

# Methodological Naturalism in Metaethics

Daniel Nolan

Penultimate Draft. The final version is in McPherson, T. and Plunkett, D. (eds) 2017. *The Routledge Handbook of Metaethics*. Routledge, New York, pp 659-673. Please cite the published version.

## 1. Introduction

The philosophical movement of *naturalism* is normally divided into two strands. One is a methodological one: roughly, that philosophy (or some area of philosophy) should proceed using the methods of the sciences, particularly the natural sciences. The other is a metaphysical one: roughly, that the theory of the world that results from our philosophical inquiries should match, or at least not conflict with, what the sciences tell us there is. This chapter focuses on the first strand of naturalism. What does investigation of metaethics and ethics look like if we take a naturalist approach?

The focus of this chapter will be on methodological naturalism as an approach to metaethics. Part of metaethics, however, involves questions about the epistemology and method of ethics. These include questions about how we know what is morally right or wrong, good or bad, virtuous or vicious, etc., and what method we should use for determining answers to tricky questions about what we should do, how we should be, and what sort of world to aim for (morally speaking). Methodological naturalism about metaethical questions would typically go with methodological naturalism in the epistemology and method of ethics itself, though as we will see below one important strand of contemporary metaethical thinking treats *metaethics* as a naturalistic enquiry but does not treat *ethics* in this way.

A wide range of methodological approaches are claimed to be naturalistic by their proponents. Perhaps this is not surprising—there are many ways an inquiry can be like scientific inquiry, and when there is a lot of disagreement in metaethics, there is likely to be a lot of disagreement even among self-styled naturalists. Rather than trying to determine who are the *real* naturalists, it is more useful to look at the different ways different metaethical theorists take the guiding idea of naturalism in different directions.

Naturalism, of one sort or another, about the epistemology and method of ethical theorising is these days very widespread, but I think this state of affairs would have come as a great surprise to early analytic ethicists. In the first few decades of the twentieth century, one of the great divides in metaethics was between G.E. Moore and other non-naturalists, on the one hand, and the non-

cognitivists such as Wittgenstein, Ayer and Stevenson. Despite their deep disagreements, each camp was convinced that ethics was *not* like science, either in its methods or results.

Moore and other non-naturalists in his camp saw moral intuition as the core of ethical method though metaethics involved conceptual analysis as well: for example, Moore thought it was analytic that right action was a matter of maximising the good (Moore 1993 p 196). While it is somewhat controversial exactly how Moore saw moral intuition, it was clearly a very different process from scientific inquiry: intuitions were more like verdicts or seemings directly produced by reflection on scenarios, whether actual or imagined, than the results of theorising or observation.

While use of our supposed faculty of intuition was in the first instance a method for determining ethical questions rather than metaethical ones, it is natural to see it as indirectly providing some of the materials for metaethical conclusions as well. For example, why should Moore think that subjectivism about ethics is false? In part because he intuits that the amount of moral goodness of an act of feeding the hungry does not depend very much, or perhaps at all, on who approves of it. Likewise for intuitions about a range of other scenarios where the difference between cases is only who judges things are good or bad, or who approves or disapproves. (Moore would have likely found this pattern even in scenarios where Moore himself has different attitudes. To adapt a case of Russell's, Moore would have likely intuited that the suffering in bull-fighting is bad even where, counterfactually, Moore himself approves of it.) A good explanation of this pattern (given intuitionist assumptions) is that the goodness of an outcome does not depend very much on who takes it to be good, or who approves of it, and so appears to count against subjectivism.

Non-cognitivists, on the other hand, were equally convinced that ethics was not like science, but for quite different reasons. They rejected a non-naturalistic faculty of moral intuition, along with the idea that there were a special range of moral facts to be discovered. Instead, moral attitudes and moral disagreements involved something other than competing opinions about the facts: clashes of attitudes, perhaps, or conflicting imperatives. One illustration of the methodological anti-naturalism of non-cognitivists is found in an influential paper by Charles Stevenson. Stevenson endorsed a strong form of naturalism about every factual question, but because, in his view, moral disagreements involved clashes in desires as well as beliefs, "scientific methods cannot be guaranteed the definite role in the so-called normative sciences that they may have in the natural sciences" (Stevenson 1963 p 8). The suggestion that ethics is not concerned with matters of fact, particularly the sort discoverable by the sciences, is an

even stronger motif in Wittgenstein's *Lecture on Ethics*: a complete scientific book of the world would "simply be facts, facts, and facts but no Ethics" (Wittgenstein 1965 p 7).

As the twentieth century went on, the idea that good ethical and metaethical inquiry had little in common with the sciences was more widespread even than those who signed up to either Moorean non-naturalism or the non-cognitivism of Ayer, Stevenson, Wittgenstein, etc. Traditional methods of ethical inquiry, both in metaethics and in first-order ethics, included at least three important components. The first was the use of philosophical and moral *intuitions*, direct seemings that perhaps required some experience to prepare an intuiter, but did not seem to require any scientific investigation to produce or evaluate. The second was the use of *conceptual analysis*, which seemed to involve nothing more than *a priori* reflection either on the meaning of what we say or the concepts we employ in ethical thought. Later in the twentieth century, a third method, the method of *reflective equilibrium*, became popular: this was seen largely as a matter of harmonising the moral principles we accept with the mass of individual moral judgements we accept, together with various other philosophical commitments we may have. (Reflective equilibrium was most famously presented by Rawls 1971.) This coherentist method, especially if the inputs were delivered by intuition or proposed conceptual analyses, again seemed very different from the methods of e.g. chemistry or demography.

