

**Alma Mater Studiorum-Università di  
Bologna  
Università degli Studi di Parma**

Dottorato di ricerca in Filosofia Analitica (XIX ciclo)

*Moral Explanations and Ethical  
Naturalism*

Coordinatore  
Chia.mo Prof. Paolo Leonardi

Relatore  
Dott. Mario Ricciardi

Dottorando  
Dott. Andrea Viggiano

Esame finale anno 2007  
Settori Scientifico-Disciplinari MIUR: M-FIL/03, M-FIL/05

# Contents

<i>Introduction</i>	1
<i>Chapter one</i>	
<b>THE CHALLENGE</b>	6
1. The condition for observational testing	9
2. Why moral principles do not satisfy the condition for observational testing	20
<i>Chapter two</i>	
<b>NATURALISTS' REPLIES</b>	30
1. Sturgeon's reply	31
2. (Some of) Sturgeon's examples	37
3. An empirical challenge	43
<i>Chapter three</i>	
<b>A MODEL OF MORAL EXPLANATIONS</b>	49
1. Substantive constraints on a defense of moral explanations	50
2. Moral explanations	59
3. Some objections	72
<i>Bibliography</i>	82

## Introduction

Ethics is different. This is a quick way of stating a point many philosophers would agree on. Different from what? Ethics is different from the empirical sciences. Our commonsensical picture of the empirical sciences is that they deal with a world whose existence and nature are largely independent of our theorizing about it: they aim at finding out about various kinds of facts which obtain in the natural world, and (at least sometimes) succeed in giving us knowledge about natural facts. Suppose we grant that this picture of the empirical sciences is correct: what about ethics? According to many philosophers, we cannot tell the same story about it: there is something distinctive and peculiar about ethics, such that it cannot be put in the same category as physics, biology, psychology, or economics. According to some, ethics is different from the empirical sciences in that it does not aim at finding out about facts in the world or, even if it does, there are no such facts to be found out about: there are no moral facts. According to others, there are moral facts, and the difference is that they are not facts of the same general kind as those studied by the empirical sciences: moral facts are not natural facts, and knowledge about them cannot be accounted for by the same methods and procedures which account for our scientific knowledge of the natural world.

Ethical naturalism is the claim that there are moral facts, and they are natural facts, which can be known in pretty much the same way in which scientific facts are known. According to ethical naturalism, our commonsensical picture of the empirical sciences can be applied to ethics too: ethics is not different. What should we think of such a position? Should we accept ethical naturalism, or should we reject it and instead accept the claim that ethics is different? In what follows, I'll deal with an argument that has been raised as an objection

against ethical naturalism, and as a reason to accept the view that ethics is peculiar, different from the empirical sciences.

Objections to ethical naturalism are often constructed on the following pattern. The major premise is taken from some philosophical area different from moral philosophy (metaphysics, epistemology, philosophy of language, philosophy of science, and so on), and amounts to a general claim about natural facts, to the effect that they have a certain property P, or that knowledge of them has a certain property P, or that language dealing with them has a certain property P, and so on. A minor premise is then added about moral facts, to the effect that moral facts (moral knowledge, moral language, and so on) are not P. This second premise is the part of the argument dealing specifically with ethics: it can be presented as an obvious truth about ethics, not in need of any articulated defense, or it can be explicitly defended by means of a sub-argument within the general argument against ethical naturalism. The conclusion is then drawn that moral facts (moral knowledge, and so on) differ(s) from natural facts (knowledge of natural facts, and so on): ethics is different. The purpose of this type of arguments against ethical naturalism is thus to point to a relevant respect (what I called the property P) under which moral facts differ from natural facts. Various implementations of this argumentative strategy have been proposed in the meta-ethical literature, yielding different arguments against ethical naturalism. Here are a couple of examples:

- 1) Disagreements about natural facts often get settled, and people's beliefs about natural facts tend (at least in the long run) to converge.
- 2) Moral disagreements do not get settled very often, and people's moral beliefs show no tendency (even after centuries of moral debates) toward convergence.
- 3) Therefore, moral disagreement is not disagreement about any matter of natural fact.

- 1) Beliefs about natural facts are by themselves inert: they do not provide those holding them with any motive or reason for action.
- 2) Moral beliefs are not inert: they do by themselves provide those holding them with a motive or reason for action.
- 3) Therefore, moral beliefs are not beliefs about any matter of natural fact.

According to the first argument, what goes wrong with ethical naturalism is its account of moral disagreement. If ethical naturalism were true and moral facts were natural facts, then disagreement about them should have the same characteristics as disagreement about any matter of natural fact: specifically, it should prove tractable and we should be able to get rid of it over time. But we are not: moral disagreement is different. According to the second argument, it is instead what we might call the distinctive practical role of our moral beliefs that proves fatal to ethical naturalism and gives us reason to believe that ethics is different.

The argument I'll be dealing with focuses on moral epistemology. According to ethical naturalism, moral facts are natural facts, and they can be known in pretty much the same way in which scientific facts are known: one way of objecting to it is therefore by claiming that moral epistemology looks instead quite different from the epistemology of the empirical sciences. What this argumentative strategy boils down to will depend, of course, on what one takes the epistemology of the empirical sciences to look like. In the first half of the last century, for instance, a widely held view was that scientific principles are tested by testing their observational implications. This view provided the major premise for the following epistemological argument against ethical naturalism:

- 1) Scientific principles have observational implications, and are tested by testing such implications.
- 2) Moral principles have no observational implications.
- 3) Therefore, moral principles are not tested in the same way scientific principles are.

Nowadays the major premise of this argument is largely discredited, at least if meant as the argument did: i.e. as claiming that scientific principles are tested by testing the observational implications they have by themselves, if taken in isolation. However, other views about the epistemology of the empirical sciences have been proposed as the bases for different, better epistemological arguments against ethical naturalism: one such argument is the one I'll be dealing with.

In dealing with this epistemological argument against ethical naturalism, I'll make use of the distinction between theoretical and observational facts, so it may be useful to explain here briefly what I mean by this distinction. Observational facts are non-theoretical facts we take to be relevant in order to reach some conclusion about theoretical facts. To take an example that will often recur in the following pages, suppose I am a physicist developing a theory about protons: the fact that there is a vapor trail in my cloud chamber is a fact I might take to be relevant to the question of whether or not a free proton is going through the cloud chamber. Or suppose I am a moralist, and am thinking about whether a given action was right or wrong: I'll regard various non-moral facts (what type of action it was, whether it was done on purpose, and so on) as relevant to answer my moral question. In both cases, I have a question about a given type of facts (about whether a free proton is going through the cloud chamber, or about whether the action was a morally wrong action) and take facts of some different type to be relevant in order to reach an answer to that question: these latter facts count as observational facts. Taken in this way, the distinction between theoretical and

observational facts does not, I believe, commit one to any controversial claims in epistemology. In particular, it does not commit one to the claim that observational facts are rock-bottom facts, through which we can have a direct access to reality, without the mediation of any part of our conceptual scheme, or to the claim that beliefs about observational facts are epistemologically privileged, incorrigible, or anything of the sort, as is shown by the fact that the distinction between theoretical and observational facts is a distinction relative to the kind of theoretical inquiry under consideration, so that the same fact can count as theoretical or as observational in different contexts: whether someone acted in a certain way out of this or that motive counts as a theoretical question for the psychologist, and as an observational one for the moralist. Our beliefs about observational facts can go wrong just as much as our theoretical beliefs: this fact does not mean, however, that we cannot distinguish between theoretical and observational facts within our theories, nor that observational facts do not play a legitimate and important evidential role in the construction of our theories.

## The challenge

The dispute about moral explanations began with an argument put forward by G. Harman in the first chapter of Harman [1977]. Harman claimed that there is «a basic philosophical problem about morality», namely «its apparent immunity from observational testing» (p. vii), and traced this problem to the lack of explanatory power of moral facts: «moral principles cannot clearly be tested by observation, since they do not appear to help explain observations» (p. 8). Harman's argument has since been discussed by several philosophers. While agreeing that the argument raises an interesting point, though, different philosophers give different reconstructions of what Harman meant, or at least should have meant in order to express his point in its full force<sup>1</sup>. I'll begin by presenting my own reconstruction. I believe Harman was "on" to something, but his formulations are sometimes misleading, so I'll begin by presenting what in my opinion is the best way of taking his claims and examples. My reconstruction will drop some of the things Harman actually said, while adopting, revising, or emending others. I'll write about "Harman's argument" or "Harman's premises" because the materials I'll be drawing on come from Harman's actual claims and examples, but I do not claim mine is the only, or the best, way of rendering Harman's actual text. What I do claim is that my emended Harman's argument is a better way of expressing the interesting point the real Harman was "on" to in those passages. That argument poses a powerful challenge to ethical naturalism, and that is the argument I am interested in reconstructing and discussing.

---

<sup>1</sup> See for instance Sturgeon [1988], Wright [1992, chap. 5], and Thomson [1996, chap. 6].

Harman's argument deserves careful consideration for two reasons. In the first place, it is a common theme in the writings of many ethical naturalists that some popular and widely accepted objections against ethical naturalism presuppose what is instead bad philosophy of language, or bad philosophy of science, or bad epistemology. For instance, ethical naturalists commonly argue that our inability to provide non-moral synonyms for moral terms should be considered no objection to ethical naturalism, because it is only a bad philosophy of language that can lead us to think the naturalist is committed to the existence of such non-moral synonyms. Or they argue that the fact that moral principles are devoid of observational implications, if taken in isolation, is no objection against ethical naturalism, because scientific principles too are devoid of observational implications, if taken in isolation: it is just bad philosophy of science to require moral (or scientific) principles to have observational implications by themselves, if they are to be empirically tested<sup>2</sup>. I find these naturalistic rebuttals persuasive, and it is therefore an interesting feature of my emended Harman's argument that it cannot be rebutted along similar lines. Harman focuses on a certain feature of scientific principles - their observational testability, which he thinks depends on the explanatory power of scientific facts - , in order to argue that moral principles do not share it, and that this constitutes «the basic problem about morality» (p. viii). Even though I do not find Harman's argument persuasive in the end, I think he is right in claiming that the feature he focuses on is a real feature of scientific principles: his argument does not rely on bad philosophy of science. Harman's argument, therefore, poses a legitimate challenge to the ethical naturalist: it sets up a constraint that moral theory does have to satisfy, if ethical naturalism is to look plausible. The naturalist can reply that, *contra* Harman, moral facts are

---

<sup>2</sup> The general idea is a major theme of Boyd [1988]. For the first kind of rebuttal see Boyd [1988], Brink [1989, chap. 6], and Sturgeon [2003; 2006a]. For the second kind of rebuttal see Sturgeon [1988] and Brink [1989, pp. 136-138 and 183-184].

not devoid of explanatory power, and therefore moral theory does satisfy that constraint. What she cannot do, though, is to discharge the burden by denying the legitimacy of the challenge<sup>3</sup>.

My emended Harman, moreover, raises a challenge that is not only legitimate, but also crucial. As I hope will be clear from my discussion of the argument, the fact that the theoretical principles of an empirical science undergo observational testing, and that its theoretical facts have explanatory power, is a crucial feature of the process of empirical theory-construction, as this process is exemplified in our scientific theories. These features account for our confidence that such a process is responsive to the observational evidence available (that it is a process of *empirical* theory-construction), that we hold the scientific theories we do because they are supported by observational evidence, and in particular by their explanatory power with regard to such evidence. It is hard to deny that these features should be regarded as crucial epistemological virtues for any theoretical enterprise claimed to be analogous to that of the empirical sciences. So the ethical naturalist, if she wants to maintain that moral facts are known in pretty much the same way in which scientific facts are known, has to consider the defense of the observational testability of moral principles, and the explanatory power of moral facts, as crucial to the feasibility of her philosophical project. However one is to judge Harman's argument in the end, what one can learn from it is how the epistemology of ethics has to be like, if ethical naturalism is to turn out true. Harman claims he has located the basic problem about morality: I do not think what he has located constitutes any insuperable problem for ethical naturalism, but I do think his argument locates a crucial feature moral epistemology has to have, if the naturalistic project is to be feasible.

---

<sup>3</sup> There is another possible option for the ethical naturalist: accept that scientific facts are explanatory, accept that moral facts are not, and deny that this difference is relevant to the tenability of ethical naturalism. Ethical naturalism claims that moral facts are natural facts, and is committed to the claim that morality resembles science in those philosophically relevant aspects of it which account for its ability to give us knowledge of natural facts, not to the claim that morality resembles science in *every* respect. The feature of science on which Harman's argument focuses, though, is, as we shall see, an entirely typical feature of scientific theories as they are currently developed, so defenders of ethical naturalism should regard this option as a last resort, and should be ready to offer a quite powerful motivation for it. For an argument that, if this option were pursued, the epistemology of ethics might, under certain assumptions, look not too different from that of the empirical sciences after all, see Sturgeon [1998].

## 1. The condition for observational testing

Harman's claim is that moral principles cannot be tested in the same way we test scientific principles. We have a characteristic way of testing scientific principles, according to Harman, namely observational testing. Moral principles differ from scientific ones because they cannot undergo observational testing in the same way scientific principles can.

How does observational testing work for scientific principles? and why is it important that scientific principles can be tested in this way? Suppose we have a scientific principle P we want to test: for instance, let P be the claim that under certain specified lab conditions C, atoms of gas X emit protons. Since P says that a certain theoretical fact occurs whenever certain conditions are realized, the obvious way to expose it to observational testing is to observe whether the occurrence of the theoretical fact really accompanies the realization of those conditions: set up the conditions C described by P, and observe whether there occurs the theoretical fact that P says will occur.

There is a problem, though, that threatens to jeopardize our attempt to test P in this way. The problem issues from the fact that we cannot observe (not in any direct way) a proton. Scientists do talk about "observing" protons, but what really happens when they "observe" a proton is that they observe something else that they (given their theoretical training) see as a proton. More generally, we cannot observe in any direct way whether a given theoretical fact occurs: we always "observe" theoretical facts by observing something else that we (given our theoretical assumptions) see as a theoretical fact. The point is just a version of a (by now) common claim among philosophers. Theoretical facts are, by definition, not observational. We do ascertain their occurrence by making use of observations, but we do that by observing the occurrence of some observational fact that we take as evidence for the theoretical fact in question. And, as philosophers sometimes put it, assessments of evidence are theory-

dependent. In order to take the occurrence of an observational fact O as evidence for the occurrence of a theoretical fact T, we have to rely on background assumptions, assumptions that are usually quite complex and often implicit (so much so that we are typically unable to specify them in a fully explicit form) and that invariably include other theoretical beliefs. There is no way to go from O to T by pure meaning analysis, or any other theoretically neutral way. The only way to assess a given piece of evidence (to decide what theoretical conclusion, if any, it is evidence for) is by relying on some body of further theory<sup>4</sup>.

How does this threaten to jeopardize our attempt to test P? We want to expose P to observational testing, so we set up the conditions C described by P, in order to “observe” whether there occurs the theoretical fact that P says will occur. No such “observations”, though, occur in the void: they are always theory-dependent. We need to rely on some body of background theory in order to make them. A problem then might arise, depending on the type of background beliefs we rely on in making our “observation”: the background beliefs we rely on might not be of the right type for our “observation” to provide a test for P. Specifically, our “observation” will not provide such a test, if the background beliefs we rely on are such as to allow us to take the very fact that gas X is under conditions C as the observational evidence relevant to determine whether any protons are emitted<sup>5</sup>.

To illustrate the point, let’s suppose that our background beliefs include principle P\*, which says that under conditions C\* atoms of gas X emit protons, where conditions C\* are such that they subsume conditions C. Given our background beliefs, once we see gas X under conditions C, and realize that conditions C\* subsume conditions C, we can take that as

---

<sup>4</sup> In what follows I’ll use inverted commas in writing about “observations” of theoretical facts: such use is not meant to suggest any unfavorable contrast with our observations, it is instead just a reminder that it is always in this, theory-dependent, way that we use observations to ascertain the occurrence of theoretical facts.

<sup>5</sup> The discussion in the text assumes that the conditions described by P are observational. What if they are theoretical? That is, what if the scientific principle we want to test says that a certain theoretical fact T occurs whenever another theoretical fact T’ occurs? How does observational testing work for principles of this type? In order to check whether the occurrence of T really accompanies the occurrence of T’, there would need to be (for the reasons referred to in the text) some observational fact O’ we take as evidence for the occurrence of T’. We would then have to check whether the occurrence of the theoretical fact T really accompanies the occurrence of O’ (and therefore of T’), which brings us back to the situation discussed in the text.

evidence for the emission of protons, and therefore “observe” «There goes a proton». Does our “observation”, in such a case, provide us with a test for P? is it enough to conclude we have tested, and confirmed, P? It seems clear that it is not: principle P has not undergone observational testing. If *that* is the observational evidence we use to “observe” the emission of protons, what we are really comparing P with are our beliefs: we have a more general, subsuming belief that applies (among other cases) also to gas X under conditions C, and in performing our experiment *that belief* leads us to judge that some protons are emitted. But that is not what we are looking for: we want observational testing to compare P with the world, not with our beliefs about it.

The point is not that, in this way, we would always end up confirming P. Our background beliefs might as easily be in contrast with P, and so our “observation” might be that the theoretical fact predicted by P does not occur (just suppose, for example, that our background beliefs include, instead of P\*, its contrary): what that would show, though, is just that some of our beliefs speak against P, not that the world does. The point is that, insofar as we take the occurrence of conditions C as the observational evidence relevant to determine whether any protons are emitted, the most our “observation” can show – whether it is in accordance or in contrast with P - is that P does / does not conform to our beliefs, not that it does / does not conform to the way the world is. If the background beliefs we rely on in making our “observation” take *that* as the relevant observational evidence, the most our “observation” can do is to test P against our beliefs, while we want observational testing to test P against the world<sup>6</sup>.

