no more than vacuous ‘sham axioms’, in Sidgwick’s phrase (c.f. Levy-Bruhl 1905, 173; Ross 1939; Lillehammer 2010). A second (and related) objection is that all substantive moral truths (e.g. ‘justice is fairness’) are the results of contingent social construction, and so at best instances of what has come to be known as a ‘historical’ a priori, i.e. claims a commitment to which may constitutive of some historically actual, but not every possible, moral epoch or sensibility (c.f. Levi-Bruhl 1905; Hacking 2002). I do not propose to pursue these objections further here. Instead, I shall consider what, if anything, might lead the contemporary armchair philosopher to stop short of a commitment to the ‘holy grail’ of a priori self-evidence, while nevertheless holding out for an ‘armchair’ response to potential debunking explanations motivated by claims made by the moral sciences, such as sociology, psychology or evolutionary biology.

I can think of at least three problems that might have worried someone like Ross in face of the rejection of his claim that moral knowledge can be grounded in a priori self-evident principles. In each case, I think many contemporary philosophers would respond that there is less room for concern than Ross may have thought, and therefore more reason to formulate the response to traditional debunking challenges in less epistemologically demanding terms.

First, it might have been worried that in the absence of a foundation in a priori self-evidence, our claims to moral knowledge would extend only to contingent truths, thereby undermining the widely accepted belief that some moral claims are necessary. On this
point, the intervening period between the time of the classical intuitionists and the present time has made a substantial difference to how the claim that moral truths have one modal status rather than another is likely to be understood. No serious philosopher after Putnam and Kripke is going to assume without argument that all necessary truths are knowable \textit{a priori}. If it were actually true that goodness is pleasure, then it could be necessarily true that goodness is pleasure, provided that terms like ‘goodness’ and ‘pleasure’ are interpreted as rigid designators, and therefore refer to the same property in all possible worlds (cf. Kripke 1980). The fact that pleasure is what ‘good’ rigidly designates could be something we are only able to settle \textit{a posteriori}, for example on the basis of an evolutionary or neuroscientifically informed moral science.\footnote{The claim that ‘good’ functions as a rigid designator in the relevant sense could still, of course, be something that we are able to discover from the ‘armchair’ (c.f. Jackson 1998). Yet this issue is equally controversial.}

Second, it might have been worried that in the absence of a grounding in \textit{a priori} self evidence, there is no way to account for how we can have knowledge of anything genuinely \textit{normative}.\footnote{The claim that all normative knowledge is basically \textit{a priori} is at least comparatively less controversial than the claim that all \textit{a priori} knowledge is basically normative. For a defence of the latter claim, see Skorupski 2011.} This worry embodies an assumption that many contemporary philosophers would reject, namely that all basic normative truths (or truths about ‘reasons’) are \textit{a priori}. Yet even if we accept this controversial claim, we do not thereby have to accept that all basic normative truths are self evident. We would only be forced to this conclusion if there were no other way to show that our basic moral knowledge is \textit{a priori} (such as an \textit{a priori} grounding of moral knowledge in a privileged set of explanatorily coherent principles). It is not at all obvious that we must accept this
assumption. If we do not, then we have no overwhelming reason to accept the claim that our basic moral knowledge requires *a priori* self evidence in the sense defended by the classical intuitionists.

Third, it might have been worried that in the absence of a foundation in *a priori* self-evidence, our moral beliefs would be unjustified. This concern arguably embodies at least one of two assumptions that many contemporary philosophers would reject. The first is the ‘aprioristic’ assumption that no self-evident moral truths are *a posteriori*. This is a controversial assumption that even some intuitionists would deny (consider some of the obviously unacceptable things that actual people have found it possible to do to each other during the course of human history. The list is virtually endless.). The second is the ‘foundationalist’ assumption that the only way for our moral beliefs to be justified is for them to be derivable (either directly or indirectly) from a set of basic moral principles, or axioms, on which they asymmetrically depend. This is also a controversial assumption that even some intuitionists would deny (e.g. because they would claim that the set of basic moral truths is not capable of any interesting and psychologically realistic axiomatization). The denial of either of these assumptions does, however, come at a cost. The cost of giving up the first assumption is that a certain amount of pressure is put on the idea that all basic moral knowledge is accessible from the armchair. The cost of giving up the second assumption is that a certain amount of pressure is put on the idea that we are able to identify a set of comparatively simple moral beliefs on which the rest of our moral knowledge somehow ‘depends’. Either way, a certain amount of pressure is put on the idea that justified moral beliefs arrived at from the armchair are immune to
debunking explanations informed by moral beliefs that can only be only arrived at from a standpoint beyond its reach. For some contemporary philosophers, this possibility may provide just enough motivation to leave the sitting room and move into the garden. Although I do not propose to argue for this claim here, I believe that doing so does not seriously endanger our entitlement to at least some of our substantial moral beliefs (c.f. Lillehammer 2010). Just because a certain range of moral claims that were previously thought to be empirically intractable are no longer considered to be beyond all reasonable doubt from the comfort of the armchair, it does not follow that the same range of claims would fail to pass the same test from the comfort of the garden chair. If so, there may be room for a moderate position in moral epistemology that is robust enough to withstand the more radical challenges of \textit{a posteriori} debunking without relying on a flat-footed \textit{a priorism} about moral knowledge.