These methods were not just central to how first-order ethics was practiced, but in metaethical discussions as well. It is easy to find entire metaethical works with very little discussion of scientific questions, but plenty of arguments that look like appeals to conceptual analysis, ethical and metaethical intuitions, and attempts to harmonise clashing parts of our ethical and metaethical starting points. One influential example of this sort of work in metaethics is Michael Smith's *The Moral Problem* (1994), which on the face of it at least proceeds in a very different way from typical scientific investigations.

There are some clear reasons naturalism about metaethical method is attractive, despite it having been far from orthodoxy at times in the past century. The sciences, particularly natural sciences like physics and biology, have discovered all sorts of facts that were not even dreamed of a hundred years ago. Furthermore, science is good at producing rational agreement: while there will always be scientific disputes, few seriously doubt e.g. that there is oxygen in the air or that antibiotics can cure bacterial infections. If we could turn these very successful methods to metaethical and ethical questions, we might hope for some of the same progress in discovery and in resolution of disagreements. Another reason to hope that naturalistic methods will be useful in metaethics is possessed by people who are

*holists* about inquiry. Answers we give in one area of inquiry often seem to have consequences in other areas of inquiry, and any particular piece of evidence we get can be handled by modifying different parts of our overall theory. If our picture of the world faces the tribunal of experience as a whole, as for example Quine would have it (Quine 1951 p 38), then it seems reasonable to think that our methods for going from our experience to our theories should have important similarities across subject matters. This suggests that our methods in metaethics and should not be too different from our methods in science. This line of thought will be particularly pressing for those who are naturalists about other parts of philosophical inquiry. So much of metaethics tackles questions about ethical language, or epistemology of ethics, or the metaphysics of ethics, or questions about moral judgements and moral reasoning that if we are naturalists about philosophy of language, epistemology, metaphysics, and philosophy of mind, naturalism about metaethics is virtually forced on us.

A third reason may stem from metaphysical naturalism. If we became convinced e.g. that all phenomena in the world were ultimately physical, or that our nature was exhausted by biological and social features of us, we might come to expect that any ethical truths about our situation would also ultimately have a physical or biological or social basis. And if they do, that might suggest they are best investigated by inquiries like those we use for physical or biological or social facts. Of course, establishing this as a reason for methodological naturalism in ethics would require an argument that facts about us are ultimately physical, or biological, or social (or a combination): but arguments for physicalism and other forms of methodological naturalism are familiar from other areas of philosophy, including metaphysics, philosophy of science, and the philosophy of mind.

Pursuing metaethics and ethics in a naturalistic spirit may be appealing: but how, exactly, ought we bring to bear the methods of the natural and social sciences? One way is through critique. Standard methods of ethics can be assessed for how naturalistic they are, and to the extent they are not a methodological naturalist may wish to modify, supplement, or even reject them. Another way is through developing positive naturalistic metaethical research projects employing scientific methods. Before discussing critiques and naturalistic projects, however, it will be worth doing some ground-clearing, to see where methodological naturalism may go beyond uses of science that virtually everyone agrees are relevant to ethics, and distinguishing naturalism approaches to metaethics from naturalism about some other topics.

## 2. *The Role of Science in Ethics: Preliminaries*

Everyone (or almost everyone) thinks that scientific methods are *somewhat* relevant to resolving ethical decisions. When deciding whether to administer a treatment to a patient, a doctor needs to take into account scientific evidence about what the likely outcomes of that treatment will be. When a government is contemplating a change in social welfare payments, one thing they should consider is the economic implications: what will this change do to inflation, employment, food availability to those affected, etc. Sometimes estimating these impacts will require specialised economic knowledge. When deciding what countries or individuals should do about climate change, the science of the impact of carbon dioxide on world weather patterns is relevant. And so on. What is at issue between the naturalists and others is whether, as well as these "descriptive" matters, there are "pure" moral principles concerning what we should do, and how we are to determine what *those* principles are.

One might be suspicious that moral questions and “descriptive” questions can be divided up into entirely descriptive and purely moral claims. One way to separate out descriptive and moral questions is to use conditionals. For a given case, let the uncontroversially descriptive facts be represented by D, and some moral verdict be represented by M. Then the question of whether “if D, then M” is correct can be a matter of disagreement even for two people who both accept D (or both reject D, for that matter). We can disagree about the conditional "if John took the car, (and such-and such else was descriptively the case), then he was wrong to take it", even if we agree on whether or not he took the car, for example. (Or for that matter two people could agree with the conditional even if we disagreed about whether John took any cars.) The conditional seems to capture a moral commitment that goes beyond descriptive commitments (or at least the specific commitment of D, at any rate).