---

<sup>6</sup> Since our purpose is to compare scientific and moral principles, the following comparison can be useful. The situation described in the text is analogous to what often happens when we want to test a moral principle by checking how well it fares in a particular case: what we often really do in these situations is to check whether the principle is in accordance or in contrast with some other moral beliefs we already hold. For example, suppose that, while wondering whether torturing animals is morally wrong or permissible, I see a group of young hoodlums set fire to a cat. That might remind me that an animal is, after all, a sentient being, and so torturing it cannot be morally permissible. In such a case, it seems that what I really do is to test the principle “Torturing animals is morally permissible” against some other moral belief I already hold (like “Torturing sentient beings, human and non-human animals alike, is never morally permissible”), and to realize that it is in contrast with it.

True, we might have good observational support for our more general theoretical beliefs, and so (in a derivative way) for our acceptance / rejection of P based on them: for example, we might have good observational support for P\*, and therefore for P, and that would be enough to have P exposed to, and confirmed by, observational testing. But that would just push our question one step further: we want to understand how observational testing works for scientific principles, so we pick out principle P and ask “How does observational testing work for P?”; we saw that, in order to expose P to observational testing, we need to rely on a body of background theoretical claims; and that if these background claims take the very occurrence of conditions C as the observational evidence relevant to determine whether any protons are emitted, we do *not* expose P to any observational test, *unless* we have already tested, and found observational support for, the more general principle P\*, under which P can be subsumed. But our question then becomes “How do we expose P\* to observational testing?”, and we can run the same argument again. We still have to understand how we manage (as we certainly do) to expose scientific principles to observational testing<sup>7</sup>.

Let me illustrate the point in a slightly different way. Suppose we are performing our experiment: we have set up conditions C, and are checking out what happens to gas X, in order to “observe” whether any protons are emitted. Suppose one of the scientists we are performing the experiment with - let’s call him John – relies, in making his “observations”, on principle P\*. By relying on P\*, John sees gas X under conditions C and so “observes” that some protons are emitted (it is obvious how to extend these considerations to the case in which John relies, rather than on P\*, on its contrary, and so “observes” that no protons are emitted). Is John’s “observation” enough to conclude we have tested, and confirmed, P? It

---

<sup>7</sup> Going back to our moral parallel: I start by asking “Should I accept that torturing animals is morally permissible?”, see the hoodlums’ action, and realize that the principle “Torturing animals is morally permissible” is in contrast with the more general principle “Torturing sentient beings is never morally permissible”. That just turns the original question into “Should I accept that torturing sentient beings is never morally permissible?”. After all, someone unmoved by the cat’s suffering might claim that that is precisely one of the cases in which the more general principle fails.

seems clear that, being made by relying on P\*, John's "observation" does not provide us with a test for P: what it tests P against are really John's beliefs, not the world. John has a more general belief that applies (among other cases) also to gas X under conditions C, and in performing the experiment *that belief* leads him to judge that some protons are emitted. John might, of course, get convinced of the truth of P in this way. If he did, though, he would not get convinced by any observational support he has got for P: he has got none. He would not, for instance, be able to convince us that our test has succeeded, that we have exposed P to observational testing and found observational support for it, and so that we should accept P: he has nothing to point to as such observational support. All he could point to is the more general principle P\*.

That would be enough, of course, if we had already got observational support for P\*. But that was not what we, John included, set out to do. We set out to see whether P can get any observational support: so far we have found none. If someone, say John, now proposes that P inherits its observational support from P\*, that does not answer our question yet: we have not yet been shown the observational support for P, until we are shown the observational support for P\*. So that just turns our question into whether P\* can get any observational support.

"Observations" of theoretical facts are theory-dependent: they require a background theoretical claim that enables the "observer" to take some observational fact as evidence for the occurrence of the theoretical fact in question. We saw that, if we want to test P, the theoretical claim we rely on in making our "observations" cannot take the occurrence of conditions C as the relevant observational evidence: what we need, therefore, is a theoretical claim that takes some observational fact different from the occurrence of conditions C as evidence for the emission of protons. What we typically use for this purpose, according to my emended Harman, are observational facts that protons *cause*. Harman's example is a vapor trail in a cloud chamber. We believe that a free proton going through a cloud chamber causes

a vapor trail we can observe. When we set up an experiment to test P, we use this theoretical belief to run the experiment: once we have set up conditions C, we can “observe” whether any protons are emitted by checking whether we observe any vapor trails. If we do, and on the basis of *that* observational evidence we “observe” «There goes a proton», our “observation” does provide us with a test for P, and we can conclude that P has been tested, and confirmed.

Let’s go back to our example about John: we have set up conditions C, are checking out what happens to gas X in order to “observe” whether any protons are emitted, and John “observes” «There goes a proton». Let’s suppose this time that the observational evidence he has for his “observation” is not “Gas X is under conditions C (so under conditions C\*)”, but “There is a vapor trail”. This time John *has* got observational support for P, and has something to point out to us, in order to convince us that our test has succeeded and we should accept P: we wanted to check whether, under conditions C, a proton would be emitted, and John is pointing to one of its observational effects.

What is the difference between relying on claims like P\* or its contrary, on the one hand, and relying on claims like “Free protons cause vapor trails”, on the other, in making our “observations”? In both cases we have two theoretical claims: the one we want to expose to observational testing, namely P, and the one we rely on in making our “observations”. In the first case, though, we have two theoretical claims both applying to gas X under conditions C: they can or cannot be in accordance with each another, they can or cannot agree about what will happen to atoms of gas X under conditions C, but they do not generate any observational claim. In the second case, on the other hand, the two theoretical claims generate the observational claim “If gas X is under conditions C, some vapor trails are produced”, which then the world can or cannot be in accordance with. In the first case, the confirmation / rejection of P comes from our background theory: we have P, and we test it against some more theory. In the second case, the confirmation / rejection of P does not come from our background theory, but from the world: we have P, and let the world speak about it. True, in

order to get an observational claim, and so compare P with the world, we need a further bit of theory: we need to rely on some background theoretical claim. But that is just the general point that “observations” of theoretical facts are always theory-dependent: we can never expose *single* theoretical claims to observational testing, we always need to rely on some more theory. Given that that is so, given that we cannot but rely on some more theory, the question is: what do we do with this further bit of theory? Do we use it to compare P with *it*? Or do we use it to compare P, *via it*, with the world? It is only if we do the latter that P can undergo observational testing. And in order to do the latter, the theoretical claims we rely on have to be of the right type: what we typically use for this purpose are causal claims stating that some observational fact, different from the occurrence of the conditions described by P, is causally explained by the theoretical fact predicted by P.

Let's imagine a discussion slightly different from the previous one between John and us. Two physicists, A and B, discuss about the truth of P, and each of them has already a view on the matter: A thinks P is true, B thinks it isn't. We believe they have a way to try to solve their dispute, a way that involves appeal to observational testing: we think observational evidence can be brought to bear on scientific disputes, and that it is relevant - characteristically and crucially relevant - to our choice among competing scientific principles. But how does observational evidence play this role? Suppose A and B set up conditions C and try to argue by appealing to what they “observe”. It would not solve the dispute to appeal to “observations” made by relying on background beliefs that take the occurrence of conditions C as the relevant bit of observational evidence: A, for instance, relying on P\*, and B on its contrary. A would “observe” that some protons are emitted, and if asked «Why?», she would answer «Look, gas X is under conditions C, so under conditions C\*». B, of course, would be totally unmoved by such an “observation”: «So what? No wonder you “observe” that some protons are emitted, if you believe that P\* is true. But that provides no test for P, and in particular it gives me no reason to change my mind about it (that only proves that the one

under consideration is precisely one of the cases in which P\* fails)». And A would be allowed to think the same, *mutatis mutandis*, about B's "observation" that no protons are emitted.

If we want to bring observational evidence to bear on the dispute, it has to be some bit of observational evidence different from the occurrence of conditions C: Harman's suggestion is that what we typically use for this purpose are the observational facts protons cause. If A and B agree at least on the claim that free protons in a cloud chamber cause vapor trails, observational evidence can be brought to bear on their dispute in the following way: if, after setting up conditions C, they observe a vapor trail, A can point to that fact as a confirmation of P; if instead, under those conditions, no trails are observed, B can point to that as a rejection of P<sup>8</sup>.

Let us now generalize these considerations about the testing process in the empirical sciences:

- (a) in order to expose a scientific principle SP to observational testing, we need some observational fact whose occurrence we can take as evidence for the occurrence of the theoretical fact predicted by SP;
  
- (b) the observational facts we can use for this purpose have to be different from the occurrence of the observational conditions described by SP. Let's label these observational facts "observational criteria" of the occurrence of the theoretical fact predicted by SP;

---

<sup>8</sup> Of course it is still open to B (or A) to claim that, in the case under consideration, there might be a different explanation of the trail (or the absence thereof): the purpose here is not to give observational evidence a conclusive role in the dispute, but to give it a role to play.

It might be objected that A and B have to agree at least on "Free protons cause vapor trails", in order for A's pointing to a vapor trail to force B to change his mind, and that therefore there is no real difference with the other scenario: if B agreed with A at least on P\*, then A's pointing to the fact that X is under conditions C (so under conditions C\*) would force B to change his mind. The difference is that, though in both cases B gets convinced of the truth of P, it is only in the first case that B gets convinced by some observational support he gets for it, while in the second case he gets none: as I put it above, it is only in the first case that A and B use the background theoretical claim they share to compare P, via it, with the world; in the second case they share a background theoretical claim too, but they use it to compare P with it.

- (c) what we typically use as observational criteria are observational facts that are causally explained by the theoretical fact in question.

Points (a) – (c) amount to the following condition for observational testing:

(COT)        *a scientific principle SP can undergo observational testing if the theoretical fact it predicts causally explains at least some observational facts we can use as observational criteria for it.*

COT states a sufficient condition: as such, it does not say that the *only* way a scientific principle can undergo observational testing is by meeting the condition it specifies. In particular, it is because of point (c) above that COT states only a sufficient condition and not also a necessary one: what we need is an observational fact we can take as evidence for the theoretical fact predicted by SP, and in principle we can use any such observational fact (different from the observational conditions described by SP), whether it is caused by the theoretical fact in question or not. In the case of our principle P, for example, one possible observational criterion would be an observational fact O (different from the occurrence of conditions C) which occurred whenever some protons are emitted, even if its occurrence were not caused by their emission. Point (c), though, reflects one of the central features of the process of empirical theory-construction, as we see it exemplified by our current scientific theories: namely, the idea that «scientific principles can be justified ultimately by their role in explaining observations, [...] by their explanatory role» (Harman [1977], p. 9). What is true about our scientific theories (for instance, our claims about protons) is not only that the way in which we develop them is responsive to, and constrained by, observational evidence, but also, more specifically, that this constraint is quite often an explanatory one: scientific

theories are constrained by observational evidence also, and characteristically, in the sense that they aim at providing a causal explanation of the observational evidence available (for instance, of the occurrence of vapor trails in a cloud chamber). Scientific theories aim at describing the theoretical reality which underlies and causally explains the empirical phenomena we observe, and therefore we are justified in accepting them, and in believing in the existence of the theoretical entities they postulate, because and insofar as they provide causal explanations for such empirical phenomena. It is true that meeting the condition specified in COT is not the only way in which a scientific principle can possibly undergo observational testing, but it is the fact that scientific principles do meet that condition, and are tested in the way described by COT, what accounts for our confidence that our scientific theories give us a causal explanation of the empirical phenomena, and that we can believe in them (and in the existence of the theoretical entities they posit) because of that.

The relevance of COT is not diminished, therefore, by the fact that meeting the condition specified in it is not the only way in which a scientific principle can possibly undergo observational testing. It is not diminished by the fact that scientific principles can (and quite often do) undergo a different kind of test either. Observational evidence is not the only kind of evidence available for scientific theories: theoretical considerations too are (and are treated by scientists as) evidential<sup>9</sup>. Several parts of the theoretical picture scientific theories describe are not tested directly against observational evidence, but are posited on the basis of theoretical considerations, licensing inferences from previously acquired theoretical knowledge to new theoretical conclusions<sup>10</sup>. For instance, the causal-explanatory link between

---

<sup>9</sup> This is a controversial claim in philosophy of science: those who do not accept it will find the case I am making for the crucial relevance of COT even more persuasive. For a defense of the evidential role of theoretical considerations in science, both in general and with regard to specific aspects of the process of theory confirmation, see Boyd [1973; 1982; 1983; 1985].

<sup>10</sup> This amounts to the claim that we can get evidential support for a theoretical hypothesis by testing it against previously established theoretical claims: the difference between this case and the case I exemplified above with the two principles P and P\*, about which I claimed we cannot get any observational support for P merely by subsuming it under P\*, is that the inferences I am talking about now are licensed by theoretical considerations, not merely by logic. These theoretical considerations are (as I put it above) what one can point to as the evidential support for the hypothesis under examination. This point is also what differentiates the case

the theoretical realm described by our scientific theories and the empirical phenomena it aims to explain is not always so simple and direct as the discussion about our principle P might lead to think: the theoretical realm postulated by our sciences is typically highly structured, and it is not the case that every single theoretical entity, mechanism, process, etc. has to earn its place in it by its direct contribution to the explanation of some observational phenomenon. The view that the only thing scientists do to solve their disputes and gain further scientific knowledge is to run an experiment and appeal to observational evidence is inadequate to actual scientific practice. But even if observational evidence is only one element in the mix of evidential considerations scientists avail themselves of, it is still an essential element of that mix: a theoretical enterprise which did not proceed by testing its total set of theoretical claims not only against theoretical considerations but also against observational evidence would not count as empirical science<sup>11</sup>.

So even if we have other ways, beyond the one described by COT, of testing a scientific principle, COT identifies a crucial feature of the testing process in the empirical sciences. Any theoretical enterprise which lacked this feature of the scientific enterprise would look very different from it: observational testing would be either absent, or anyway working in quite a different and much more limited way than the way it works within empirical science. This is, according to Harman's argument, how things are with ethics. The problem with ethics is that moral principles do not satisfy the condition stated in COT, therefore they cannot undergo observational testing in the same way scientific principles can: we cannot test moral principles in the same way we test scientific principles.

---

exemplified by P and P\* from those in which scientists make use of what is usually called their "intuition". Scientists sometimes test a scientific hypothesis against their "sense" of what should happen in certain circumstances, a "feel" they have acquired in virtue of their professional training and working familiarity with both theory and scientific practice, and accord these tests evidential weight. I believe they are justified in so doing, but also believe the evidential value of their "intuition" is an instance of the evidential value theoretical considerations have in science, so I think that what goes on in such cases is quite different from the P-P\* case: when scientists use their "intuition", they do something much more complex, and crucially theory-dependent, than just subsuming a scientific hypothesis, according to a rule of logic, under a more general scientific claim.

<sup>11</sup> Something stronger, I believe, is true: the evidential value of each of the two kinds of tests depends on that of the other. If this is so, then in the absence of observational tests, theoretical considerations would lose their evidential character too.

## 2. Why moral principles do not satisfy the condition for observational testing

According to COT, we can expose scientific principles to observational testing because scientific facts causally explain at least some observational facts we can use as observational criteria for them. If we want to expose moral principles to the same kind of observational testing scientific principles undergo, the same has to be true about moral facts. The source of the problem with ethics is that this is not the case: moral facts do not cause any observational facts we can use as observational criteria for them.

Consider the moral fact that a given action  $a$  is morally wrong. Like with any other theoretical fact, the “observation” of this moral fact is theory-dependent: we can “observe” that  $a$  is morally wrong only by observing some observational fact that we (via our background moral beliefs) take as evidence for the occurrence of the moral fact in question<sup>12</sup>. In ethics, though, something more specific appears to be true: namely, that the observational facts we can use as constituting such evidence are facts about the non-moral properties of  $a$  that make  $a$  a wrong action. For instance, we can “observe” that  $a$  is wrong by observing that it is an act of torture toward an animal, and relying on the background moral belief that torturing animals is wrong; or by observing that  $a$  is a theft, and relying on the background moral belief that stealing is wrong. In these examples we “observe” that  $a$  is wrong by observing it has those non-moral properties that make it wrong: the background moral belief we rely on in making our “observation” is a belief to the effect that all actions having those non-moral properties are wrong because of that, i.e. a belief in what we call a moral principle. Is there another way we can “observe” that  $a$  is wrong? It seems there is not. Let’s take Harman’s example: I round a corner, see a group of young hoodlums pour gasoline on a cat and ignite it, and “observe” «That’s wrong». What observational facts could I take as

---

<sup>12</sup> The application to ethics of the thesis about theory-dependence predates its defense as a general claim in philosophy of science, and is commonly referred to as the doctrine of the autonomy of ethics or the is-ought gap.

evidence for the moral fact I “observe”? Maybe the observational fact I take as constituting such evidence is the fact that the hoodlums’ action is an act of torture toward an animal; or it might be the more general fact that their action is an act of torture toward a sentient being. Someone else might also “observe” that the hoodlums’ action is wrong, but take very different observational facts as evidence for the occurrence of the moral fact she “observes”. For example, the observational fact she takes as constituting such evidence might be that the hoodlums’ action is made in public, stains the street and presumably causes noise that will annoy people in the neighborhood; or it might be the fact that the action expresses nothing but a desire for violence and destruction on the hoodlums’ part. These are, of course, just very few examples of the vastly different observational facts one could take as evidence for the moral wrongness of a given action *a*. Still, it seems to be true that, different as they may be, all such observational facts share this characteristic: they are all about the non-moral properties of *a* that make it the case that *a* is a wrong action. How could I “observe” that *a* is wrong, if not by relying on some belief as to what makes it wrong? More generally, it seems to be true that, for any moral property M and any object of moral evaluation *a*, we can only “observe” that *a* is M by observing that *a* has some non-moral property N that makes it the case that *a* is M: the background moral belief we rely on in making our “observation” has to be a moral principle (to the effect that all N objects are M because of their N-ness). Let’s label this thesis the “M-making property thesis” (MMP). MMP has seemed to many philosophers a very important and fundamental truth about ethics, usually presented as an obvious truth, a truism about morality. It is the key premise in my emended Harman’s argument for the conclusion that moral principles do not satisfy the condition stated in COT, and so cannot undergo observational testing in the same way scientific principles can<sup>13</sup>.