\textit{Appendix: one way to read the Open Question Argument}

One of the more puzzling aspects of Twentieth Century moral philosophy is how Moore’s main argument for ethical non-naturalism could have been so massively influential if, as many subsequent critics (including Ross) have pointed out, it is so obviously implausible on its own terms. Thus, in the last few decades, the existence of \textit{a posteriori} identities has frequently been cited as one of the two main objections to Moore’s so-called ‘open question argument’ for ethical non-naturalism, according to which the property of goodness cannot be identified with any natural property (or any other property for that matter) because whichever natural (or other) property (e.g. being
pleasurable or desired) we attribute to some object, it will always remain an ‘open question’ among competent speakers whether that property is good. When considered on the purely semantic and metaphysical terms on which it is often proposed, this objection may well be decisive. (I take no view on this matter here.) When considered in the light of the preceding discussion of Moore’s response to the moral sciences of his time, however, this objection is arguably misguided, at least on Moore’s own terms. Suppose that what Moore was looking for was not just some property with which ‘good’ is necessarily coextensive (or necessarily coextensive and grasped through some particular mode of presentation), but a property with which ‘good’ is necessarily coextensive (or necessarily coextensive and grasped through some particular mode of presentation), and its identification with which can play the psychological role of a basic foundational claim in a system of moral knowledge actually possessed by conceptually competent human adults. On that assumption, it is much less plausible to claim that any \( a \textit{posteriori} \) identity will do, at least if the discovery of that identity depends on a substantial amount of empirical theory that is not accessible from the ‘armchair’. On this reading, the \( a \textit{posteriori} \) identity objection, at least as it is normally formulated, is at best dubiously on target with respect what Moore’s ‘open question argument’ is meant to show.\(^{29}\)

Although I do not mean to claim that this reading of ‘the open question argument’ is the only plausible (or even the best) way to make sense of everything Moore says in

\(^{29}\) Even if we reject Moore’s ‘aprioristic’ epistemology, we could accept his ‘open question argument’ as proving that no naturalistic (or other) definition of ‘good’ is possible on his own ‘aprioristic’ terms. Our disagreement with Moore would now be one about his moral epistemology, and his view about how epistemology relates to semantics and ontology.
Principia Ethica, I do want to claim that there is some textual and contextual evidence that supports it. First, the 'property' targeted by the argument is the property of 'good', which stands at the very foundation of Moore’s ethical system, and therefore (by implication) what he thought of as the basic core of our actual moral competence. As Moore himself puts it, "'good' is the notion upon which all Ethics depends", where the dependence in question is most plausibly read as obtaining in the order of knowledge, and not necessarily in the order of being (Moore 1903, 142). Second, and as Thomas Baldwin has pointed out, Moore and his contemporaries were notoriously slippery in their use of terms such as 'notion', 'idea', 'concept', 'property', 'predicate' and 'object'; frequently sliding across the 'sense/reference distinction' (Baldwin 2004). The claim that the 'property' the term 'good' refers to is something that can form a part of a relation in extension in the way required for a purely metaphysical (as opposed to a conceptual) identity cannot, therefore be taken for granted on Moore’s own terms. Another possibility is that Moore is primarily looking for an a priori conceptual identity in the realm of 'notions' or 'ideas'. Third, the philosophy of language practiced by some of Moore’s contemporaries had a strongly epistemological element (c.f. Russell’s conception of logically proper names ('this' and 'that') as picking out the sense data of immediate experience (Sainsbury 1979)). Partly in light this fact, Baldwin interprets 'the open question argument' as having a non-eliminable epistemological element, when he writes