Specifying “purely” moral issues in this way avoids two tricky issues. One is whether “pure” moral claims can be captured by entirely general moral principles with no apparent non-moral commitments: one might doubt this can be done because of a general suspicion about moral principles of that generality. The other issue avoided is finding a way of once-and-for all drawing a line between moral and non-moral vocabulary. Many think that so called “thick” moral terms bring with them both moral and descriptive commitments: to describe a policy as cruel says something descriptive about the motives and potential for producing suffering of the policy, but also something moral about those who engage in it, for example. (See Sundell's chapter on thick concepts in this volume.) Some think these “thick” terms cannot be disentangled into those two components, at least in any language we have. If

we wish to isolate “pure” moral claims using the conditional strategy, we can often leave any allegedly thick moral expressions in the consequent, thus allowing us to state moral disagreements, including ones stated in thick terms, between people who agree on all of their descriptive commitments.

Note that in dividing up claims into descriptive ones and moral ones, I do not mean to imply that the truth of moral claims is not fixed by the truth of the descriptive ones. Moral facts might even *just be* descriptive facts, for all that has been said here. (One simple example of a theory like this: one could be a hedonistic utilitarian that *identifies* the property of being the morally best outcome with the property of being the outcome having the greatest balance of pleasure over pain.) Nevertheless, even those who wish to identify facts specified with moral language and facts specified by non-moral language face controversy about whether, in settling a question posed in non-moral vocabulary, science goes very far towards settling questions posed in moral vocabulary, since the science of pleasure and pain is unlikely to itself tell us that hedonism is the correct moral theory.

The next things to be careful about, when working out the role of science in ethics, is to keep in mind two distinctions. Let us first distinguish morality and *mores*: the difference between what is morally right or wrong, and what *the conventions of a given society treat as* morally right or wrong. Many people believe that what is right, on the one hand, and what a society currently treats as right, on the other, come apart. Some societies have had (and have) the death penalty for consensual sodomy: that by itself does not establish that consensual sodomy is sometimes in itself gravely wrong. Some societies have had (and do have) permissive attitudes towards slavery: that by itself does not show that slavery was morally permissible. Scientific discoveries about what a group’s mores are are only controversially of relevance to the question of what really is right or wrong. Even for those few who do who identify morality and social mores, the question of what justifies that identification, at least, is presumably not settled by discovering what the mores in fact are.

The second distinction worth keeping in mind is the distinction between morality and *altruism*. The exact definition of altruism is contested, but at a first pass altruistic behaviour is behaviour that benefits others at a cost to oneself. Sometimes in the popular media discoveries about the “science of morality” only concern the science of *altruism*: how it developed, why evolutionary pressure does not eliminate altruistic dispositions and why they are sometimes maintained by evolutionary mechanisms, and so on. We should not assume too quickly, however, that discoveries about altruistic behaviour will be discoveries about morality. For one thing, some altruistic behaviour seems immoral: nepotism and

corruption are often engaged in to help relatives or friends, and the fanatic who slaughters innocents for The Cause might well be behaving altruistically even if what he does is horrendously evil. Even when altruistic acts are morally good, understanding and explaining their altruistic nature may not give us much insight into why they are morally good.

So achievements in the anthropology or sociology of conventional standards, or the psychology or biology of altruism, do not seem to answer moral questions in any *obvious* way. It is tempting to think that we can gain some better answers to moral questions once we are sensitive to these discoveries, but getting those answers requires some further philosophical work. One valuable contribution naturalistic metaethics can make is to explain how and why we can draw conclusions about ethical and metaethical questions from evidence about the workings of mores and altruism. But we should not think that ethical and metaethical work is done *just* by investigating mores and altruism.

### ***3. Metaethical Naturalism: Two Examples***

There are many naturalistic projects, small and large, that are being carried out in metaethics. To get a sense of the range, let us look at two general approaches that stress methodological naturalism but which come to very different conclusions. The first is realist and cognitivist about ethical claims, the second anti-realist and expressivist.

One influential approach to metaethics in the late twentieth century was pursued by a group of philosophers known as the Cornell realists, due to connections they had to the philosophy department at Cornell University. Major figures in this tradition included Richard Boyd (1988), David Brink (1989) and Nicholas Sturgeon (1988), while other philosophers like Peter Railton (1986) have metaethical views that resemble the Cornell realists in important ways. What unified the Cornell realists was the view that we could learn about moral language and moral metaphysics by looking at what scientific realist philosophers of science had to say about scientific language and metaphysics. Part of this involved the thought that we can judge theories using so-called “inference to the best explanation”. Something similar can be used in metaethics to choose between theories: and just as we can use scientific method to tell us about all sorts of physical and social phenomena, we can use similar methods to tell us about moral phenomena.

The best theory that enables us to predict and explain the spread of diseases, for example, is one that employs talk of bacteria and viruses in its predictions and explanations. When we test our theories of disease and they pass those tests, and our theory yields satisfying explanations of initially puzzling disease phenomena, that provides us with good reason to *believe* our theories of bacteria and viruses, or to believe those theories are approximately correct. Or at least that is a picture of inquiry, known as "inference to the best explanation" (IBE), which many scientific realists would accept. Cornell realists hold that we can apply something like this approach to investigating morality. Some packages of ethical and metaethical views do better in explaining the truth of our secure moral judgements, the behaviour of agents, and various other things. And we can test these explanations through further investigation, comparison with rival explanations, and in other ways. When we come upon a package that passes our tests and offers satisfying explanations of previously puzzling phenomena, we can use roughly the same method as we use in the sciences to infer, using inference to the best explanation, that we are onto the truth in our moral and metaethical theorising. (Or we are at least close to the truth.) Furthermore, just as the facts about viruses and bacteria are objective, and do not depend on what we happen to believe or desire about viruses, so to are the facts about morality objective, not depending on the theorist's standpoint or preferences.