---

<sup>13</sup> What I labeled “M-making property thesis” is sometimes referred to as the thesis of the supervenience of the moral on the non-moral. Several different claims, though, are called “supervenience” claims in meta-ethics, and more than one of them has been thought to bear on the issue of moral explanations: in order to avoid confusion I prefer to use a new name. Moore [1903] (with regard to intrinsic value) and Hare [1952], to mention just two examples, both accept, despite all their differences, the truth and the character of truism of MMP.

Suppose we want to expose a moral principle to observational testing: for instance, let the moral principle be the claim that torturing animals is wrong. We see Harman's hoodlums set fire to a cat and set out to check whether their action really is, as our principle says, wrong. If we want to test our principle along the same lines we test scientific principles, we need an observational criterion for the moral fact that the hoodlums' action is wrong: that is, a non-moral fact, different from the fact that the hoodlums' action is an act of torture toward an animal, that we (via our background moral beliefs) can take as evidence for the fact that their action is wrong. Since MMP holds, such an observational criterion has to be a fact about the non-moral properties of the action that make it a wrong action: the background moral belief we rely on in taking such a fact as evidence for the action's wrongness has to be a moral principle. Do we have any such moral principles? There is no reason why we should not. Suppose, for instance, our background moral beliefs include the following principle: all actions tending to decrease the agent's sensitiveness to human pain are wrong. Such a background belief does provide us with an observational criterion for the moral fact predicted by the moral principle we want to test: we can "observe" whether the hoodlums' action is wrong by checking whether it has any tendency to decrease the hoodlums' sensitiveness to human pain. If we run some psychological tests on the children and observe that it does, we can point to that as a confirmation of the principle that torturing animals is wrong; if instead it does not, we can point to that as a (tentative) rejection of it<sup>14</sup>.

Or take the following example, due to N. Sturgeon<sup>15</sup>: consider the utilitarian principle that only actions producing a lesser net balance of pleasure over pain than some available alternative act are wrong. Suppose we have an action *a* which is such that no alternative act

---

<sup>14</sup> Of course, even if we do not observe any tendency of the hoodlums' action to decrease their sensitiveness to human pain, we might still try to find observational support for the moral principle under consideration in some other way, by appealing to a different background moral principle and a different observational criterion. As in the case of scientific principles (see n. 8), the purpose here is not to give observational evidence a conclusive role, but to give it a role to play in the testing process of moral principles.

<sup>15</sup> See Sturgeon [1988, pp. 231-232]. The example in itself is a standard counter-example to act utilitarianism: Sturgeon uses it to illustrate the epistemological point I am interested in.

open to the agent would have produced a greater net balance of pleasure over pain: we want to check whether it really is the case that, as our principle says, *a* is not wrong<sup>16</sup>. Suppose our background moral beliefs include the principle that murder is wrong: such a background belief does provide us with an observational criterion to “observe” whether *a* is wrong. If *a* is an act of murder, we can point to that as a refutation of the utilitarian principle; if it is not, we can point to that as a (tentative) confirmation of it.

So moral principles can undergo observational testing. We saw above, though, that there is more than one way a theoretical principle can undergo observational testing, and that what is crucial about scientific principles is that they undergo observational testing along the lines described by COT. So far we cannot say the same about moral principles. The mark of scientific principles is that they «can be justified ultimately by their role in explaining observations». We accept principle P, that under conditions C atoms of gas X emit protons, because its truth causally explains why, when we set up conditions C, we observe a vapor trail: under conditions C gas X emits protons, and these protons cause the vapor trails we observe. Let’s consider now the moral principle that torturing animals is wrong, and suppose that the observational test just described confirms it: we run our psychological tests on the hoodlums, and observe they really are less sensitive to human pain, so we conclude their action is wrong and the moral principle has been confirmed. In such a case we do expose the moral principle to observational testing, but we do not accept it because it causally explains why acts of torture toward animals tend to decrease sensitiveness to human pain. We do not accept that acts of torture are wrong because their wrongness causes their observed tendency to decrease sensitivity to human pain: we accept they are wrong because this observed tendency is what makes them wrong acts.

---

<sup>16</sup> So far my examples of observational tests have been about theoretical principles predicting the occurrence of a theoretical fact, tested via an observational criterion of the occurrence of that fact. This example is instead about a theoretical principle predicting the non-occurrence of a theoretical fact, tested via an observational criterion of the occurrence of that fact. And other combinations are possible too. These cases may require slight modifications to some of the claims I made on the previous pages. I omit them for brevity’s sake: it is easy enough to make the necessary modifications, if needed.

Or take our second example. Suppose action *a* is an act of murder: by relying on our background moral belief we “observe” that *a* is wrong, and so conclude the utilitarian principle has been rejected. Compare this situation to the following one: we set out to test the principle that under conditions C atoms of gas X emit no protons, set up conditions C, observe a vapor trail, and conclude the principle has been rejected. In the scientific case, we reject the principle because we observe something (a vapor trail) which is caused by a free proton, while the principle predicted no protons would be emitted. In the moral case, on the other hand, we observe *a* is an act of murder, but it is not the case that *a*’s being an act of murder is caused by its being wrong: rather, *a*’s being an act of murder is what makes it the case that *a* is wrong. But then our rejection of the moral principle is not justified (as our rejection of the scientific principle instead is) by its role in explaining, causally speaking, what we observe.

MMP is true: we can only “observe” that an object of moral evaluation *a* possesses a moral property M by observing that *a* has some non-moral property N that makes it the case that *a* is M. That is: we can only “observe” that *a* is M by relying on some background moral principle to the effect that all N objects are M because of their N-ness (for some non-moral property N). So, if we want to test the moral principle that torturing animals is wrong, we need another moral principle to rely on, in order to “observe” whether the hoodlums’ action actually is, as our principle says, wrong. So far we have considered examples of background moral principles that provide us with observational criteria, like the principle that all actions tending to decrease the agent’s sensitiveness to human pain are wrong. We saw that, through them, moral theory can be constrained by observational evidence, but not in the same way scientific theories are: namely, causally. Of course, there is no reason why we should not have other background moral principles as well, providing us with observational facts that we can take as evidence for the moral fact that the hoodlums’ action is wrong, but that are not observational criteria of that fact: for instance, the principle that torturing sentient beings, human and non-human animals alike, is wrong. Since they do not provide us with

observational criteria, though, these background moral beliefs do not allow us to expose moral principles to observational testing. By relying on the background moral belief that torturing sentient beings is wrong, we can “observe” that the hoodlums’ action is wrong: such an “observation”, though, does not provide us with a test for the principle that torturing animals is wrong. As I put it above, we would not be testing the principle against the world.

So moral facts do not cause any observational facts we can use as observational criteria for them because, since MMP is true, for any object of moral evaluation *a* and any moral property M, the observational facts we can take as evidence for the occurrence of the moral fact that *a* is M (whether they are observational criteria of that fact or not) are facts of the form “*a* is N”, for some non-moral M-making property N: but then it is not the case that *a*’s being M causes its being N; rather, it is *a*’s being N that makes it the case that *a* is M. If the observational fact we use as evidence is not an observational criterion, we do not even expose the moral principle to observational testing; if it is, we expose the moral principle to observational testing, but not along the lines described by COT. In any case, moral principles cannot undergo observational testing in the same way scientific principles can.

If my emended Harman is right about this argument, then our epistemological prospects in ethics look quite dim. Suppose A and B are discussing about whether torturing cats is morally wrong: A maintains torturing cats is wrong, while B cannot see anything wrong with it. They round a corner and see a group of young hoodlums set fire to a cat, so A sets out to convince B by appealing to the fact that the hoodlums’ action is wrong. One way their discussion can go is along the following lines: A says «Look at what those children are doing: that is wrong», and if asked «Why?», she answers «It is an act of torture toward an animal». B, though, might be totally unmoved by such a remark: «Well, if you believe that torturing animals is wrong, then of course you find something wrong in torturing cats. But I still cannot see what is wrong with it. If you want to convince me, you should show me there is something wrong with torturing animals. And I cannot see anything wrong with that either».

A might insist «But it's an act of torture toward a sentient being», but B might be unmoved by such a remark too: «Of course *if* torturing sentient beings is wrong, then so is torturing cats. But that does not show torturing cats to be wrong, and does not give me any reason to change my mind about it. That only proves that the one we are considering is precisely one of the cases in which torturing sentient beings is not wrong. I agree that torturing a human being is wrong, but cannot see anything wrong with torturing non-human sentient beings: so I cannot see anything wrong with, in particular, torturing cats». In a discussion like this, observational evidence cannot even be brought to bear on the dispute, on the choice between the two competing moral positions: the moral dispute between A and B proves refractory to being constrained by observational evidence. And of course this problem would persist even if A managed to convince B in the end: even if, for instance, B, after A's remarks focused her attention on the fact that a cat is a sentient being, finally saw what is wrong with torturing them. In such a case, B would get convinced, but not by any piece of observational evidence she has been provided with by A: she has got none. She would simply test her moral belief that torturing cats is not wrong against another moral belief she already holds, find them incompatible with each other, and choose the latter over the former. So one way our moral discussion can proceed is by testing the moral principle under examination against our more general moral beliefs. This process can show us that the principle under consideration is in accordance or in contrast with some of these beliefs, and in response to that we can accept the principle and revise those beliefs, or stick to our beliefs and reject the principle. People do sometimes change their views during this process, coming to accept new things they previously rejected: the point, though, is that in such a process moral principles are tested against our moral beliefs, never against the world. We check whether the moral principle under examination is in accordance or in contrast with our background moral theory: the

confirmation or rejection of the principle comes from our background theory, not from the world<sup>17</sup>.

Another way A can go, in order to convince B, is the following: she says «Look at what those children are doing: that is wrong», and when asked «Why?», she answers «It tends to decrease their sensitiveness to human pain». If B shares A's background belief that actions tending to decrease the agent's sensitiveness to human pain are wrong, she cannot dismiss A's remark as quickly as she did before: «Really? I don't believe it does. Anyway, let's go and check». In such a case, the moral dispute between A and B is constrained by observational evidence, but the constraint is not a causal-explanatory one. If B finds out that the hoodlums' action does tend to decrease their sensitiveness to human pain, and on that basis concludes it is wrong after all, thus changing her mind about torturing cats, she will not do that in an attempt at providing a causal explanation of what she found out. So another way our moral discussion can proceed involves a constraint by observational evidence, but in a very limited way indeed, much more limited than the way scientific discussion is constrained by observational evidence. Science aims at explanation, so the constraint observational evidence imposes on the process of theory-construction in the empirical sciences is explanatory. This kind of constraint is absent in ethics, and so are its epistemological advantages. An explanatory constraint on the process of theory-construction is much stricter than a general constraint to be in accordance with the observational evidence available, so it is much more

---

<sup>17</sup> It might be objected at this point that there must be something wrong with the argument, because the process I have just described is what in the literature is called (following Rawls [1971]) the method of reflective equilibrium, and surely this method is a way of getting evidential support for moral principles. While I accept that the method of reflective equilibrium has evidential value, I deny it comes to the process I describe in the text: general considerations of coherence, broadly construed as not just mere consistency, and the appeal to the plausibility of general principles of various sorts, to mention two elements, play an important role in the method of reflective equilibrium. Actually, my own view of this method closely parallels my view of the use of theoretical considerations and scientific "intuition" in the empirical sciences (see n. 10): it consists in testing a moral principle against a background of previously established moral conclusions on the basis of inferences licensed by background moral theory, and not just in subsuming a moral principle under a more general one. Moreover, since my emended Harman's argument is an argument against ethical naturalism, it is dialectically relevant, in response to this objection, to observe that the view according to which the method of reflective equilibrium is closely parallel to the use of theoretical considerations and "intuition" in science is the view adopted by the best defenders of ethical naturalism who have written on the subject: see Boyd [1988] and Sturgeon [2002] (I have been heavily influenced by these papers).

useful when it comes to bringing observational evidence to bear on the choice among competing theoretical positions. We can take advantage of this in science, we cannot in ethics.

How adequate is this picture to our actual everyday experience? Isn't it true, after all, that moral discussion is as limited as Harman's argument depicts it to be? We saw the key premise in the argument is MMP. MMP says that, for any object of moral evaluation  $a$  and any moral property  $M$ , we can only "observe" that  $a$  is  $M$  by observing that  $a$  has some non-moral property  $N$  that makes it the case that  $a$  is  $M$ . MMP thus limits the observational facts we can take as evidence for the occurrence of a moral fact to one type only: unfortunately for our epistemological prospects in ethics, these observational facts are not caused by the moral fact for whose occurrence they are evidence. What we would need, in order to expose moral principles to the same kind of observational testing we expose scientific principles to, are observational facts caused by the moral fact that  $a$  is  $M$ : if there were such facts, we could take *them* as observational evidence for the occurrence of the moral fact causing them. But isn't it an obvious mistake to suggest we can do this in ethics? It amounts to the suggestion that we can "observe" that an object of moral evaluation  $a$  possesses a moral property  $M$  by observing an observational effect of its  $M$ -ness, that we can say « $A$  must be  $M$  because, look, there is one of the observational effects of its  $M$ -ness» in the same way we can say «A proton must have been emitted because, look, there is one of the observational effects of a free proton: a vapor trail». Isn't it obvious that we cannot detect  $a$ 's  $M$ -ness in this way, without having to rely on any belief about  $a$ 's  $M$ -making properties? Isn't MMP right, after all, in saying we can only judge that  $a$  is  $M$  by observing it has the non-moral properties that make it an  $M$  object?

The feasibility of the naturalistic project in ethics depends on the plausibility of a negative answer to these questions. This is what I meant when I wrote that, whatever one's final assessment of Harman's argument, what one can learn from it is how the epistemology of ethics has to be like, if ethical naturalism is to turn out true. If the naturalist wants to

maintain that moral facts are known in pretty much the same way in which scientific facts are known, she has to claim that ethical theory shares the crucial feature of the scientific enterprise COT identifies. My emended Harman's challenge to ethical naturalism then is: does ethical theory share that feature? can the naturalist make a good case for the claim that moral facts do cause some of the observational facts we can use as observational criteria for them? In particular: if this is what has to prove true, for the naturalistic project to be feasible, the naturalist has to deny that our epistemological access to the moral domain is limited in the way MMP says it is. She has to claim we can ascertain the occurrence of the moral fact that  $a$  is M by observing the occurrence of some non-moral fact the former causes, and not by observing some fact about the non-moral properties of  $a$  that make it an M object: can this claim be made plausible?

## Naturalists' replies

Moral principles cannot undergo observational testing in the same way scientific principles can, unless moral facts causally explain at least some observational facts we can use as observational criteria for them. But MMP is true, so moral facts do not cause any observational facts we can use as observational criteria for them. Therefore, moral principles cannot be tested in the same way we test scientific principles. So argues my emended Harman. As I hope the preceding discussion has shown, this argument poses a crucial challenge to ethical naturalism: how can the naturalist rebut it? If we look at naturalists' replies to Harman, we find they do not provide the resources for an adequate answer to it. As I said in the previous chapter, different philosophers take Harman's argument about moral explanations to be different things, so it would not be fair to criticize those of them who discuss the argument in order to rebut it because they do not address the problems my own reconstruction of Harman's argument raises: they do not set out to rebut *that* argument in the first place. What is nonetheless legitimate, once I have presented my emended Harman's argument and argued it poses a powerful challenge to ethical naturalism, is to ask whether naturalists' discussions of Harman's worry provide any help with this challenge. I'll ask this question about the most thoughtful reply to Harman available in the literature, namely Sturgeon's<sup>1</sup>. My claim will be that what one can learn from Sturgeon's reply is still not

---

<sup>1</sup> See Sturgeon [1988] (originally published in 1985). A further exchange between Harman and Sturgeon consists in Harman [1986] and Sturgeon [1986a]. Sturgeon's arguments provide not only what might be called the standard reply to Harman's worry on behalf of ethical naturalism, but also the best sustained defense of the acceptability of moral explanations and their relevance for both meta-ethical debates (in particular as part of an argument for ethical naturalism) and substantive ethical discussions. Although my discussion here will be mainly critical, I have been deeply influenced by Sturgeon's ideas, as is clear from my evident debts to him throughout this work.

enough to answer my emended Harman's challenge, and so that my emended Harman points to a problem about moral explanations ethical naturalists have not answered yet. Isolating Harman's argument from other arguments one might mount against ethical naturalism, Sturgeon writes: «What I *show* is that any remaining reservations about the existence of moral facts must be based on those *other* skeptical arguments [...]. In short, there may still be a “problem with ethics”, but it has *nothing* special to do with moral explanations»<sup>2</sup>. If I am right, the objection from moral explanations is instead still alive and well; the naturalist has still some work to do on the issue of moral explanations: quite important work indeed, given how crucial observational testability along COT's lines is to the scientific enterprise.