30 Moore defines the concept of ‘duty’ as follows: ‘Our ‘duty’… can only be defined as that action, which will cause more good to exist in the Universe than any possible alternative’ (Moore 1903, 148). Notoriously, he claims that ‘universal propositions of which duty is a predicate, so far from being self-evident, always require a proof, which it is beyond our present means of knowledge ever to give’ (Moore 1903, 181). In The Right and the Good, Ross objects to Moore’s definition, claiming that the question whether what is our duty is what will produce the most good is just as open as the question whether what is good is what will produce pleasure (Ross 1930). This definition of ‘duty' in terms of ‘the good’ is absent from Moore’s later treatment of the topic (c.f. Moore 1912)
that ‘Moore’s argument for the indefinability of good is focused on the unacceptable epistemological implications of ordinary analytic and theoretical definitions’, because ‘one primary purpose of a definition of good is… [that it] …should enable us to decide on non-ethical grounds what is good’ (Baldwin 2004, 322-3) (Baldwin 2004, 326).\footnote{Baldwin also writes that ‘anyone who accepts the definition in the context of the theory is committed to holding that the definition provides a new and superior epistemology for ethical questions’ (Baldwin 2004, 323), and that ‘a central element in Moore’s thesis of the indefinability of good is the epistemological claim that ethical reflection plays an inescapable role in moral judgement’ (Baldwin 2004, 324). These claims are both consistent with the reading of the ‘open question argument’ offered here, but stop short of attributing to Moore the stronger thesis that the definition in question should play the role of an \textit{a priori} self evident foundation for our moral knowledge.}\footnote{31} According to Baldwin, these epistemological assumptions are sufficient to justify Moore rejection of any definition of good in terms of an \textit{indefinite} disjunction of natural properties, although it leaves his argument undefended both against possible definitions in terms of an \textit{infinite, but well-defined} disjunction of natural properties (Baldwin 2004, 326), and against \textit{a posteriori} ‘natural kind’ style definitions (Baldwin 2004, 323). On the reading of Moore’s argument that I am considering here, the strongly epistemological constraints on an acceptable definition would allow him to reject these alternative definitions as well, because neither an infinite disjunction (however well-defined) nor an \textit{a posteriori} identity can arguably play the epistemologically foundational role that Moore is asking any putative definition of ‘good’ to play. On this reading of the argument, what Baldwin’s interpretation leaves out is the role played by the ‘aprioristic’ foundationalism that lies at the heart of Moore’s intuitionist project; a view which, although it is inextricably connected with Moore’s commitment to ethical non-naturalism, is logically distinct from it.
A reading of ‘the open question argument’ along the lines just proposed also provides the basis for a possible response on Moore’s behalf against the second apparently fatal objection often given to that argument, namely that he failed to realize the significance of the fact that there are non-obvious *a priori* truths, and that some naturalistic (or other) definition of ‘good’ could be one of those. On my reading of Moore’s argument, this objection is not so much wide of the mark as based on a misunderstanding of what Moore (rightly or wrongly) thought he could establish. One of the salient facts about the conceptions of self evidence espoused by the classical intuitionists is that there is no necessary connection between a truth being *self evident* on the one hand, and that truth being *obvious* on the other (c.f Sidgwick 1907; Ross 1930). What is crucial to the notion of self evidence is the idea of a truth that is either incapable of (e.g. Moore) or that stands in no need of (e.g. Bentham) proof, and our characteristic mode of access to which is in some way ‘direct’ (e.g. by way ‘acquantance’ as opposed to ‘description’, or by being present to the mind as a ‘clear and distinct’ idea). If Moore were operating with such a notion of self-evidence, then the mere fact that a definition of ‘good’ could be correct, be *a priori* knowable, but also non-obvious would present no fundamental difficulty for his view. On the present interpretation, Moore’s view is consistent with this fact. What he is claiming is that there is no definition of ‘good’ (including a complex *a priori* network analysis (see e.g. Jackson 1998)) that satisfies these criteria, while also being able to play the requisite epistemological role. Although Moore could reasonably be accused of having made this claim on a flimsy inductive basis, he cannot (on this reading) be

---

32 Tracing the history of assumed connections between self evident truths and knowledge by acquaintance would take me too far afield here. For further discussion, see e.g. Audi 2004.
accused of simply having missed the possibility that some *a priori* truths are far from obvious.

It has not been my intention in this Appendix to defend Moore’s ‘open question argument’, either in this or in any other form. Nor has it been my intention to recommend the epistemologically driven philosophy of language on which I suspect it may have been based. What I have tried to do is provide a contextually sensitive reading of the open question argument according to which it is less obviously implausible on its own terms than many contemporary presentations of the argument make it out to be. This reading of the argument gains further support from the fact, as illustrated in the main body of this paper, that the idea of *a priori* self evidence plays a central part elsewhere in Moore’s moral philosophy, as well as in the moral philosophy of his intuitionist contemporaries.

**References**


Press.


Spencer, H., 1894, *The Data of Ethics*. Cheap Edition (Or otherwise the seventh and eighth thousands.) Williams and Norgate.