The inputs, or "moral observations", that go into our ethical and metaethical theorising presuppose some willingness to categorise the world we encounter in moral terms: presumably an alien observer could detect all the same movements, masses, and even colours and sounds without necessarily recognising any actions as just or unjust, right or wrong. But this is not a barrier to treating moral observations as being like scientific observations, according to Cornell realists, since scientific observations presuppose theoretical commitments as well. An observer with no concepts of distances or planets can look through a telescope and see nothing but spots and regions of coloured lights, but this should not make us think that using telescopes to do astronomy is somehow unscientific. Indeed, for any experimental or observational science, we must make assumptions about the world to gather information from observation: we can test those assumptions piecemeal, but we cannot do without them altogether. Since this is the situation when it comes to any scientific observation, it is no worry in principle that it also obtains in the case of moral observations.

Likewise with moral explanations. Someone who does not recognise that anything is right or wrong, good or bad, will not feel the need to explain anything they encounter in moral terms. But likewise, a hypothetical alien observer who can only detect living cells and chemical interactions might never feel

the need to offer psychological or economic explanations for phenomena that they observe, and this should not make us sceptical that there are minds or economies, or make us think that psychology or economics cannot be sciences. Cornell realists tend to go even further, and argue that there are explanations in moral terms of phenomena described in non-moral terms (Sturgeon 1988): the failure of a certain rescue mission, for example, can be explained on the basis that its leader is "no damned good" (p 243). If our best explanations, by scientific standards, of some phenomena described in non-moral terms turn out to be moral explanations, all the better for the Cornell realist story about confirming moral theories, and the metaethical package that goes along with those theories.

The second form of naturalism we will look at in some detail takes a radically different approach to the subject-matter of ethics, but also in a way that permits us to see metaethical investigation as fundamentally like a scientific enterprise in many respects. This approach is the quasi-realist approach of figures like Simon Blackburn and Allan Gibbard. While Gibbard is more explicit about tying his approach to metaethics to an understanding of ourselves enriched by the social sciences, (especially in Gibbard 1990 and Gibbard 2003), other quasi-realists such as Blackburn (Blackburn 1993) do not present their work as having a particularly scientific methodology.

In very rough outline, Blackburn and Gibbard see the central role for moral language as expressing certain attitudes other than belief: patterns of approval and disapproval for Blackburn, and plans for Gibbard (Gibbard 2003). However, despite this primarily expressive function of moral language, it behaves as if it states facts and makes claims to objective truth. After telling a story about how to express an attitude (or pattern of attitudes), a further story can be told about how to express a pattern of attitudes through calling various moral claims true or false, moral arguments valid or invalid, expressing claims about the connections between moral matters and non-moral matters (e.g. organised dogfights are wrong in part because of the unnecessary suffering of the dogs involved), and so on. Despite its core function as expressing attitudes other than belief, moral thought and language can come to have many of the features moral realists have claimed for it. (Arguably, Gibbard 2003 ends up slipping back into a kind of moral realism, though he would not describe the view that way.)

What should expressivists say about moral method: how should we work out what to think about moral questions, and argue with others about them? Blackburn and Gibbard say a limited amount about this topic, but plausibly, this will be a matter of engaging with our own non-cognitive states, together with an evaluation of whether they would be different were we more ideal. Are any the product of false or

dubious factual beliefs? Are any produced by distorting processes that would be absent in more idealised versions of ourselves? Are their conflicts between our patterns of approvals or patterns of planning, and if so what is our attitude to how they should best be resolved? And so on. This internally focused project of self-improvement, together with related interpersonal exchanges aimed at this sort of thing and perhaps at exercising more-or-less respectful pressure on each other, seem like a radically different sort of 'inquiry' from scientific investigation of fully factual questions. (This seems to fit Gibbard's discussion of moral method in Gibbard 1990 p 253-337, and his explicit rejection of moral inquiry using just the methods of social science in Gibbard 2003 p 161.) So when it comes to making moral judgements, the method that seems to mesh best with quasi-realism is more like that proposed by traditional non-cognitivists such as Stevenson rather than those who want to assimilate moral inquiry to scientific inquiry.