### 1. Sturgeon's reply

Here is what Sturgeon takes the worry about moral explanations to amount to. Our “observations” of scientific facts can be justified by their truth's role in explaining why we make them. Suppose a physicist observes a vapor trail, and takes that as evidence for the emission of a proton, thereby “observing” «There goes a proton». Is the physicist's belief in the emission of a proton justified? We think it is, provided his making that “observation” is explained by the actual occurrence of the scientific fact in question; but that is exactly what seems to go on in the physicist's case: the actual emission of a proton is what explains the physicist's “observation” «There goes a proton». If no proton had been emitted, the physicist would not have made that “observation”. Suppose now we see Harman's hoodlums set fire to a cat, and “observe” «That is wrong»: can our moral belief be justified in the same way? In

---

<sup>2</sup> Sturgeon [1988, p. 238], his emphasis.

this case, the actual wrongness of the hoodlums' action seems to play no role in the explanation of our making that "observation": it seems our "observation" is rather explained by the fact that the children are torturing an animal and by some moral belief we hold, condemning torture toward animals. Had the children action not been wrong, we would still have made that "observation", given our moral beliefs. Scientific beliefs can therefore be justified in a way moral beliefs cannot: by being explained by their truth. We can be justified in believing that a given scientific fact occurs by the fact that its occurrence explains why we "observe" it; the occurrence of a moral fact, instead, plays no role in the explanation of why we "observe" it obtains.

Sturgeon's reply to this argument goes as follows. The argument claims that our "observation" «That action is wrong» is not explained by the moral fact that the children's action is wrong. In order to assess this claim, we need to distinguish two cases, according to whether we do or do not, in assessing it, trust our moral beliefs. Suppose that, in evaluating the explanatory relevance of the moral fact in question, we do trust our moral beliefs. Then it seems the moral fact is explanatorily relevant: if the hoodlums' action had not been wrong, it would not have been an act of torture toward an animal (since this is what our moral beliefs tell us), so we would not have "observed" it is wrong. If instead we evaluate the explanatory claim without assuming our moral beliefs, then the moral fact does look irrelevant to the explanation of our "observation": if the children's action had been an act of torture toward an animal, but (contrary to what our moral beliefs tell us) it had not been wrong, we would still have "observed" it is wrong, given our moral beliefs. So Harman's claim should be qualified: it is right, *if* we are not allowed to trust our moral beliefs. But:

1) exactly the same point holds about the physical case Harman wants to contrast the moral case with. If the physicist had observed a vapor trail, but, contrary to what his physical theory tells him, no proton had been emitted, he would still have "observed" «There goes a proton», given his physical beliefs: the "observation" would have been made even if the

scientific fact in question had not occurred. We do believe the physicist's "observation" is explained by its truth, and so the physicist is justified in his physical belief, but that is so because we trust his physical theory in evaluating the explanatory relevance of the physical fact: if no proton had been emitted, there would have been no vapor trail (since this is what his physical theory says), so he would not have "observed" the emission of a proton. Harman, therefore, does not locate any difference between ethics and physics: if what he points to is any problem for ethics, it is a problem for physics too;

2) this point is no problem anyway, for either physics or ethics, since it is no reason for the claim that we cannot trust our physical (moral) theories, nor therefore for the claim that physical (moral) facts are explanatorily irrelevant;

3) there is no reason (apart from some bad epistemology, which Harman himself rejects) why we should evaluate whether moral facts are relevant to the explanation of our "observations" without assuming our moral theory: there is no privileged point of view, neutral among substantive moral theories, from which to decide whether moral facts are or are not explanatorily relevant. The only way we have of assessing their explanatory relevance is by assuming what seems to us the best moral theory, and we saw that, if we allow ourselves to do that, moral facts are often relevant to the explanation of our "observations";

4) of course, if we had reason to doubt our moral theory, this point would make us doubt the explanatory relevance of moral facts too. But Harman does not give us any such reason.

Therefore, Harman does not give us any independent argument for the claim that moral facts are explanatorily irrelevant. And he does not because he cannot: any argument for such a claim has to rest upon a prior, and different, skeptical argument. If there is a problem with ethics, it has nothing to do (directly) with moral explanations. Sturgeon's suggestion is that «the only argument [...] with any independent weight is the argument from the difficulty of settling disputed moral questions», which points to «the apparent difficulty of *settling* moral

disagreements, or even of knowing what would be required to settle them, a difficulty thought to be noticeably greater than any found in settling disagreements that arise in, for example, the sciences»<sup>3</sup>, so that is where discussion should focus.

Isn't Sturgeon's reply enough to rebut my emended Harman's argument? Can't a naturalist who accepts (as I do) Sturgeon's claims about the explanation of our "observations" find in this reply the resources he needs to answer my emended Harman's challenge? I do not think he can: the problem my emended Harman locates lies, so to speak, at a different level from the problem Sturgeon sets out to solve, and so from the type of problems in reply to which a Sturgeon-like strategy would be adequate. Sturgeon's aim is to show that our "observations" of moral facts can be explained by the actual occurrence of the facts "observed": he argues they can be so explained, provided we rely on our moral beliefs, and claims that our relying on moral assumptions in this way should not be considered problematic in the spirit of a naturalized epistemology (rather, it is exactly what is to be expected), and cannot be the reason why ethics differs from science anyway, since an exactly parallel point holds about the explanation of our "observations" of scientific facts. My emended Harman's question is whether moral principles can undergo observational testing in the same way scientific principles can. A Sturgeon-inspired naturalist might try to answer that they can, provided we rely on our moral beliefs, and claim that our relying on moral assumptions in this way should not be considered epistemologically objectionable, and that it cannot be what makes ethics different from science anyway, since an exactly parallel point holds about the observational testability of scientific principles. The problem with such an answer, though, would be that my emended Harman agrees with these claims: they are part of his argument. My emended Harman agrees that we can only expose scientific principles to observational testing by relying on our scientific beliefs: his argument *starts* with the claim that "observations" of theoretical facts are always theory-dependent, and so we can never

---

<sup>3</sup> Sturgeon [1988, pp. 254 and 229].

expose single theoretical claims to observational testing, but we always need to rely on some body of background theory. He agrees, therefore, that we can only expose moral principles to the same kind of observational testing we expose scientific principles to (if we ever can do that) by relying on our moral beliefs: he does not object to our relying on background moral assumptions in this way. So much granted, my emended Harman argues that these background moral beliefs we do have to rely on, if we want to expose moral principles to the same kind of observational testing scientific principles undergo, have to be of a specific type: they have to be causal claims about the observational criteria moral facts cause. The problem with ethics is that moral facts do not cause any observational facts we can use as observational criteria for them.

If we do not rely on background moral beliefs, we cannot expose moral principles to observational testing: it is possible to make different uses of this point, in developing an argument against ethical naturalism. One might think this point in itself is a problem for the observational testability of moral principles: if we want to bring observational evidence to bear on a moral issue, we cannot do that in a morally neutral way; observational evidence cannot provide an independent foundation for moral theory. So one might mount an argument against ethical naturalism starting from this point<sup>4</sup>. Against *this* argument, it would be adequate to reply that our relying on background moral beliefs in order to expose moral principles to observational testing is not problematic at all, and that this can hardly be a good starting point to differentiate ethics from science anyway, since an exactly parallel point holds about the observational testability of scientific principles. My emended Harman, for his part, does not object to our relying on moral beliefs in this way: rather, he takes it for granted that it is epistemologically all right that we so rely on them, and that this parallels what we do when we expose scientific principles to observational testing. The mere fact that observational testability is possible in ethics only if we rely on moral assumptions, he agrees, does not

---

<sup>4</sup> This is what the standard verificationist objection, pointing to the fact that moral principles are devoid of observational implications, did.

generate any problem in itself. He takes a step further, though, and argues that, even if our relying on moral assumptions is not problematic in itself, we have good reasons to require of these assumptions, if we are to be able to test moral principles in the same way we test scientific ones, that they be causal claims about the observational criteria moral facts cause. And this does generate a problem for ethical naturalism, since we also have good reasons to believe that moral facts do not cause any observational facts we can use as observational criteria for them.

The same considerations would hold if we tried to apply a Sturgeon-like strategy to the question “Do moral facts cause any observational facts we can use as observational criteria for them?”<sup>5</sup>. Since assessing a scientific causal claim is as theory-dependent as exposing a scientific principle to observational testing (that is, we can assess a causal claim, even in the sciences, only against a background of scientific assumptions about the field in question we already accept), a Sturgeon-inspired naturalist might reply that the only way we have of answering this question in the affirmative is by relying on some background moral beliefs, and claim that our relying on moral assumptions in this way should not be considered epistemologically objectionable, and cannot be what makes ethics different from science anyway. My emended Harman, though, does not need to deny these claims, because his argument does not object to our relying on background moral assumptions in assessing the causal power of moral facts. Granted that it is epistemologically all right that we rely on moral assumptions in this way, his argument is that since MMP is true, the observational facts we can take as evidence for the occurrence of a moral fact (including its observational criteria) are not caused by it.

So it is possible to see a problem about moral explanations even if one accepts the theory-dependence of our justification procedures in the sciences (and so accepts that the problem with our justification procedures in ethics cannot be that they are theory-dependent),

---

<sup>5</sup> As will be clear in a short while, this is Sturgeon’s actual strategy with regard to another example he discusses, in addition to Harman’s example of the hoodlums igniting a cat.

and even if one (in the spirit of a naturalized epistemology) believes the theory-dependence of these procedures does not prevent them from being *justification* procedures, from being able to yield justification for our beliefs. Sturgeon is right in claiming that the mere fact that our justification procedures in ethics are theory-dependent is not enough to generate an epistemological problem, and so in asking for a further reason why we should believe there is such a problem: my emended Harman, though, provides exactly these further reasons. Considerations from the epistemology of the empirical sciences give us reasons to believe that, if we want to expose moral principles to the same kind of observational testing we expose scientific principles to, the background moral assumptions we rely on have to be causal claims about the observational criteria moral facts cause. And considerations relative specifically to ethics give us reasons to believe that moral facts do not cause any observational facts we can use as observational criteria for them. This battery of considerations gives the objection from moral explanations independent weight: the argument from disagreement is certainly an objection ethical naturalists need to reply to, but it is not the only independent objection against their position. Even if the ethical naturalist's epistemology is naturalized, the objection from moral explanations is still an independent problem for her, worth independent discussion and rebuttal<sup>6</sup>.

## 2. (Some of) Sturgeon's examples

Sturgeon's claims, I have argued, do not provide an ethical naturalist with the resources for an adequate reply to my emended Harman's argument: as I put it above, the type of

---

<sup>6</sup> The issue of the consequences for meta-ethics of a thoroughgoing naturalism in epistemology is raised again in the following chapter.

problem Sturgeon sets out to solve lies at a different level from my emended Harman's problem. Some of the examples Sturgeon uses to illustrate his claims, though, can be used in a slightly different way, a way that might be helpful to the naturalist in search of a reply to my emended Harman's challenge: we must now turn to these examples.

Sturgeon's target, as we saw, is the claim that our "observations" of moral facts cannot be explained (as our "observations" of scientific facts instead can) by the actual occurrence of the facts "observed". He argues against this claim by using our "observation" that Harman's hoodlums' action is wrong as an example, in order to claim that that "observation" can be explained by its truth, provided we rely (as it is all right that we do) on our moral belief that torturing animals is wrong. Another example Sturgeon uses to defend his claim is the "observation" «Hitler was a morally depraved person», an "observation" many of us make after learning about Hitler's deeds. Of course, argues Sturgeon, if Hitler had done all he did, but (contrary to what our moral beliefs tell us) he had not been morally depraved, we would still have "observed" he was depraved, given our moral beliefs: this is how Harman can seem right in denying explanatory relevance to the moral fact in question; but we have no reason not to trust our moral beliefs, and so we can argue that, if Hitler had not been a morally depraved person, he would not have done what he did (since this is what our moral beliefs tell us), so we would not have "observed" he was depraved. That is, the moral fact in question is relevant to the explanation of our making the "observation" we do, provided we rely on our moral beliefs.

Sturgeon uses this example to illustrate another point too. He remarks that Harman's claim that our "observations" of moral facts are never explained by the actual occurrence of the facts "observed" commits him to denying that moral facts ever explain any non-moral facts at all, since if some moral fact MF explained some non-moral fact NMF, it would then be possible for us to "observe" the occurrence of MF by observing the occurrence of NMF, and in such a case the occurrence of MF would be relevant to the explanation of why we

“observe” it obtains. So Harman must deny that moral facts ever explain any non-moral facts at all, and not just our making the moral “observations” we do: for instance, he must deny that Hitler’s moral depravity explains his doing what he did. Sturgeon argues that Harman is wrong about this too, and his argument parallels his argument against Harman’s main claim that the occurrence of a moral fact never explains why we “observe” it. The reason why Harman can seem right in denying that Hitler’s depravity explains his deeds is that the moral fact in question does look explanatorily irrelevant *if* we are not allowed, in evaluating its relevance, to rely on background moral beliefs; but this point can hardly mark any difference between science and ethics, and it is (in the spirit of a naturalized epistemology) entirely irrelevant to the question of whether Hitler’s depravity explains why Hitler did what he did, since evaluating explanatory claims is a theory-dependent practice even in the sciences, and scientific facts also would turn out non-explanatory if we were not allowed to rely on background scientific assumptions in assessing their explanatory relevance.

So Sturgeon’s use of his Hitler example does not differ from his use of the children example: in both cases, his point is that our procedures in ethics are theory-dependent, but this theory-dependence cannot be what makes ethics different from science (since our procedures in the sciences too are theory-dependent), and should not (in the spirit of a naturalized epistemology) be considered epistemologically objectionable. As we saw, these claims are not helpful with my emended Harman’s argument, since my emended Harman agrees that the theory-dependence of our procedures in ethics is not problematic: the problem with ethics stems, according to him, from some further considerations. The Hitler example, however, can be used in a different way, namely to question these further reasons my emended Harman offers: it can be used as a counter-example to MMP.

MMP is the claim that for any moral property M and any object of moral evaluation *a*, we can only “observe” that *a* is M by observing that *a* has some non-moral property N that makes it the case that *a* is M (that is, the background moral belief we rely on in making our

“observation” has to be a moral principle, to the effect that all N objects are M because of their N-ness). MMP is the reason my emended Harman offers for the claim that moral facts do not cause any observational facts we can use as observational criteria for them: given the moral fact that *a* is M, the observational facts we can take as evidence for it (whether they are observational criteria of that fact or not) are facts of the form “*a* is N”, for some non-moral M-making property N; but then it is not the case that *a*’s being M causes its being N: rather, it is *a*’s being N that makes it the case that *a* is M. In the case of our “observation” «Hitler was a morally depraved person», though, we do not “observe” that Hitler was morally depraved by observing some fact about the non-moral properties of Hitler that made him depraved: we do not, for instance, “observe” he was depraved by observing he had psychological traits XYZ, which constitute depravity. We make our “observation” by observing a non-moral fact (his doing what he did) we take as evidence for the moral fact in question because we believe it is caused by that fact.

MMP limits the observational facts we can take as evidence for the occurrence of a moral fact to one type only: unfortunately for our epistemological prospects in ethics, these non-moral facts are not caused by the moral fact for whose occurrence they are evidence. What the Hitler example shows is that our epistemological access to the moral domain is not limited in the way MMP says it is: *contra* MMP, we can “observe” the occurrence of a moral fact by relying on a causal claim about its effects, rather than on a moral principle. The Hitler example is just one of a type of examples one can use to make this point, examples taken from our talk about both individual and social virtues and vices. It is quite common for us to draw conclusions about someone’s character, including their moral character, starting from what they did, or omitted to do; and our thinking about, say, social justice is shaped by the assumption that, other things equal, justice stabilizes a society, while injustice is a

destabilizing factor<sup>7</sup>. In all these cases, we can “observe” that a person is, say, mean (or that a social arrangement is unjust) not by observing that she has the non-moral properties that make her mean (or that the arrangement in question has the non-moral properties that make it unjust), but by observing what her meanness leads her to do (or what the unjust character of the arrangement leads some of the people living in that society to do). And apparently we can use these “observations” to test competing accounts of, say, what non-moral features of a social arrangement make it just: a theory of justice according to which it is realized in a society having non-moral features XYZ would face some difficulties if it turned out that social arrangements with those features are regularly characterized by a high degree of unstableness and strong pressure for reform.

MMP claims that the only type of observational facts we can take as evidence for the occurrence of the moral fact that *a* is M are facts about the non-moral M-making properties of *a*, hence facts that are not causally explained by the fact that *a* is M. If MMP is true, therefore, it follows that, as philosophers sometimes put it, our moral discourse and practice do not obey explanatory constraints: our attributions of moral properties are not shaped by explanatory purposes. As I said, MMP has seemed to many philosophers a truism about morality: my emended Harman points to the consequences of this supposed truism for the issue of the explanatory pretensions of morality (namely, that morality has no such pretensions), and remarks that these consequences generate a deep problem for ethical naturalism (given how crucial to the scientific enterprise observational testability along COT’s lines instead is). But despite its seeming a truism to so many philosophers, MMP is false, as examples like the ones I have mentioned show: our attributions of moral properties are sometimes shaped by

---

<sup>7</sup> These examples too are taken from Sturgeon, who appeals to this type of examples in different places: see, for instance, Sturgeon [1991; 2006b]. For the reasons discussed above, it should be clear that, though I think these examples are useful in a discussion of the objection from moral explanations, I do not think they are useful for the reason Sturgeon identifies.

explanatory purposes, our moral discourse and practice do obey explanatory constraints<sup>8</sup>. Moral principles cannot undergo observational testing in the same way scientific principles can, unless moral facts causally explain at least some observational facts we can use as observational criteria for them; but moral facts do not causally explain any such observational facts - so argues my emended Harman - because morality does not serve explanatory purposes: morality has no such explanatory pretensions. This reason, though, won't do, because it is false: as (some of) Sturgeon's examples are sufficient to show, important parts of our moral discourse do have explanatory pretensions.