Despite this, when it comes to methods for resolving *metaethical* questions, quasi-realists are methodological naturalists. While *moral* questions are distinctively different from scientific ones, questions about *what it is to make a moral judgement*, or *the function of moral language*, or in general the psychologies of moral agents and the language used by moral communities are naturalistic questions. Indeed, one motivation for expressivism is a naturalistic conception of ourselves and our responses, and a desire to avoid postulating non-naturalistic moral facts, or non-natural facts about the nature of rationality. That motivation in the first instance comes from metaphysical naturalism rather than methodological naturalism, but it goes hand in hand with thinking that scientific means are the way to improve our theories of human agents and their interactions. Quasi-realists should think, and do think, that psychology and linguistics will reveal what mental states we have and how we express them in language. Philosophers of language and mind and those who work on metaethics and moral psychology have contributions to make here too, but as part of an integrated project of understanding ourselves as natural creatures in natural environments, rather than seeking a special philosophical realm of objective facts about morality or rationality. Ethics does not look much like science from a quasi-realist perspective, but metaethics does. In treating meta-ethics naturalistically despite not taking a naturalistic approach to ethical questions, quasi-realists resemble some earlier non-cognitivists such as Stevenson 1963, though I would hesitate to classify some other traditional non-cognitivists, such as Wittgenstein, as meta-ethical naturalists.

#### ***4. Naturalism and Traditional Methods of Metaethical Inquiry***

Naturalism in ethics can seem like a radical doctrine, especially against the background of insistence by people like Moore and Wittgenstein that ethics cannot be science. So you might have thought, and some naturalists do think, that methodological naturalism precludes relying on the methods traditionally used by analytic philosophers to answer metaethical and ethical questions. However, some self-identified naturalists have tried to vindicate the use of moral intuitions, reflective equilibrium, and even conceptual analysis as important parts of ethical inquiry. This section will look at some attempts to reconcile these methods with naturalism.

Debates in moral philosophy frequently feature appeals to “intuition”, and some theories of the source of moral intuitions make using intuitions radically different from engaging in scientific method. G.E. Moore’s theory of moral intuition, for example, maintained that moral intuition was a distinct psychological faculty, that gave us access to a special kind of property, even when contemplating non-actual scenarios. Some theories of this sort of intuition take it to be a special rational faculty that delivers us information about the world: see Bealer 1998 for a defence of this view of philosophical intuitions in general. Metaethical positions are also often justified by appeal to “intuition”: it is supposed to be our intuitions about the subject matter that support the doctrine that moral facts are objective, for example. Understandably, naturalists have often wanted to reject the psychological picture on which we have this sort of special faculty, or other distinctive moral detectors such as a “moral sense”.

However, there are other theories of intuitions and appeals to intuitions that make them seem less anti-naturalistic. Perhaps intuitions are just our judgements about cases: David Lewis claims about philosophical intuitions that “[o]ur intuitions are simply opinions” (Lewis 1983 p x). Or perhaps they are a special class of judgements, such as the relatively unreflective ones or the ones not arrived at by explicit inference from other judgements. Or perhaps they are *inclinations* to make such judgements. Seen in these ways, starting from our moral intuitions may be no more mysterious than starting from our ordinary opinions in any other inquiry. Historically, this seems to have been where many successful scientific inquiries came from: even physics and chemistry have their origins in plausible speculations about motion and the interactions of matter. Even today, many mature scientific inquiries still employ elements of good sense, hunches, and other moves that look like they are supported by something like educated “intuition”. It would make no sense for a methodological naturalist to demand that ethical and metaethical theorising have no place for intuition if intuitions can be found playing a helpful role even in paradigm scientific enterprises.

The method of reflective equilibrium, very roughly, is the method of beginning by articulating particular moral judgements and pre-theoretic moral generalisations, perhaps supplementing both of these with moral intuitions, and then bringing the resulting theory into “equilibrium” by working out the strengths of the various judgements and in cases of clashes, modifying those held less strongly so as to cohere with the more certain ones. As mentioned above, this procedure can seem suspicious to naturalists: it does not seem to be constrained by what the world is like, as opposed to what we are inclined to believe about the world. It suggests that ethics is constructed, rather than discovered, which seems a stark contrast from scientific inquiry. Indeed, Rawls himself seems to have favoured a constructivist approach to ethical truth. Finally, the starting point of reflective equilibrium involves giving a central role to intuitions, so suspicion of intuitions naturally carry over to suspicion of reflective equilibrium.

However, it is far from obvious that there is anything naturalistically objectionable about using reflective equilibrium. Historically, Rawls’s conception of reflective equilibrium derives from Goodman’s 1955 account of how we should establish the correct theory of inductive inference. Goodman’s project was, in part, giving a theory of good scientific method: and so trying to do the same thing for ethics looks like a methodologically *naturalist* move, not an anti-naturalist one. Those sympathetic to Goodman, at least, are likely to see scientific inquiry itself as engaging in something like reflective equilibrium, trading off views about the soundness of particular judgements about the world and principles of unifying and revising those particular judgements. So it is not clear that employing reflective equilibrium is somehow unscientific.

When evaluating how naturalistic reflective equilibrium is, it can matter whether we conceive of reflective equilibrium in a *narrow* or a *wide* way. (For the distinction see Daniels 1979.) Narrow reflective equilibrium looks only at one’s own particular and general moral judgements, and trades them off to reach a stable equilibrium. *Wide* reflective equilibrium takes into account many more considerations when trying to reach a stable resting place, including what judgements others are likely to make, different conjectures that could explain the judgements of oneself and others, and beliefs other than moral beliefs, e.g. beliefs about how it is reasonable to criticise initial judgements, beliefs about human psychology and sociology, and so on. At its widest, reflective equilibrium could be performed on all our beliefs at once. A wide enough reflective equilibrium can take into account anything a naturalist might want to bring to bear on moral inquiry, so whether a particular project of reflective

equilibrium is objectionable on naturalist grounds is likely to have more to do with how narrow it is than that it involves reflective equilibrium *per se*.

Using conceptual analysis to reach metaethical conclusions is widespread, and seems to often be what is going on when philosophers offer an "analysis" of justice, or supererogation, or moral goodness in terms of reasons, etc. Conceptual analysis is often thought to be anti-naturalist, both because the results of conceptual analysis are meant to be *a priori* (and many methodological naturalists are suspicious of the *a priori*), and because resolving a question through conceptual analysis seems to be a method that insulates that inquiry from what science has to tell us, and from scientific testing of that answer. These suspicions of conceptual analysis are not particularly to do with conceptual analysis as used in ethics: many naturalists are suspicious of conceptual analysis across the board. See e.g. Devitt 1996 or Kornblith 2002.

However, a number of theorists have wanted to defend conceptual analysis as naturalistically respectable. Just two examples are the accounts of conceptual analysis offered by Rey 1998 and Jenkins 2008, 2013. While these accounts have many psychological and epistemological differences, a shared idea is that what concepts we have and what those concepts represent about the world is a matter of past engagement with the world itself: and once we think of concepts as mental representations bearing reliable information about the external world, a connection produced by ordinary causal interactions, it does not sound particularly anti-naturalistic to think that conceptual analysis is in part a process of making that information available for belief. If conceptual analysis is widely used, at least implicitly, in the sciences, that would make it even more naturalistically acceptable. For example, if Jenkins 2008 is right that conceptual analysis, as she conceives of it, is the source of our basic arithmetical knowledge, the results of conceptual analysis sit at the core of most contemporary natural and social sciences: neither would get very far without numbers.

One strand of recent metaethical theorising that would describe itself as naturalist, despite relying heavily on conceptual analysis, is the *moral functionalism* defended by Frank Jackson and Philip Pettit (see Jackson and Pettit 1995, Jackson 1998). The core of metaethical inquiry, according to this view, consists of two steps. The first is to assemble "commonplaces" about ethics, particularly generalisations about ethical matters that we could reasonably think are shared among those who use ethical concepts and are at least implicit, if not explicit, in their use of moral expressions. Once we have our commonplaces assembled, we can treat them as implicitly defining the subject matter of our

inquiry. This network can be treated as implicitly defining many pieces of moral vocabulary together. For example, if we define just actions partly in terms of which distributions of resources are morally good, for example, then information we have about what are paradigm examples of just actions are might help constrain what counts, in our "folk" theory of morality, as being a morally good outcome.

Armed with our networked implicit definition of moral vocabulary, the second step is to examine the world to determine what properties and objects best satisfy that network definition of our moral terms. Jackson and Pettit 1995, for example, suggest that the right actions are the ones we would approve of under various (descriptively specified) improvements to our situation. Jackson 1998, on the other hand, suggests that the property of goodness, as implicitly defined, might end up being maximising desire satisfaction. The question of what sequence of properties fits the networked implicit definition supplied by moral commonplaces, however, is in principle a matter for scientific inquiry, particularly social and psychological scientific inquiry.

The theory that results from the moral functionalist procedure is intended to be *metaphysically* naturalistic, or “descriptivist”, since the aim of the moral functionalist is to determine what it is in the world, specified in non-moral vocabulary, which plays the functional roles specified for moral properties like goodness and obligatory action and justice. But it is natural to see it is as *methodologically* naturalistic as well—or at least intended to be by its proponents. The first step of the project is one they think is needed in a wide variety of areas to fix the subject matter of an inquiry: indeed, they model their strategy on that of the *analytic functionalists* in the philosophy of mind, who think that this kind of approach will provide the foundations for cognitive psychology, by offering us a way of understanding what e.g. beliefs and desires are. And the second step, locating the things that best play those roles in the world that science tells us about, looks like a fairly standard scientific inquiry. Despite the heavy reliance on conceptual analysis in the first step of establishing the networked interdefinition of moral vocabulary, the moral functionalist approach is intended to be part of a methodologically naturalistic approach to metaethics.

This section has attempted to establish, not that naturalists must by sympathetic to methods such as reliance on intuitions, use of reflective equilibrium, and use of conceptual analysis, but rather that being a naturalist about moral methodology does not *per se* mean that a theorist must reject any of these three tools for moral inquiry. Methodological naturalism *need* not be a revolutionary doctrine when it comes

to the question of how to do moral inquiry.

However, those who retain any of these methods but who wish to be methodological naturalists may want to treat them in particular ways. It would be natural to not *just* rely on intuitions in metaethics, especially if intuitions are something like trained hunches or what seems plausible before detailed inquiry. Methodological naturalists are likely to think that conceptual analysis cannot be the full story about the metaphysics and epistemology of morality or the key to moral psychology: all of these areas seem to have substantive questions in them to be resolved by theoretical inquiry and not just reflection on concepts. Finally, insofar as reflective equilibrium plays a role, it is more plausible that information about ourselves and the world other than just the particular judgements and principles we begin with should be the inputs into our theoretical trade-offs.