So what a naturalist in search of a reply to my emended Harman can learn from Sturgeon is that MMP is false: not a little result, after all. MMP is the reason my emended Harman gives for the claim that moral facts do not causally explain any observational facts we can use as observational criteria for them, so it certainly is quite important for our naturalist to reject it. But, then, isn't that enough to conclude that my emended Harman's challenge to ethical naturalism can be answered, and that that challenge is not so damaging as it might at first appear to be? I do not think our naturalist is entitled to this conclusion. One of the main reasons why I reconstructed Harman's argument in the way I did in the first chapter is that I believe many of those who find the worry about moral explanations compelling have MMP in the back of their minds: so it is quite important to realize that MMP is false. But a further reason why I reconstructed Harman's argument as I did is that that reconstruction shows how the challenge to ethical naturalism can stand even after we reject (as I think we should) MMP. That is, I believe the problem for ethical naturalism Harman points to is deeper than many of those who find Harman's worry compelling take it to be. I must now try to explain why.

---

<sup>8</sup> Examples about virtues and vices are of course the first type of examples that come to one's mind, in order to show that our moral discourse has explanatory pretensions, but they are not the only type of examples one can use for that purpose: I'll give a different type of examples in the following chapter.

### 3. An empirical challenge

A crucial feature of the testing process in the empirical sciences is that scientific principles undergo observational testing along the lines described by COT. Therefore, if the ethical naturalist wants to maintain that moral facts are known in pretty much the same way in which scientific facts are known, she has to claim that moral principles can undergo observational testing in the same way: can they? They cannot, unless moral facts causally explain at least some observational facts we can use as observational criteria for them: do moral facts explain any such observational facts? My emended Harman is entirely right in pressing this question on the naturalist: what I claimed in the foregoing section is just that he is wrong in the reason he gives for answering it in the negative. Since morality has no explanatory pretensions, says my emended Harman, the answer has to be “No”: but, our naturalist can reply, morality does have explanatory pretensions. That, however, does not answer my emended Harman’s question yet, and we saw the naturalist owes us an answer (and a positive one) to that question. Do moral facts causally explain any observational facts we can use as observational criteria for them? Morality does have explanatory pretensions, sure enough, but what if such pretensions were not vindicated? Consider a type of examples I used above against MMP: what if it turned out that (as some social psychologists maintain) our tendency to explain one’s actions in terms of one’s character, including one’s moral character, is a fundamental error in reasoning we are subject to (the fundamental attribution error, as it has been called), people’s actions are not explained by their character but by situational factors, and there is no empirical evidence supporting the existence of relatively stable and long-term dispositions to act and feel in distinctive ways (as character traits are

supposed to be)<sup>9</sup>? In such a case, the answer to my emended Harman's question would still be negative: the naturalist's position would still prove defective, and defective exactly on the crucial point my emended Harman focuses his attention on.

Even if he is wrong in appealing to MMP, my emended Harman is entirely right in focusing on the question "Do moral facts causally explain any observational facts we can use as observational criteria for them?": if scientific facts causally explain some observational facts we can use as observational criteria for them, and if the fact that they do is a crucial feature of the scientific enterprise, then he is right in pressing that question on the naturalist, whose claim is that moral facts are known in pretty much the same way in which scientific facts are known. And, again even if MMP is false, it might very well turn out that the answer to that question is "No". The reason my emended Harman gives for believing that question is to be answered in the negative is a reason of principle: moral facts do not causally explain any observational facts we can use as observational criteria for them because any observational fact we *can* take as evidence for the occurrence of a given moral fact (whether it is an observational criterion of that fact or not) is such that it is not causally explained by it. The set of the observational facts *in principle* available in ethics is limited in such a way that it is true of any of them that it is not caused by any moral fact for which it can be taken as evidence. Denying MMP amounts to claiming that there is no such limitation: there are in principle no reasons why moral facts cannot causally explain any observational facts we can use as observational criteria for them. But, even if there are in principle no reasons why they cannot, it might still be the case that they, *as a matter of fact*, do not: and if so, that would be equally bad news for the naturalist. The reason why I believe my emended Harman's challenge to ethical naturalism stands even after we reject MMP is that the sting to the naturalist is in the

---

<sup>9</sup> For work in social psychology casting doubt on the existence of character traits, see Darley & Batson [1973] and Ross & Nisbett [1991]. It is interesting to note that Harman is one of the first philosophers who pointed to the relevance of these empirical studies for moral philosophy (though not, to my knowledge, with reference to the problem of moral explanations, his target being character-based versions of virtue ethics): see Harman [1999].

question my emended Harman presses on her. In order for her project to be feasible, the naturalist needs that the following be true: moral facts causally explain some observational facts we can use as observational criteria for them. Is that true? do moral facts causally explain any such observational facts? The naturalist cannot be satisfied with the answer “There are in principle no reasons why they cannot”.

A good way of seeing that she cannot is the following. Since we are interested in possible ways in which ethics can resemble the empirical sciences, it is useful to consider the case of abandoned empirical theories: take, for instance, phlogiston theory. It is of course false that phlogiston theory had no explanatory purposes: it certainly did. The problem with it was that such purposes were misdirected: as a matter of fact, the phenomena it purported to explain are instead explained by a different theory, appealing to entirely different theoretical facts and entities (facts about oxygen, rather than phlogiston). The negative epistemological consequences that, according to my emended Harman, follow from such lack of explanatory power hold true, of course, of phlogiston theory: principles about phlogiston cannot be «justified ultimately by their role in explaining observations». The interesting point is that they are true of it not because the theory had no explanatory pretensions (which is false), but because its pretensions were not vindicated. Therefore, the naturalist who denies MMP by appealing to the explanatory pretensions of our ethical discourse cannot believe her job is done: she might have taken morality out of the “theories with no explanatory purposes” frying-pan, just to put it into the “theories with misdirected explanatory purposes” fire (as we saw about the case of explanations in terms of character traits, this risk might be quite real indeed). The ethical naturalist wants to maintain that ethics resembles the empirical sciences: her claim would certainly not be vindicated by making ethics similar to phlogiston theory.

Even if we reject MMP, therefore, my emended Harman’s challenge survives as an empirical challenge to ethical naturalism. More accurately: the question to the naturalist “Do moral facts causally explain any observational facts we can use as observational criteria for

them?” has been an empirical question all along; the reason my emended Harman gives for answering it in the negative, though, is a reason of principle. If we reject that reason, and so believe that there are in principle no reasons why moral facts cannot explain any observational facts we can use as observational criteria for them, we still have to discuss whether, as a matter of fact, they do. If ethical naturalism is to turn out true, the naturalist needs to maintain that morality’s explanatory pretensions are vindicated: that there are some non-moral facts whose occurrence is causally explained by the occurrence of some moral facts, so that we can use the former as observational criteria for the latter (as we can use the occurrence of a vapor trail as observational criterion for the occurrence of a free proton, and as we cannot use the combustion of a flammable material as observational criterion of the liberation of phlogiston). The claim that moral facts are part of the causal-explanatory network of the world is controversial enough, even in the absence of any reason of principle to believe they cannot be part of that network. What non-moral facts are causally explained by the occurrence of a moral fact, so as to be available as observational criteria for it? What does ethics explain? These questions have a hold as such, even in the absence of any reason of principle to believe they cannot be answered as the naturalist would. At the same time, they are so crucial when it comes to assessing the plausibility of her position that the naturalist cannot avoid confronting them.

I have argued that the naturalist needs to defend the claim that morality’s explanatory pretensions are actually vindicated, if she wants to defend her position from my emended Harman’s challenge. What can the defense of that claim amount to? I take it that it is part of the naturalist’s (and, more generally, the philosopher’s) task to confront questions about what causally explains what in a way continuous with the methods and results of the natural and social scientist, as a part (maybe a distinctively general and abstract part) of the empirical inquiry about the world carried on by the sciences. Our naturalist has therefore to integrate her claims with the scientist’s ongoing empirical findings, and so to argue that the explanatory

pretensions our ethical discourse and practice have are, or can be, vindicated by our best empirical theories. She needs to develop a model of moral explanations and show that these explanations can be integrated into our best empirical explanatory picture of the world. One task she might take on is to argue that nothing in that picture, as described by current science, speaks against her proposed model of moral explanations. As we saw in our example of explanations appealing to moral features of people's character, this task can require the naturalist to question the appropriateness of some of the conclusions drawn from some empirical results, and to argue that those results, when properly assessed, do nothing to undermine her claims: in such a case, the standards the naturalist appeals to, in order to defend her assessment of the empirical results under examination, need to be internal to the empirical discipline in question (in our example, they have to be standards internal to psychological inquiry), and her case is of course strengthened if she can point to alternative assessments of those results available in the relevant empirical literature<sup>10</sup>. Another task the naturalist might take on is to argue that an empirical theory about the types of facts that morality (according to her) purports to explain has already begun, at least tentatively, to be developed, and that its explanatory claims do appeal to moral facts indeed. Of course, the naturalist can try to do all this only with reference to the methods and results of the current, ongoing empirical sciences: nothing guarantees that future developments will not reject all moral explanations as defective. But this is a pervasive feature of any empirical inquiry, one that the naturalist's (and, more generally, the philosopher's) conclusions share, when arrived at via methods continuous with those applied in the sciences, with the scientist's.

At the end of the first chapter I summarized my emended Harman's challenge to ethical naturalism in two questions. The first question was: can the naturalist make a good case for the claim that moral facts do cause some of the observational facts we can use as

---

<sup>10</sup> For a critique along these lines of the empirical work casting doubt on the existence of character traits (again, not with reference to the problem of moral explanations, but in defense of character-based versions of virtue ethics), see Sabini & Silver [2005].

observational criteria for them? The second question was: can she plausibly deny that our epistemological access to the moral domain is limited in the way MMP says it is? If the argument of the present chapter is correct, these two questions do not stand or fall together: while it is relatively easy for the naturalist to answer the second, she has more, and more complex, work to do in order to adequately deal with the first. What non-moral facts are causally explained by the occurrence of a moral fact, so as to be available as observational criteria for it? for the explanation of what in the world do we need ethical theory?

## **A model of moral explanations**

Moral principles cannot undergo observational testing in the same way scientific principles can, unless moral facts causally explain at least some observational facts we can use as observational criteria for them: do moral facts explain any such observational facts? This is the question my emended Harman presses on the ethical naturalist. MMP is not true: *contra* MMP, we can in principle “observe” the occurrence of a moral fact by relying on a causal claim about its effects, rather than on a moral principle. There are thus in principle no reasons why moral facts cannot causally explain any observational facts we can use as observational criteria for them. However, the question still looms large whether they, as a matter of fact, do: an ethical naturalist still needs an answer to this question, if she wants to answer my emended Harman’s challenge. In this chapter I’ll try to provide such an answer, by putting forward the model of moral explanations I find most promising.

Briefly described, the model of moral explanations I’ll try to defend is the following: facts about human flourishing, about the satisfaction of important human needs and the realization of human potentials, causally explain facts about our feelings, attitudes and actions (and sometimes even facts about the occurrence of some physical symptoms in our bodies). Before trying to defend this model of moral explanations, I’ll need to explain my reasons for focusing on this specific type of explanatory claims. I’ll then outline the kind of empirical inquiry within which facts about human flourishing, if they are to prove amenable to empirical investigation, find their place. Finally, I’ll argue that this kind of inquiry is actually carried on in our empirical sciences, and that its explanatory claims appealing to facts about

human flourishing are vindicated by our current empirical theories, and I'll deal with some objections that might be raised against my answer to my emended Harman's challenge.

### 1. Substantive constraints on a defense of moral explanations

At the end of the previous chapter I claimed that the ethical naturalist's defense of morality's explanatory pretensions is subject to the following constraint: it has to be carried on as a part of the general empirical inquiry about the world carried on by the empirical sciences. The naturalist's methods have to be continuous with those employed by the natural and social scientist, and her results must be capable of being integrated into their ongoing empirical findings. The naturalist's defense of moral explanations is, I believe, subject to a different type of constraint as well. The appropriateness of this further constraint is best understood if we consider the wide variety of moral judgments we make in our everyday life: to mention the few examples I made in the previous two chapters, we judge actions as right or wrong, institutions and social arrangements as just or unjust, people's character as virtuous or vicious (or in terms of specific virtues and vices). Correspondingly, we have a variety of types of moral principles (principles about what features of an action make it right / wrong; about what types of institutions or arrangements are just / unjust; about what features of a person's character constitute virtuousness / viciousness), about each of which my emended Harman can ask his question: can this type of moral principles undergo observational testing in the same way scientific principles can? do moral facts consisting in an action's being right / wrong (in an institution's being just / unjust; in a person's character's being virtuous / vicious) causally explain any observational facts we can use as observational criteria for

them? Distinct types of moral facts raise distinct questions about explanatory power: does our naturalist need to answer all of them? If not, then how are we supposed to select the one(s) she does need to answer? I believe the type(s) of moral facts an ethical naturalist has to focus on, in order to argue that they causally explain some observational facts we can use as observational criteria for them, is (are) determined by the answer to a difficult substantive moral question, about what might be called the theoretical structure of ethics: what is the place, within a substantive ethical theory, of the different types of moral notions we commonly employ in our everyday moral judgments? Are they all vindicated by a systematic substantive inquiry? And, among those that are vindicated, are there any that are more fundamental than others? For instance: is the right a function of the good? or is it rather a function of the virtuous? or is rightness an independent and fundamental moral notion<sup>1</sup>? Suppose an ethical naturalist's substantive conception of morals takes virtuousness to be prior to rightness: right actions are those a virtuous agent would choose to do. For such a naturalist, my emended Harman's question is most pressing if asked about facts on virtuousness: if facts of this type lacked explanatory power, so that principles about virtuousness could not be «justified ultimately by their role in explaining observations», then it seems the naturalist would (given her conception of morals) be forced to accept that our epistemological prospects are quite dim about principles of rightness as well. So, if a naturalist with this view of the theoretical structure of ethics sets out to answer my emended Harman's challenge and vindicate morality's explanatory pretensions, she will need to focus on the explanatory power of facts about virtuousness. Suppose now a second ethical naturalist believes instead that right actions are those maximizing good outcomes. Her conception of morals goes as follows: moral assessment, she claims, is basically assessment of actions in terms of the goodness / badness of their consequences; though in our everyday life we do assess people's character as

---

<sup>1</sup> This type of substantive questions lies at a more abstract level than those commonly cited as examples of substantive moral questions (such as "What types of actions are right?", "Is abortion wrong?", "What types of states of affairs are good?", "Is the experience of pleasure the only good thing?"): I borrow the useful phrase "questions about the theoretical structure of ethics" to refer to them from Brink [1989, p. 1].

well, she adds, this kind of assessment is quite peripheral to the most central concerns of morality and can be quite misleading with respect to them (any trait of character usually conceived of as a virtue can lead to very bad as well as very good actions, she remarks), and perhaps it should not even be classified as *moral* assessment proper, but as serving some different evaluative purpose. Given her substantive conception of morals, the latter naturalist cannot regard the former naturalist's defense of the explanatory power of facts about virtuousness as providing an adequate answer to my emended Harman's challenge: she cannot, even if we suppose she regards that defense as successful (even if, for instance, she takes the former naturalist to have convincingly rebutted the worries recent work in social psychology has raised about the existence of stable character traits). Even if the latter naturalist takes the former to have shown that facts about virtuousness can be integrated into our empirical explanatory picture of the world, she cannot think that is enough to answer my emended Harman's challenge: given her view of the theoretical structure of ethics, what an adequate defense of morality's explanatory pretensions would require is to show facts about goodness (as a property of outcomes), not about virtuousness, to have explanatory power.

As an ethical naturalist engaged in vindicating morality's explanatory pretensions, I too am subject to the constraint that my defense be adequate in light of the theoretical structure of ethics. My views about this matter will thus inevitably affect my argument: given my substantive conception of ethics, my emended Harman's question is most pressing if asked about a particular type of facts, so it is the explanatory power of this type of facts that I need to focus on and defend. I therefore need to give an (inevitably very sketchy) account of my substantive conception of morals: this account will provide my reasons for focusing on the defense of the explanatory power of facts about human flourishing. Other defenses of ethical naturalism from the worry about moral explanations might take different routes: for instance, they might set out to vindicate the explanatory power of facts about individual or social virtues and vices. I am not committed to viewing all these attempts as doomed to empirical

failure: each of them will have its specific empirical problems to face (and so will mine), but many of them may very well succeed in solving them. Naturalistic defenses of the explanatory power of moral facts, however, are subject not only to a constraint of empirical adequacy, but also to a different, substantive, type of constraint: it is with this multiple standard in mind that I wrote above that my model of moral explanations is the one I find most promising.