Methodological naturalism in metaethics offers more than a new take on traditional methods, however. Bringing scientific methods to bear on metaethical questions gives us more options than just pursuing metaethical business-as-usual. Let us turn to examine some of these new options.

### ***5. New Uses for Science in Metaethics***

Perhaps naturalists can keep some or all of the traditional methods of ethics, and approaches to metaethics, employed in analytic philosophy. But naturalism can make a distinctive contribution if there are new, promising, methods available through applying scientific techniques to address ethical and metaethical questions.

We have already seen one suggested above in the discussion of Cornell realism. It is that we can use inference to the best explanation (IBE) to reach ethical conclusions. Cornell realists talk less about the justification for their approach to metaethics, but it is natural to see this as a matter of applying IBE as well. Note that the general kind of project pursued by the Cornell realists could be pursued by naturalists who disagree with particular views the Cornell realists may have had about science, language, or ethics: just as rival scientific theories can appeal to the same canons of scientific method, rival metaethical approaches can share a view about what it would take to vindicate one approach over the other.

A second use of the methods of the sciences, particularly the social sciences, can be found in the

contemporary "experimental philosophy" or "xphi" movement. Narrowly construed, this movement seeks to discover what non-philosopher's judgements are about topics of philosophical interest, for example to test whether "intuitions" philosophers rely upon are widespread in the general public, and to do so through empirical surveys. Why would the results be valuable for metaethics? If intuitions or the results of conceptual analysis were legitimate inputs to metaethical theorising, then drawing intuitions from a broader range of informants and observing how they deploy concepts that they are presumptively competent with could provide better data for those inputs. On the other hand, if there is no particular reason to trust intuitive judgements or information apparently drawn from concepts, it is less clear why the judgements of the person on the street are relevant. Experimental philosophy might still be able to serve a debunking function if it turned out that philosophers' intuitions were not widely shared beyond the ivory tower, for example. But its role in positive theorising about the nature of ethical facts, the epistemology of moral judgement, their objectivity etc. would likely be limited.

Despite its name, "experimental philosophy", in the most common use of that term, covers a comparatively narrow range of experimental work being drawn on by contemporary philosophers. Another major source of evidence is coming from psychological investigations into how we make moral judgements. Some of that work takes the form of asking informants for their moral judgements and theorising about what best explains those patterns (i.e. xphi in the more narrow sense). But psychologists can take advantage of more sources of evidence than that. To give one example, one debate with metaethical relevance is the debate about how much moral behaviour is influenced by relatively transitory features of an environment that an agent finds herself in. This debate, about *situationism*, seems to have implications for whether there are moral virtues, or at least whether moral virtues can play the kind of central role that virtue ethicists have traditionally thought: if it is an illusion that there are kind people and cruel people, as opposed to people acting kindly or cruelly because of local situational factors, then a virtue ethics that builds an account of moral behaviour around virtues such as kindness and vices such as cruelty would be refuted.

The experimental evidence assembled about situationism is typically drawn from putting experimental subjects in a variety of conditions and seeing what they do. A famous early experiment in this area is the notorious experiment carried out by Stanley Milgram, where experimental subjects would administer what they thought were severe, and potentially fatal, electric shocks to other people as punishment for mistakes in a test, merely because the experimenter told them sternly to do so. These punishments continued despite the screams and pleas of those apparently affected by those shocks

(Milgram 1974). More recent experiments have demonstrated significant differences in moral behaviour based on relatively small differences in circumstances: for example, the likelihood someone will stop to help an apparently injured man who has dropped some books is radically decreased if there is loud ambient noise, such as a mower running nearby (Mathews and Canon 1975). The cumulative effect of these sorts of experiments leads Doris 2002 to argue that the psychological presuppositions of traditional virtue ethics must be rejected.

Of course, experimental evidence on its own does not settle the issue of whether traditional virtue ethics has been refuted, and Doris is sensitive to the need do serious theoretical work to show that these experiments bear on the viability of virtue ethics (Doris 1998). There are still debates to be had about the connection between evidence about people's behaviour in morally charged situations and what conclusions we can draw about whether there is moral character, and if so what it is like. Doris and Stich 2014 have a good introductory discussion of the back-and-forth about situationism and virtue ethics, as well as a discussion of many other places where psychological results (including more narrowly x-phi ones) play into metaethical debates. Whatever conclusions should be drawn in the end about situationism and virtue theory, it is clear that psychological experiments other than surveys have a role to play in metaethical debates, as well as xphi narrowly conceived.

There is no reason why psychological research is the only place where metaethicists should find useful scientific resources, either. Investigation of the evolution of human behaviour, and how it is influenced by genetics and transmitted culture are others. Both economics and political science have, in their own ways, always been close cousins of ethical inquiry, and the project of working out what sort of thing moral agents are and how they best interact is naturally one that disciplines such as anthropology and sociology can also contribute to. There is often careful theoretical work to be done to show how scientific results bear on traditional metaethical questions. But in this respect metaethics need be no different from science either: the links between results of experiment and general theoretical questions often require careful work to make out, despite the mythology of entire theories being immediately swept away by single "crucial experiments".