Before turning to my view of the theoretical structure of ethics, though, I need to deal with an objection that will probably be raised at this point against my way of setting things up<sup>2</sup>. I claimed that any attempt at arguing in defense of moral explanations, and so in defense of ethical naturalism, must rely on a substantive thesis (though of a fairly abstract kind). But, the objection goes, should we not be suspicious of this intertwining of meta-ethical and substantive views? Ethical naturalism is a meta-ethical thesis, and meta-ethics is substantively neutral: arguments for or against meta-ethical views should not commit one to, and should not depend on, any substantive view, and it should be possible to assess them without engaging in substantive moral inquiry. If this were not so, after all, solving meta-ethical disputes might require solving substantive disagreements, and there surely must be something wrong with the idea that one's assessment of a meta-ethical position may turn on one's substantive moral views. My first point in reply to this objection is that, if we consider the naturalist's task (i.e. showing that moral facts have explanatory power) bracketing for a moment claims about what the relationship between meta-ethics and substantive ethics should be, it appears that relying on some view about the theoretical structure of ethics is, for the naturalist, inevitable: our moral discourse and practice deal with a wide variety of moral facts, and I can think of no attempt at arguing for the explanatory power of moral facts which does not focus (not necessarily via an explicit choice) on a (few) specific type(s) of such facts to claim that *they* have explanatory power. Moreover, a choice is required for empirical reasons too: moral facts

---

<sup>2</sup> As should be clear from my discussion of Sturgeon's reply to Harman in the previous chapter, this objection is not one my emended Harman would raise, since it is a version of the idea that, in evaluating whether moral facts have explanatory power, we cannot be allowed to rely on moral assumptions. It has, however, often been raised at this point in discussions I have had of the topic of moral explanations, so I believe I need to deal with it.

about, say, people's character and the outcomes of actions are quite different types of facts, and they probably have different types of effects (if they have any), so a defense of the explanatory power of the former will look quite different from a defense of the explanatory power of the latter, and will have different empirical problems to face. But if a choice is to be made, the only way of justifying it is through a conception of morals that selects the chosen type(s) of facts as central in the theoretical structure of ethics: if a naturalist attempts at integrating moral facts into our empirical explanatory picture of the world by arguing that facts about virtuousness have explanatory power, she must at least implicitly be relying on a conception of ethics assigning a central place to facts about virtuousness. If we now look at the reason the objection gives for rejecting this account of what is required for a defense of moral explanations, I believe it is fair to remark that the objection relies on a view about the relationship between meta-ethics and substantive ethics that (though dominant among meta-ethicists for many decades of the last century) has more recently come largely under attack. It is of course both possible and useful to distinguish between meta-ethics and substantive ethics as domains of inquiry: what is at issue, though, is the claim that one's views about the two domains can have no bearing on one another. This view was mainly motivated by the idea that philosophical problems about ethics are confined to problems about the meaning of moral terms (which were taken to exhaust the field of meta-ethics) and by a belief in the existence of a sharp divide between questions of meaning and questions of substance, a divide enabling philosophers to discuss the former without engaging in discussions about the latter. Both the view of philosophy as mere analysis of language and the existence of a sharp divide between questions of meaning and questions of substance have by now been deeply criticized, and such criticism has as profound consequences for moral philosophy as it does for other fields of philosophical inquiry: a meta-ethicist who relies, in his arguments, on substantive moral claims should look no more suspect than an epistemologist who relies on substantive

scientific claims<sup>3</sup>. Since these consequences, however, are usually less explicitly dealt with in the moral case, I'll dwell a little longer on the problem, raised by the objection I am considering, of how the dependence of meta-ethical claims on substantive views affects the discussion of meta-ethical issues.

It is certain that, once we accept that any attempt at vindicating morality's explanatory pretensions has to rely on a substantive conception of ethics, the discussion about moral explanations becomes more complex: the field of play (so to speak) is broadened. The ethical naturalist has to claim that moral facts causally explain at least some observational facts we can use as observational criteria for them, in order to maintain that moral principles can undergo observational testing in the same way scientific principles do. The naturalist's opponent finds this claim highly questionable: for the explanation of what non-moral facts can we appeal to a moral fact? The naturalist tries to answer by putting forward her model of moral explanations and defending it, but it certainly follows from what I am claiming that, empirical adequacy aside, her answer can always in principle be questioned by questioning the conception of ethics on which it relies. This fact, however, will sound disappointing only if one expected meta-ethics to be able to provide an independent foundation for moral theory: I believe that what should look suspect are those expectations, because there is no such foundation (again, there is none for the empirical sciences either). Moreover, does this fact mean that, in the meta-ethical dispute between the naturalist and her opponent, each disputant's meta-ethical arguments will necessarily look valueless, because misdirected, to the other, if the two disagree on substantive matters, so that all they can do in such a case is to acknowledge the existence of a deep substantive disagreement preventing further meta-ethical discussion, and turn to substantive moral inquiry? I do not think it does. Assessment criteria for the meta-ethical contest between the naturalist and her opponent (as for so many other

---

<sup>3</sup> See Boyd [1988] for a detailed exposition of the thesis that epistemology has to rely on substantive scientific claims, for the view that an adequate defense of scientific realism has to rely on substantive scientific findings, and for a defense of ethical naturalism against various meta-ethical objections which depends (and which explicitly theorizes its dependence) on a substantive conception of morals.

philosophical contests) will have to be holistic: each of the two disputants will have to lay out her meta-ethical claim, along with her package of arguments and views (including substantive views) designed to support it, for a judgment of overall plausibility. The naturalist's opponent finds it hard to believe that moral facts belong in the causal-explanatory network of the world: if the naturalist is able, by relying on a conception of ethics chosen among the standard alternatives commonly discussed in substantive ethics, to mount an argument to the effect that moral facts do turn out to belong in the causal-explanatory network of the world, then her conclusion is certainly worth serious philosophical attention, whether her opponent shares that conception of ethics or not. Suppose the naturalist's defense of moral explanations were to rely on a very peculiar conception of ethics: in such a case, her argument would be of limited interest indeed, if it were not supplemented by a defense of that substantive conception. But if the naturalist's defense relies on (a plausible version of) one of the standard options available on the market of substantive ethics, then the naturalist's package is a legitimate competitor in the contest, whatever her opponent's substantive views.

I do not think, therefore, that the doctrine of the substantive neutrality of meta-ethics is a good reason to deny the appropriateness of a substantive constraint on a defense of moral explanations: for the reasons I gave above, the naturalist's defense of morality's explanatory pretensions has to be adequate in light of the theoretical structure of ethics, so it has to rely, *pace* that doctrine, on some view about what that theoretical structure is. What, then, is my substantive conception of ethics? And what types of facts does it assign a central place to? I can summarize my conception of ethics in the following way<sup>4</sup>. Human beings have important needs that have to be satisfied in order for them to be fully healthy people, thrive, or (as philosophers sometimes put it) flourish. Some of these needs are physical; some of them are psychological; still others are social. Some of these needs are for the satisfaction of some types of desires; others are for the exercise and development of some types of capacities.

---

<sup>4</sup> For a conception of ethics similar to mine, see Brink [1984] and especially Boyd [1988].

Some needs are nearly universal among human beings; others may vary from person to person; in any case, the ways in which many of them are satisfied are likely to be different in different cultures and ways of living. What should we include in the list of these important human needs? For example, what types of capacities and desires do beings such as humans need to exercise and satisfy, if they are to flourish? If my conception of ethics is correct, these are among the most important substantive moral questions, they are some of the questions about which it would be most important to get the answers right in order to get things right in substantive ethics; moreover, if the meta-ethical position I am going to defend is correct, these are (complex and presumably difficult) empirical questions. What I can do, in order to give a clearer idea of what I mean by “human flourishing” and “human needs”, and so clarify my substantive conception of ethics, is to mention some of the answers that have been given to these questions: given the social nature of human beings, the need to form inter-personal bonds, and in particular the need for friendship and love, have long widely been regarded as fundamental human needs; the need for artistic expression and appreciation is an example of a kind of needs which have been recognized as such more recently; a need to exercise control over one’s life is commonly recognized in western societies, but it might be claimed that there really is no such need and that belief in its existence is a product of our individualistic culture. Ethics is about the satisfaction of human needs and the realization of human potentials for thriving. Human goods are the constituents of a flourishing life, things which satisfy important human needs; moral assessment of actions, character traits, institutions, social arrangements, etc. is concerned, from a point of view equally engaged on behalf of all those potentially affected, with the extent to which these objects of moral evaluation contribute to the realization of those human goods and so to human flourishing. Of course, I cannot try to defend this substantive conception of morals here; however, in compliance with what I maintained above about the interplay between substantive and meta-ethical views in the debate between the ethical naturalist and her opponent, my substantive conception of ethics is

- I believe - a plausible (though, of course, hardly uncontroversial) version of one of the standard positions available in substantive ethics, namely consequentialism.

The reason why, I argued, an ethical naturalist needs to rely, in her defense of morality's explanatory pretensions, on a substantive view (though of a fairly abstract kind) is that she needs to select the type(s) of moral facts on which her defense should focus: she needs, as I put it, a view of the theoretical structure of ethics, a view about what types of moral notions are vindicated by a systematic substantive inquiry and about which (if any) of them are more fundamental than others. As any consequentialist theory, my substantive conception of morals selects a type of evaluative (as opposed to normative) facts as central in the theoretical structure of ethics: facts about human flourishing and needs, about what people's needs are and about what the constituents of a flourishing life are<sup>5</sup>. Is a need for friendship a fundamental human need? is friendship one of the human goods, one of the constituents of a good or flourishing life for nearly every human being? Is the need for artistic expression one of John's basic needs, so that a life as a painter would be more fulfilling for him, allow him to flourish more than a life as an accountant? On my conception of morals, these evaluative questions are central for substantive ethics: if my emended Harman is right about them, if competing answers to these questions cannot undergo observational testing in the same way scientific principles can, then moral facts can hardly be known in pretty much the same way in which scientific facts are known. It therefore is the explanatory power of facts about human flourishing and needs that I have to be ready to defend, if I want to answer my emended Harman's challenge.

---

<sup>5</sup> I am here using the terms "normative" and "evaluative", in accordance with one common usage in moral philosophy, to mark a distinction that it is standard to make within the moral field (with rightness, duty and obligation, on the one hand, and goodness, on the other, as typical exemplifications). According to a broader usage of the term "normative", all moral questions (whether about what is good or bad, or about what is right or wrong) count as normative questions: this is, for example, the sense in which it is standard to distinguish between meta-ethics and normative ethics (i.e. what I refer to in the text as substantive ethics). The term "evaluative" is used with a broader sense too: when some philosophers claim that evaluative terms are not descriptive, for example, they mean their thesis to apply to "ought" and "right" as well as to "good".

## 2. Moral explanations

In the second chapter I claimed that, contrary to what MMP entails, important parts of our moral discourse and practice obey explanatory constraints: morality does have explanatory pretensions. My examples there were taken from our talk about individual and social virtues and vices: we can in principle “observe” that a man is courageous or that a society is just by observing some non-moral facts about, say, that man’s actions or that society’s relative stability, non-moral facts we can take as evidence for the moral fact in question because we believe them to be caused by it. So our attributions of courage or justice are shaped by explanatory purposes with regard to the courageous man’s actions or the just society’s stability. Of course, these explanatory pretensions might very well not be vindicated; if, however, one wanted to argue that facts about courage or justice do turn out to belong in the causal-explanatory network of the world, one would likely start from these pretensions, and investigate whether, for example, being courageous does sometimes actually explain doing the actions we commonly take to be explained by it. What, now, about facts on human flourishing and needs, which are, on my conception of morals, the central facts in the theoretical structure of ethics? what are the explanatory pretensions of our discourse about human flourishing? We can get an idea of such pretensions if we look at some of our commonsensical explanations in which we appeal to facts about the satisfaction of human needs and the realization of human potentials for thriving. Even if probably not exactly in these terms, these explanations are not uncommon: here are a couple of examples.

Nick and Linda are talking about their son John. John is a high school student: he is a model student, his grades are the best in his school and his intelligence and diligence are admired by his teachers and classmates. He spends his time at home doing his homework and reading, because he genuinely enjoys learning new things and going into what he is taught at

school. Nick and Linda are of course glad to have a child who takes so strong an interest in his education, and are proud of John's performances at school. They are, though, a little anxious for him. He is often nervous and sometimes stressed, and every now and then he goes through periods during which he loses any interest for anything: he just sits passively in the living room, staring blankly at the images on TV. These periods are always quite short, two or three days at most, after which he is again the usual John everybody knows, always ready to tell you about the recent successes of genetic studies (of which he is especially fond) he read about in the newspaper. John has just got back to school after three days spent hanging around in the house, between the living room and his bedroom, and Nick and Linda are talking about what may lie behind John's mood and behavior. Nick has a definite idea about that: «Of course he's so nervous and stressed: he'd need to do something else, besides sitting in his room, studying and reading. He would need to do some sport, for example; and above all he would need to go out and see some people every now and then».

Peter and Richard have known each other since they were eight years old. Peter has always enjoyed Richard's company: Richard is very sociable, and chatting with him is always funny. More than anything else, Peter likes the fact that Richard is one of those people who never let misfortunes discourage them: he never loses his positive attitude, and his enthusiasm is contagious. Richard is a promising young musician, who has already won several important prizes for young pianists. Even though he has never been fond of classical music, Peter has often thought, while looking at Richard talking about music and its expressive powers, that he would make anybody feel like listening to his beloved Beethoven. Peter therefore is quite surprised when Richard tells him he will give up music to go to law school, in order to comply with his father's wishes. Peter goes to college as well, but the two keep on meeting, when they are both back in town for their holidays. Peter, however, finds that Richard's company is no longer so pleasant as it used to be: he has turned bitter, and is now all too ready to complain and blame others when things go wrong (as unfortunately they often do at

law school: thus far he has been able to pass very few exams). While leaving their usual café at the end of their latest meeting, Peter could not avoid thinking «He was such a brilliant guy, he's been turning more and more bitter since he decided to study law and become a respected judge. But he's not cut out for that: he's got an artistic spirit».

In looking for an explanation of John's mood and behavior, Nick turns to considerations about his son's needs. His opinion is that John's mood and behavior are explained by the fact that his son's way of living does not satisfy all his needs: given the kind of life John is living, some of his fundamental needs are not satisfied, and this fact reveals itself in John's being at times nervous, stressed and without any interest in anything. Nick has also some positive ideas about what it is, in John's life, that prevents his fundamental needs from being all satisfied, and about what it is that John is neglecting. While John is certainly thirsty for knowledge and enjoys learning new things, one cannot forget, so Nick thinks, that he is a young boy, who needs to have friends and maybe do some exercise, just like anybody else in his age: devoting his time only to study prevents these other needs of his to be satisfied, and in the long run it proves detrimental to John's desire for knowledge too, since he refuses to do anything during those days he spends sitting before TV (that is how he missed the meeting with that famous geneticist his school organized last month).

Peter is struck by the changes in Richard's character: he is no longer the sociable, pleasant person, always ready to face difficulties with optimism and confidence, he used to be. Peter attributes these changes to Richard's decision to give up music for a career as a judge. That is not, however, Peter's full explanation: after all, Richard might have been such that that choice would have produced no such effects (he might have been like Mark, who used to play the triangle: he also gave up music to go to law school, but he is still the same pleasant guy he used to be at high school). The point is, so Peter thinks, that Richard chose a career he is not "cut out" for. Peter has known Richard for years, and he knows how much Richard loves music, how good he is at playing the piano, and how fulfilling playing the

piano and studying music are for him: a career as a pianist is what Richard is “cut out” for, that is the career which would “fit” him perfectly, which would be fulfilling for him as no other and would allow him to fully realize his potentials for thriving. The fact that Richard chose a career that does not allow him to flourish is what, according to Peter, explains why Richard, once a pleasant and optimistic guy, turned bitter and complaining.

As is shown by these examples, our thinking about human flourishing is shaped by something like the following assumption: other things equal, living a flourishing life leads one to be active, or serene, or satisfied with one’s life, while living a life that prevents one from flourishing leads one to be listless, or stressed, or unsatisfied. More generally, the explanatory pretensions of our discourse about human flourishing seem to me to parallel those of our discourse about (physical) health: the type of commonsensical explanations in which we appeal to facts about human flourishing parallels the type of explanations in which we appeal to facts about being more or less healthy. Our discourse about health, about whether a person *p* is more or less healthy, and about what fosters and what undermines *p*’s health, is shaped by the assumption that being more or less healthy has a characteristic pattern of symptomatic effects. In an analogous way, our thinking about *p*’s flourishing, about whether and to what extent *p*’s needs are / are not satisfied, or *p*’s potentials are / are not realized, is shaped by the assumption that these evaluative facts have a characteristic pattern of effects, effects which can be seen as symptoms of the fact that *p* does / does not thrive: for instance, effects on some of *p*’s feelings, attitudes, or actions, such as that *p* is a serene / stressed person, or that *p* is satisfied / unsatisfied with her life, and the like<sup>6</sup>.

So here is the picture that emerges if we take the explanatory pretensions of our discourse about human flourishing into account. For a person *p*, there are various things *p* has

---

<sup>6</sup> It is interesting to note in this context that, according to the definition of “health” of the World Health Organization, physical health is actually a part of flourishing (or of something very close to flourishing), to which the label “health” is reserved: «Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity» (Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 1946).

to get, do, etc. if she is to thrive: these are the things *p* has a need for, the constituents of a flourishing life for *p*. Given my substantive conception of morals, the evaluative questions “What are *p*’s needs? what does *p* have a need for?” are central substantive moral questions to ask. If *p* does not get, do, etc. one of these things, the fact that she does not live a flourishing life affects her feelings, attitudes, actions, etc.: *p* is stressed, or listless, or unsatisfied with her life, and the like. If, on the other hand, *p* does get, do, etc. one of these things, then her needs are, *pro tanto*, satisfied and she lives, *pro tanto*, a flourishing life: this fact reveals itself in *p*’s being serene, or active, or satisfied with her life, and the like<sup>7</sup>. The explanatory pretensions this picture attributes to our discourse about human flourishing would, if vindicated, provide what is needed in order to answer my emended Harman’s challenge: non-moral facts whose occurrence is causally explained by the occurrence of facts about human flourishing, and which we could use as observational criteria for such facts. Suppose, for example, that we wanted to test the claim that *p* needs to have friends, that friendship is one of the constituents of a flourishing life for *p*: we should check whether *p* thrives even without friends or instead having friends contributes to *p*’s flourishing, making *p*’s life fulfilling in a way that would be lost in the absence of friends. If the explanatory pretensions of our discourse about human flourishing were vindicated, we could do that by observing how *p* reacts to the presence and absence of friendship in her life: for instance, by observing whether her life undergoes no significant changes after she loses her friends, or she turns more nervous and stressed instead. This is what Nick does in the first of my two stories: John certainly likes learning new things and getting excellent grades at school, but he would also need, so Nick believes, to go out with some friends every now and then, which is instead something he never does. When John begins to show signs of nervousness, stress and listlessness, Nick takes them to confirm his

---

<sup>7</sup> I am simplifying things a little, since there are of course differences in how satisfying or failing to satisfy different needs affects a person’s thriving, and so in the degree of relative importance of different needs: Nick, for example, seems to think that John would certainly need to do some exercise, instead of spending all his time sitting in his bedroom, but that it is much more detrimental to John’s flourishing that he has no friends with whom to go out. These differences, however, do not affect my argument, so I’ll bracket them.

belief, and when he talks with Linda about the possible causes of John's mood and behavior, he appeals to that belief for an explanation: «Of course he's so nervous and stressed: he would need to go out and see some people every now and then». Nick "observes" that John is not living a flourishing life by observing that he is at times nervous, stressed and listless, and uses that "observation" to test his belief about John's need for friendship.

The notion of flourishing this framework employs has, in itself, nothing that limits its applicability to human beings only: rather, it seems to be possible to apply it (again, like that of health) to all living beings<sup>8</sup>. For instance, it seems it is in the same sense of "need" that we can say, about both humans and certain non-human animals, that they need parental care in the early periods of their life: so much so that one possible option in substantive ethics is to broaden my view of the theoretical structure of morals and take facts about at least some non-human animals' flourishing, besides facts about human flourishing, as central. Of course, the general notion of flourishing, along with its explanatory pretensions, will be more or less wide, depending on the kind of living beings in question, and the notion of flourishing I am interested in is a wide notion, appropriate to such complex animals as human beings are. As P. Railton once put it:

«"Flourishing" [...], even if understood as mere reproductive fitness, is not a narrow notion. In order for beings such as humans to be reproductively successful, they must as phenotypes have lives that are psychologically sustainable, internally motivating, and effectively social; lives, moreover, that normally would engage in a wide range of their peculiarly human capacities.»<sup>9</sup>

Still, the general framework (a living thing needs a number of things if it is to flourish, and the fact that its needs are / are not satisfied reveals itself in a pattern of characteristic effects)

---

<sup>8</sup> Etymology is helpful here, since the verb "to flourish" was originally used to describe trees which, growing healthy, put out flowers.

<sup>9</sup> Railton [1986a, p. 179].

is the same: the claim that facts about human flourishing turn out to belong in the causal-explanatory network of the world, and thus prove amenable to empirical investigation, is just one particular instance of the more general claim that facts about flourishing can figure in the empirical study of living beings. It seems to me this makes the prospects of a defender of the former claim look not entirely dim, since life sciences do tell us about living things' needs: such claims as that plants of type T need amount A of water every day or that animals of type T' need food F are empirically in order (and it is worth noticing that even these crude examples do not amount merely to claims about conditions for survival: that animals T' need food F could be true even if a T' could survive if fed only on a different type of food)<sup>10</sup>.

Be that as it may, I'll argue that the explanatory pretensions of our discourse about human flourishing are vindicated, in that current empirical science does appeal to facts about human flourishing in its explanations. Specifically, I'll review some recent developments in psychological research in order to show that psychological explanations vindicate the explanatory pretensions of our discourse about human flourishing and that psychologists make use of explanatory claims about facts on human needs to test hypotheses about what those needs are. In so doing, I'll deal with hypotheses and claims about needs common to almost all human beings: as should be clear from my claims and examples above, however, and as it is worth noticing explicitly, no assumption is made in my account to the effect that there are no (or very few) differences among different people with regard to their respective needs, hence with regard to the constituents of a flourishing life for each of them. On the contrary, since human beings have much in common (as members of the same species) and much that differentiates them as well (due to their different developmental histories, circumstances, etc.), what it is reasonable to expect is that, while a whole class of needs will

---

<sup>10</sup> Of course research into the physiological, chemical, etc. facts which underlie living beings' needs is an important part of our scientific understanding of living beings, but it hardly impugns the empirical status of claims such as those I mention in the text: rather, it bolsters those claims, in the same way in which research into the physical facts underlying chemical phenomena articulates and improves our understanding of those phenomena. I discuss related issues, with regard to human flourishing, in the next section.

be common to almost all of them, others will be different for different people: my dealing, in my review of the psychological literature, with the former is in no way meant to deny the existence (or diminish the importance) of the latter<sup>11</sup>. As any form of moral realism, ethical naturalism is usually taken to deny relativism, but I believe that an account of morality recognizing different components of a good life for different people can be pluralist without being in any interesting sense relativist. The parallel with health is helpful here: human beings have much in common, and much that differentiates them too, so what fosters and what undermines their physical health is partly common and partly different for different people. I do not believe, however, that a hypothetical relativist about health would convince many of us by pointing to the fact that different things are sometimes healthy or unhealthy for different people.

What does psychological theory tell us about the commonsensical explanations in which we appeal to facts about human flourishing to explain various facts about people's feelings, attitudes, actions, etc.? Does it vindicate this type of explanatory claims? Or does it show instead that they must be rejected, in that they appeal to facts that, despite what we may believe in our everyday life, contribute no explanatory insights, and have no role to play in our best empirical explanatory picture of the world? Various trends in the psychological research of the last few decades indicate that facts about human flourishing do have explanatory power<sup>12</sup>. In order to make my discussion definite, I'll focus on two types of possible *explananda*: intrinsically motivated behaviors and daily well-being.

By "intrinsically motivated behaviors" psychologists mean those behaviors humans and other animals engage in spontaneously, for the interest and enjoyment of the behaviors

---

<sup>11</sup> Of course the same considerations apply to each individual as well: while some of a person's needs will remain nearly constant over her life span, others will likely vary depending on various factors such as age, circumstances, and so on. I think it is also possible that, for a given individual at a given time of her life, two different kinds of life can be equally good, both contributing to her flourishing in the same ways.

<sup>12</sup> Of course, I cannot attempt a comprehensive survey of the psychological literature: my purpose will be to point to some trends and to their relevance for my attempt to answer my emended Harman's challenge. My discussion will rely mainly on the work of R.M. Ryan, E.L. Deci, C.D. Ryff, K.M. Sheldon and H.T. Reis: see, in particular, Ryan [1995], Ryan & Deci [2000; 2001], Ryff [1989; 1995], Ryff & Keyes [1995], Sheldon et al. [1996], Reis et al. [2000].

themselves, and which are not a function of either visceral drives or external reinforcements. Children, for instance, are often active and inquisitive, even in the absence of any specific reward; the curious exploration of new objects, and more generally the tendency to explore, learn, seek out challenges, extend and exercise one's capacities are examples of the phenomenon of intrinsic motivation. Though certainly a type of motivation that moves human beings to act, intrinsic motivation does not, however, operate uniformly: different contexts, conditions, situations, etc. are known to facilitate or undermine it. An interesting question therefore is: what types of conditions facilitate and, respectively, undermine intrinsic motivation in human beings? what accounts for variability in intrinsic motivation? Ryan & Deci [2000]'s review of various studies and results on this point posits «a parsimonious list of three basic psychological needs as a means of organizing and interpreting a wide array of empirical results, results that seemed not to be readily and satisfactorily interpretable without the concept of needs» and concludes that «social environments can facilitate or forestall intrinsic motivation by supporting versus thwarting people's innate psychological needs»; specifically, «the findings have led to the postulate of three innate psychological needs – competence, autonomy, and relatedness – which when satisfied yield enhanced self-motivation [...] and when thwarted lead to diminished motivation»<sup>13</sup>. In a similar spirit, Ryan [1995] introduces a study of intrinsic motivation and other motivational processes by claiming that «these processes are [...] shown to be facilitated by conditions that fulfill psychological needs for autonomy, competence, and relatedness, and forestalled within contexts that frustrate these needs. Interactions between psychological needs and contextual supports account, in part, for the domain and situational specificity of motivation», so that «how active, interested, and authentic versus alienated one is in a given setting is [...] not just a matter of

---

<sup>13</sup> Ryan & Deci [2000, pp. 74, 71 and 68]: “self-motivation” covers both intrinsic motivation and other types of motivation Ryan and Deci examine.

preexisting individual differences, but is also a function of immediate affordances related to basic psychological needs»<sup>14</sup>.

Daily well-being is a matter of feelings of happiness, satisfaction and vitality, as they are experienced at a day-to-day level. As each of us knows, daily well-being varies from day to day and, within the same day, from setting to setting: in order to research into what may account for such fluctuations, a series of diary studies distinguished three dimensions of daily well-being. Participants in the studies were asked to complete: a mood checklist, comprising nine mood adjectives such as “joyful”, “happy”, “pleased”, “depressed”, “worried / anxious”, “angry / hostile”; a vitality scale meant to measure the degree to which participants felt vigorous and alert, both physically and mentally; a physical symptom checklist, comprising nine items such as runny nose, difficulty in breathing, soreness. Variations in daily well-being along these three dimensions were what the studies tried to account for by looking at the types of activities participants reported they had engaged in. Again, whether and to what extent the activities participants engaged in satisfied their needs proved to be explanatorily relevant: «we propose that “good days” are those in which fundamental psychological needs are met»<sup>15</sup>; «much of this within-person fluctuation may be understood by examining how the ever-changing environment of daily life affords fulfillment of basic psychological needs»; «in this study, we investigate variables that we expect to have direct effects on within-person changes in well-being. Specifically, we investigate the hypothesis [...] that personal well-being is a direct function of the satisfaction of basic psychological needs»<sup>16</sup>.

Do these empirical results vindicate the explanatory pretensions of our discourse about human flourishing? do they show, for instance, that there is something to Nick’s explanation of John’s mood and behavior? It might be doubted that they do because, on my substantive conception of morals, human needs are for things human beings have to get, do, etc., if they

---

<sup>14</sup> Ryan [1995, pp. 397 and 417].

<sup>15</sup> Sheldon et al. [1996, p. 1270].

<sup>16</sup> Reis et al. [2000, pp. 419 and 420].

are, as I put it, to be fully healthy people, thrive, or flourish: it is facts about needs so conceived, as linked to human flourishing, that have to turn out to belong in the causal-explanatory network of the world, if I want to answer my emended Harman's challenge. In order to appreciate the relevance of these empirical studies for my attempt to develop such an answer, it therefore is necessary to look at the definitions these studies give of "need":

«In this usage, the term *need* does not refer to *any* chronic want or desire (some of which are clearly irrelevant or even harmful to health and development) but only to inputs that are important for psychological health and integrity.»<sup>17</sup>

«In this definition needs refer to the nutriments or conditions that are essential to an entity's growth and integrity. A plant needs sunlight and water to grow. Similarly a person, as a biological entity, needs food, water, and (in Rochester, NY) shelter to thrive.»<sup>18</sup>

«By our definition, a basic need, whether it be a physiological need [...] or a psychological need, is an energizing state that, if satisfied, conduces toward health and well-being but, if not satisfied, contributes to pathology and ill-being. We have thus proposed that the basic needs for competence, autonomy, and relatedness must be satisfied across the life span for an individual to experience an ongoing sense of integrity and well-being or "eudaimonia"»<sup>19</sup>

As these quotations make clear, the sense of "need" in which these studies attribute explanatory power to facts about human needs is the same as that in which, on my substantive conception of morals, facts about human needs are central in the theoretical structure of

---

<sup>17</sup> Sheldon et al. [1996, p. 1277].

<sup>18</sup> Ryan [1995, p. 410].

<sup>19</sup> Ryan & Deci [2000, pp. 74-75].

ethics. More generally, it is quite common, in this type of psychological studies, that the authors explicitly acknowledge the contiguity of the subject of their researches with that of philosophical debates about the good life, and stress the relevance of their work for all those (be they moralists, parents, policy makers, and so on) concerned with people's quality of life<sup>20</sup>. It therefore is legitimate to conclude that these studies vindicate the explanatory pretensions of our discourse about human flourishing: the picture I described above (a person *p* needs various things if she is to thrive, and the fact that her needs are / are not satisfied reveals itself in a pattern of characteristic effects) seems to be vindicated by current empirical science.

If this is so, it should be possible to expose competing hypotheses about what human beings have a need for to observational testing: if facts about human flourishing and needs explain such things as variations in intrinsic motivation or in feelings of happiness and satisfaction, then competing hypotheses about what human beings' needs are should lead to competing predictions about what fosters or, alternatively, undermines intrinsic motivation, feelings of satisfaction, etc. It is interesting to notice that this procedure for the observational testing of claims about human needs is, in some cases, explicitly described in the psychological literature. Here are two quotations, which follow, respectively, the second and the first definition of "need" quoted above:

«This usage of the term need can be applied across the life sciences, since all living things have empirically identifiable needs whose essentialness can be examined by systematically varying nutrients against the criteria of health and integrity.»<sup>21</sup>

«we suggest that the importance of a postulated need can be established by showing that it uniquely predicts criterion measures of well-being, health, or development [...]. Given

---

<sup>20</sup> On this point see in particular Ryff [1989] and Ryan & Deci [2000; 2001]. The reference to Aristotle in the definition of "need" I quoted from Ryan & Deci [2000] is hardly an exception in this respect.

<sup>21</sup> Ryan [1995, p. 410].

this definition of needs and this strategy for identifying them, the question of how many needs there are becomes an empirical one.»<sup>22</sup>

Moreover, as is evident from some of my quotations, part of the purpose of some of the studies I am reviewing is to test a specific hypothesis about human needs: namely, that human beings have a need for autonomy, competence and relatedness. And this hypothesis is tested precisely by relying on explanatory claims about facts on human needs, and reviewing existing studies, or conducting new ones, in order to check whether, say, teachers and parents who are autonomy supportive really have students and children who are more intrinsically motivated, or whether days in which people engage in activities allowing them to feel competent and effective actually are days in which they feel happier, and so on. That is: the fact that strong enough correlations are observed to obtain between autonomy, competence and relatedness, on the one hand, and intrinsic motivation or feelings of happiness, on the other, is offered as evidence for the hypothesis that human beings need autonomy, competence and relatedness, because the hypothesis would help explain those correlations, given the explanatory relevance of facts about human needs with regard to intrinsic motivation or feelings of happiness. A hypothesis about what human beings need is thus «justified ultimately by [its] role in explaining observations»<sup>23</sup>.

My substantive conception of morals selects facts about human flourishing and needs as central in the theoretical structure of ethics: if I want to answer my emended Harman's challenge, it is facts of this type that have to turn out to belong in the causal-explanatory network of the world. In discussing about what people's needs are and about whether and to

---

<sup>22</sup> Sheldon et al. [1996, p. 1277].

<sup>23</sup> For a particularly clear description of this procedure for the observational testing of claims about human needs, and an application of it to the claim that human beings need to form and maintain interpersonal relationships, see Baumeister & Leary [1995]. Of course, I need not here commit myself to this view about human needs, nor do I need to accept the claim, often made in part of the literature I am reviewing, that the needs for autonomy, competence and relatedness are innate and universal: what is relevant for my purposes is just that psychologists' empirical work vindicates the explanatory power of facts about human needs and relies on it in exposing hypotheses about those needs to observational testing.

what extent they are / are not satisfied, we do appeal to explanatory claims about facts on human needs: our discourse about human flourishing has explanatory pretensions. Are these explanatory pretensions vindicated? I have answered this question in the affirmative, by claiming that empirical science does appeal, in its explanations, to facts about human flourishing. The best way to show that it does is to appeal to the relevant empirical literature: I have thus pointed to some recent trends in psychological research, and quoted extensively from the psychological literature in order to show as clearly as possible that psychologists mean by “human needs” the same as I do, that they attribute explanatory power to facts about human needs, and that they use their explanatory claims about such facts to expose hypotheses about human needs to observational testing. I must now turn to some objections that might be raised against my claims.

### 3. Some objections

In order to answer my emended Harman’s challenge, I sketched my substantive conception of morals, and argued that the type of facts that, on that conception, are central in the theoretical structure of ethics do turn out to belong in the causal-explanatory picture of the world, as that picture is described by current empirical science. My conception of ethics, I claimed, selects facts about the constituents of a flourishing life as the central evaluative facts in the theoretical structure of ethics: so it is about facts of this type that I set out to argue that they have explanatory power. But, it might be objected, facts about human flourishing and needs are not evaluative facts at all. The objection is not, of course, that they are normative facts: what is objected is rather that they are neither evaluative nor normative, that they are a

type of plain non-moral facts. Establishing the explanatory power of this type of facts, so the objection claims, amounts to no defense of the explanatory power of moral facts.