## ***6. Conclusion***

This introduction could not hope to be comprehensive: many ethicists see themselves as naturalists to a greater or lesser extent, and the temptation to conduct metaethical and ethical inquiry along scientific

lines is perennial. Given the disputes between card-carrying naturalists, it is unlikely that a commitment to methodological naturalism all by itself can determine very much about how to carry out moral inquiry. My own view is that it is very unlikely that naturalism could provide a shortcut that is some kind of alternative to the careful inquiry and deliberation that making progress on moral issues requires. However, construing moral method in naturalistic terms does shift the focus of some methodological debates, and it can direct attention to avenues of inquiry too often neglected by traditional moral philosophy. Whatever turns out to be the truth of the matter about the methodological similarities between moral inquiry and scientific inquiry, or between metaethical inquiry and scientific inquiry, the debate between naturalists and non-naturalists, and among different naturalists, promises to both improve our understanding of moral inquiry and improve our methods for that inquiry.<sup>1</sup>

### **References**

- Bealer, G. (1998) "Intuition and the Autonomy of Philosophy" in DePaul, M. and Ramsey, W. (eds.) *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, Lanham MD: Rowman and Littlefield. pp 201-240
- Blackburn, S. (1993) *Essays in Quasi Realism*, Oxford: Oxford University Press.
- Brink, D. (1989) *Moral Realism and the Foundations of Ethics*, New York: Cambridge University Press.
- Boyd, R. (1988) "How to be a Moral Realist" in Sayre-McCord, G. (ed) *Essays on Moral Realism*, Ithaca: Cornell University Press. pp 181-228
- Daniels, N. (1979) "Wide Reflective Equilibrium and Theory Acceptance in Ethics". *Journal of Philosophy* 76.5: 256-282
- Devitt, M. (1996) *Coming to Our Senses*, Cambridge: Cambridge University Press.
- Doris, J. (1998) "Persons, Situations and Virtue Ethics". *Noûs* 32.4: 504-530
- Doris, J. (2004) *Lack of Character: Personality and Moral Behavior*, Cambridge: Cambridge University Press.
- Doris, J. and Stich, S. (2014) "Moral Psychology: Empirical Approaches". *Stanford Encyclopedia of Philosophy*, accessed 22 October 2015. URL = <http://plato.stanford.edu/archives/fall2014/entries/moral-psych-emp/>
- Gibbard, A. (1990) *Wise Choices, Apt Feelings*, Cambridge MA: Harvard University Press.

---

<sup>1</sup> Thanks to Frank Jackson, Carrie Jenkins, and Alex King for discussion and to Tristram McPherson, David Plunkett and Lea Schroeder for comments. Work on this paper was supported by the National Humanities Center.

- Gibbard, A. (2003) *Thinking How to Live*, Cambridge MA: Harvard University Press.
- Goodman, N. (1955) *Fact, Fiction and Forecast*, Cambridge MA: Harvard University Press.
- Jackson, F. C. (1998) *From Metaphysics to Ethics: A Defence of Conceptual Analysis*, Oxford: Oxford University Press.
- Jackson, F.C. and Pettit, P. (1995) "Moral Functionalism and Moral Motivation". *Philosophical Quarterly* 45.178: 20-40
- Jenkins, C.S. (2008) *Grounding Concepts: An Empirical Basis of Arithmetic*, Oxford: Oxford University Press.
- Jenkins, C.S.I. (2013). "Naturalistic Challenges to the A Priori" in Casullo, A. and Thurow, J.C. (eds) *The A Priori in Philosophy*. Oxford: Oxford University Press. pp 274-290
- Kornblith, H. (2002) *Knowledge and its Place in Nature*, Oxford: Oxford University Press.
- Lewis, D. (1983) *Philosophical Papers Volume 1*, Oxford: Oxford University Press.
- Mathews, K. E. and Canon, L. K. (1975) "Environmental Noise Level as a Determinant of Helping Behavior," *Journal of Personality and Social Psychology* 32.4: 571–577
- Moore, G. E. (1993) *Principia Ethica*, Cambridge: Cambridge University Press.
- Milgram, S. (1974) *Obedience to Authority: An Experimental View*, New York: Harper and Row.
- Quine, W. V. (1951) "Two Dogmas of Empiricism," *Philosophical Review* 60.1: 20-43
- Railton, P. (1986) "Moral Realism," *The Philosophical Review* 95.2: 163-207
- Rawls, J. (1971) *A Theory of Justice*, Cambridge MA: Belknap Press.
- Rey, G. (1998) "A Naturalistic A Priori," *Philosophical Studies* 92: 25-43
- Smith, M. (1994) *The Moral Problem*, Oxford: Wiley-Blackwell
- Stevenson, C. (1963) *Facts and Values: Studies in Ethical Analysis*, New Haven: Yale University Press.
- Sturgeon, N. (1988) "Moral Explanations" in in Sayre-McCord, G. (ed) *Essays on Moral Realism*, Ithaca: Cornell University Press. pp 229-255
- Wittgenstein, L. (1965) "A Lecture on Ethics," *Philosophical Review* 74.1: 3-12