I confess I find this objection a little surprising. The question whether facts about human flourishing and needs are or are not evaluative cannot be answered, in the present dialectical situation, with any systematic argument, since we do not have any general neutral account of what counts as an evaluative fact, or of what an evaluative term is (nor, for that matter, of what a normative fact or term is), to which we can turn in order to adjudicate contested cases. Meta-ethical theorizing, and more generally philosophical theorizing about the evaluative and the normative domains<sup>24</sup>, do not begin with a general account of the distinction between the moral and the non-moral, or between the evaluative and the non-evaluative, but with a rough and intuitive grasp of those distinctions. Disputes about whether moral facts are natural facts, for example, or about whether evaluative epistemology looks like the epistemology of the empirical sciences, begin with two rough lists of terms, and ask about the relations between them: are the terms on the “moral / evaluative” list referential as the terms on the “non-moral / non-evaluative” list are? do the terms on the first list refer to properties of the same kind as those referred to by the terms on the second, or do they refer to properties of an entirely different, *sui generis* kind? If the question is raised, about a particular term, whether it is evaluative or not, it can only be answered by asking which of the two lists we would intuitively say that term is on. The reason why I find the objection surprising is that “being a human need”, “being a constituent of a flourishing life” and the like seem to me clearly to belong in the “evaluative” list: questions about what counts as a human need or about what the constituents of a thriving life are, or claims to the effect that, say, exercising control over one’s own life is not a human need or that having friends is a constituent of a flourishing life seem to me to be clearly evaluative questions and claims. Suppose that Linda,

---

<sup>24</sup> Morality is not, of course, the only domain in which we use evaluative or normative notions (other examples are aesthetics, epistemology, the theory of practical rationality), and meta-ethical theories often study morality as one sub-field within the wider evaluative / normative field.

talking with Nick about their son John, agrees with him and says «You're right: he spends all his time studying in his room, while he would need to go out with some friends every now and then»: it seems to me that her claim would be an evaluative claim. Or suppose that someone were to maintain that no human being can thrive unless he or she is granted a considerable amount of autonomy in his or her choices with regard to things like his or her job or partner, and that someone else were to claim instead that that is true for men, but that it is no prejudice to a woman's potentials for thriving and living a fulfilling life that she can exercise no control over her own life: wouldn't the two be engaged in an evaluative dispute? If we set out to develop a philosophical theory about the evaluative, wouldn't we regard "being a human need" or "being a constituent of a flourishing life" as part of the object of our theory, as part of what we want a theory of?

Even though I cannot appeal to any general neutral account of the evaluative in support of my answer to these questions, I can appeal to a test I always find useful<sup>25</sup>. Suppose I were trying to construct a derivation of an evaluative conclusion from entirely non-evaluative premises: if I were to use a premise to the effect that this is a human need, or that that is a constituent of a flourishing life, would I be accused of smuggling a covertly evaluative premise into what was supposed to be a set of non-evaluative premises? If the answer is positive then, my test concludes, the terms in question are evaluative. And I think I would no doubt be accused of smuggling values into my premises. Here is, for example, what R.M. Hare writes at the end of the chapter in Hare [1952] in which he gives his famous restatement of G.E. Moore's open question argument:

«Let us suppose that someone claims that he can deduce a moral or other evaluative judgement from a set of purely factual or descriptive premisses, relying on some definition to the effect that V (a value-word) means the same as C (a conjunction of

---

<sup>25</sup> I claim no originality at all for this test: anyone interested in meta-ethics will already have employed it several times in his or her reflections.

descriptive predicates). We [...] have to ask him to be sure that C contains no expression that is covertly evaluative (for example “natural” or “normal” or “satisfying” or “fundamental human needs”). Nearly all so-called “naturalistic definitions” will break down under this test – for to be genuinely naturalistic a definition must contain no expression for whose applicability there is not a definite criterion which does not involve the making of a value-judgement.»<sup>26</sup>

Of course, one of the possible results of a meta-ethical theory, or of a philosophical theory about the evaluative, is that we end up with a general account of what a moral or an evaluative fact or term are: we can then use that theory-informed account to answer questions about whether a particular fact or term is moral or not. And it is certainly possible that, on the basis of our theory-informed account, we are led to expunge, from our initial list of moral terms, some terms we had intuitively judged to belong in it: one of the possible effects of developing a theory (in philosophy as in any other field) is that we are led to revise some of our intuitive judgments about what we are developing a theory of. I believe this is usually what those sympathetic to the objection I am considering actually do in raising that objection: they rely on a theory-informed account of the moral, and on those grounds they object that facts about human flourishing or claims about human needs do not count as moral facts or claims. In order to assess the objection, so taken, I'll focus on what I believe is the theory-informed account those sympathetic to it most commonly rely on: namely, the idea that what is distinctive of the moral is that it is intrinsically action-guiding<sup>27</sup>. This general claim can be interpreted in very different ways: to mention two possibilities, according to some

---

<sup>26</sup> Hare [1952, p. 92].

<sup>27</sup> As a matter of fact, this is the account relied on by all those who have raised this objection to me in discussions I have had of the topic of moral explanations. It is an account championed by a long list of philosophers in analytical meta-ethics, and which can claim as noble ancestors as D. Hume and I. Kant. I focus on the label “intrinsically action-guiding” because it is commonly used in meta-ethics, but many of those who accept this general account believe that actions are not the only possible objects of evaluation: intentions, feelings, emotions, attitudes and (outside morality) arguments and beliefs are all possible objects of evaluation (interesting complications are possible: Gibbard [1990], for example, holds that in judging actions as morally right or wrong we guide feelings of guilt and anger, rather than actions directly).

philosophers who accept it, the moral is intrinsically action-guiding in that one cannot sincerely assent to a moral claim without being *ipso facto* motivated to act in a certain way; according to others, the moral is intrinsically action-guiding in that moral facts are facts about normative reasons for action agents have regardless of their aims, desires, wants, etc. However different the various interpretations may be, the common idea is that moral claims or moral facts have a distinctive and special element of prescriptivity or (as J.L. Mackie would put it) “to-be-doneness” built into them<sup>28</sup>. The objection to my proposed model of moral explanations then goes as follows: facts about human flourishing are not intrinsically action-guiding, hence they are not moral facts at all, therefore establishing the explanatory power of this type of facts amounts to no defense of the explanatory power of moral facts<sup>29</sup>.

Before answering this objection, I believe it is useful to remind the present dialectical situation. Ethical naturalism claims that moral facts are natural facts, and as such can be known in pretty much the same way in which scientific facts are known; my emended Harman raises the objection from moral explanations against ethical naturalism, and I try to rebut that objection by arguing that facts about human flourishing have explanatory power; the objection is then raised to me that facts about human flourishing are not intrinsically action-guiding, hence not moral facts at all, so that establishing their explanatory power amounts to no defense of the explanatory power of moral facts. I certainly agree that facts about human flourishing are not intrinsically action-guiding: there is no special to-be-doneness built into them. There had better be none, if I want to claim that they are natural facts, amenable to the same kind of empirical investigation scientific facts are subject to: intrinsically action-guiding facts would hardly find a place in the natural domain. Indeed one

---

<sup>28</sup> See Mackie [1977, chap. 1].

<sup>29</sup> Philosophers who subscribe to the account of the moral domain I am considering can have very different metaphysical views, from the view that there are no moral facts to the view that there are moral facts and they are facts about to-be-doneness. This difference would likely show up in their way of stating the objection I am considering: for instance, while the latter would start from the claim that facts about human flourishing are not intrinsically action-guiding, the former would likely object that claims about human flourishing are not intrinsically action-guiding. So far as I can see, these differences do not matter to my discussion of the objection, so for brevity's sake I'll stick to the formulation in the text.

of the standard arguments against ethical naturalism (one usually championed by defenders of the account of the moral I am considering) is precisely that since moral facts are intrinsically action-guiding, while natural facts are not, moral facts are not natural. That is: the account of the moral domain my objector relies on, though certainly neutral among many meta-ethical alternatives, is not neutral with regard to the one we are discussing about, i.e. ethical naturalism<sup>30</sup>. Rather, the claim that the moral is intrinsically action-guiding offers one of the best reasons one might have to reject ethical naturalism: if we have gone so far as to accept that account, it takes just one more easy step to get to the denial of ethical naturalism. That account provides the second premise of the objection I am considering: relying on it, my objector concludes that facts about human flourishing, which (first premise) are not intrinsically action-guiding, are not moral facts at all. Since I grant the first premise, I would ask the objector about the second: why should we accept that the moral is intrinsically action-guiding? This account, we are supposing, is the result of a meta-ethical theory licensing the conclusion that what is distinctive about the moral is that it is intrinsically action-guiding: what are the reasons for such a theory? The point of my remarks about how such a theory would lead to the rejection of ethical naturalism is not, of course, that there are no such reasons: it is rather that, once they are given, they are enough to reject ethical naturalism, and not because they reject my model of moral explanations (though they certainly do), but because of their independent weight. Those who object to my proposed model of moral explanations along these lines are not defending the objection from moral explanations, they are not reinstating the worry about moral explanations: they are rather appealing to a different objection against ethical naturalism.

Suppose my objector answers my question. Various reasons have been proposed in the meta-ethical literature (and others might of course be thought up) in support of a meta-ethical

---

<sup>30</sup> See the preceding note for the meta-ethical neutrality of the claim that the moral is intrinsically action-guiding. I do not consider here the option to claim that natural facts can instead be intrinsically action-guiding, or that intrinsically action-guiding claims can be about facts which are not intrinsically action-guiding.

theory licensing the conclusion that what is distinctive about the moral is that it is intrinsically action-guiding: such a theory provides the best explanation of the role moral judgments have in our deliberations, choices and actions; or of the nature and extent of the disagreement characterizing moral discussions; and so on. Whatever the reasons my objector appeals to, they had better not include the claim that moral facts are not explanatory<sup>31</sup>, since in that case appealing to that meta-ethical theory and the general account of the moral it yields, in order to object to my proposed model of moral explanations, would be question-begging. As we have just seen, my objector has various options available to avoid being question-begging: to answer my question, she can point to the distinctive practical role of morality (the phenomenon of moral disagreement, and so on) and claim that it gives us reason to believe that the moral is intrinsically action-guiding, and so that facts about human flourishing, which (we both agree) are not intrinsically action-guiding, are not moral facts. Suppose now she persuades me: I think about the practical role of morality, and am persuaded by my objector's considerations that the best way of accounting for it is to accept that the moral is intrinsically action-guiding, that it is laden with to-be-doneness; armed with this general account of the moral, I consider my proposed model of moral explanations, and agree with my objector that it does not establish the explanatory power of moral facts. If that, however, is what goes wrong with my model of moral explanations, then the problem with moral explanations is just a derivative problem: what is wrong with ethical naturalism has nothing to do specifically with moral explanations, it rather has to do with the practical role of morality. When the discussion with my objector is over, I am certainly no longer an ethical naturalist, but what made me change my mind? Specifically, did I reject ethical naturalism because of the objection from moral explanations? I would not say so. I do believe that moral facts are not

---

<sup>31</sup> Here is one possible (sketchy) way that might go: our moral practice has explanatory pretensions, but they are not vindicated; unless we are ready to get rid of it, the best non-debunking reconstruction we can give of our practice is to say that its primary purpose is not to deal with potentially explanatory facts, but to express one's choices with regard to policies of action, and probably influence others' similar choices. And there could of course be other ways too.

explanatory, sure enough, but I would not start from that claim if I wanted to change my old naturalistic fellows' minds: what I would likely do is rather to talk about the distinctive practical role of morality.

Someone who, against my proposed model of moral explanations, objects that facts about human flourishing are not intrinsically action-guiding, hence not moral facts at all, is therefore already appealing to some reasons to reject ethical naturalism: namely, the very reasons why she accepts that the moral is intrinsically action-guiding. As we have seen, moreover, such reasons against ethical naturalism must be independent of the worry about moral explanations: if they were not, the objection would be question-begging. They will no doubt have the consequence that facts about human flourishing are not moral, and so that my model of supposed moral explanations does not vindicate the explanatory power of moral facts, but they will not be enough to reinstate the objection from moral explanations, because they will not give it any independent weight: the independent work against ethical naturalism is done by these other reasons (the objection from moral disagreement, for instance, or whatever reason the objector relies on). If this is why my model of moral explanations fails, the objection from moral explanations can hardly be claimed to be the «basic problem about morality». I of course agree that ethical naturalism has other objections to face, besides the objection from moral explanations. The only way to assess them, however, and (if any of them proves to be fatal to it) to understand exactly what is wrong with ethical naturalism (its account of moral epistemology, or its account of moral disagreement, or its account of the practical role of morality, and so on) is to discuss them one at a time. Some philosophers believe that moral explanation is the key to understand what is wrong with ethical naturalism: this is the objection I am examining here. Pointing to other objections against ethical naturalism is not a way of defending this one.

A second objection that might be raised against my model of moral explanations does not question the evaluative character of facts about human flourishing and needs: it appeals to

it, and to a widely held view about the evaluative, i.e. its supervenience upon the non-evaluative. Consider the supposedly explanatory fact that *p* is / is not living a flourishing life: if she is / is not, it must be because of the kind of things *p* does / does not do, get, etc., the psychological, physiological, etc. characteristics of *p*, *p*'s circumstances, and so on. This whole set of non-evaluative facts is what makes it the case that *p*'s way of living does / does not satisfy her needs: it provides the supervenience base for that evaluative fact<sup>32</sup>. The objection claims that, whenever it looks as if we have reason to attribute an explanatory role to the fact that *p*'s way of living does / does not satisfy her needs, what really does the explanatory work is the supervenience base of that fact. For example, if *p* feels happy and shows a high degree of intrinsic motivation, what explains these facts about her is the fact that *p* (who has the typical psychological, physiological, etc. make-up of human beings) engages in activities that allow her to exercise her competence and autonomy, not the fact that her activities satisfy her needs. Of course we can, while observing *p*, say «She feels happy because her activities satisfy her needs», but our remark is just a quick and convenient way of pointing to the really explanatory factor, i.e. the supervenience base, and can always be replaced by an explanation citing that fact instead of facts about human flourishing.

I have two comments to make in reply to this objection. In the first place, it is not true that we can always replace an explanation citing facts about human flourishing and needs with another one citing the corresponding subvening facts. Psychologists' actual practice shows precisely that we can sometimes accept the claim that what explains *p*'s mood and behavior is the fact that her way of living satisfies her needs without knowing what it is that her living a flourishing life supervenes on, and then proceed to use that claim to test competing hypothesis about what does / does not contribute to *p*'s flourishing. In the second

---

<sup>32</sup> Think again about physical health: if *p* does / does not live a physically healthy life, it is because of various facts about the kind of food she eats, the type of activities she engages in, the characteristics of her body and her environment, and so on. As I wrote in the first chapter (n. 13), MMP too is sometimes referred to as the thesis of the supervenience of the moral on the non-moral. The relationship between the two claims is the following: MMP claims that we can only "observe" that *p* is / is not living a flourishing life by observing the occurrence of what I am here calling the supervenience base of that fact.

place, it is important to realize that even granting that the subvening facts are explanatorily sufficient by themselves to account for *p*'s mood and behavior is not enough to rule out the explanatory relevance of facts about human flourishing. That it is not is shown by the fact that a parallel argument can be mounted about biological, chemical and (on a physicalist view) psychological facts, all of which supervene on facts that are by themselves sufficient to account for their effects. If the objection I am considering proved that facts about human flourishing have no explanatory power, we should conclude that biological, chemical and psychological facts too are devoid of explanatory power. If my objector were ready to accept this conclusion, I would point to the fact that the problem thus diagnosed would be a problem neither with my model of moral explanations specifically, nor with moral explanations generally, but with (likely) any explanations but those of fundamental physics (assuming there is a level of basic, non-supervening facts in fundamental physics). Such a problem could hardly be claimed to be the reason why moral facts cannot be known in the same way in which scientific facts are known. I believe we should not accept that conclusion, however, and that, whatever the right account turns out to be of the way in which facts supervening on explanatory facts can be explanatory (a point on which I have no suggestions to make), this is what happens with biological, chemical and psychological facts; and therefore that we have not been provided with any reason to believe this is not what happens with facts about human flourishing too.

I conclude that both objections to my model of moral explanations fail, and therefore that moral facts do explain some observational facts we can use as observational criteria for them. As my emended Harman claims, this is one of the respects in which moral facts have to resemble scientific facts, if ethical naturalism is to turn out true. Against my emended Harman's claim, I have argued that moral facts do resemble scientific facts in this respect: insofar as we focus on this point, ethics is not different. If there is anything wrong with ethical naturalism, it has nothing to do directly with moral explanation.

## Bibliography

AUDI, R.: 1997, 'Ethical Naturalism and the Explanatory Power of Moral Concepts', in R. Audi, *Moral Knowledge and Ethical Character*, Oxford University Press, New York, pp. 112-128.

BAUMEISTER, R.F. and M.R. LEARY: 1995, 'The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation', *Psychological Bulletin* **117**(3), pp. 497-529.

BLACKBURN, S.: 1991a, 'Just Causes', *Philosophical Studies* **61**, pp. 3-17.

BLACKBURN, S.: 1991b, 'Reply to Sturgeon', *Philosophical Studies* **61**, pp. 39-42.

BOYD, R.N.: 1973, 'Realism, Underdetermination, and a Causal Theory of Evidence', *Noûs* **7**(1), pp. 1-12.

BOYD, R.N.: 1982, 'Scientific Realism and Naturalistic Epistemology', in P.D. Asquith and R.N. Giere (eds.), *PSA 1980. Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Vol. 2, Philosophy of Science Association, East Lansing, Mich., pp. 613-662.

BOYD, R.N.: 1983, 'On the Current Status of the Issue of Scientific Realism', *Erkenntnis* **19**, pp. 45-90.

BOYD, R.N.: 1985, 'Observations, Explanatory Power, and Simplicity: toward a Non-Humean Account', in P. Achinstein and O. Hannaway (eds.), *Observation, Experiment, and Hypothesis in Modern Physical Science*, MIT Press, Cambridge, Mass., pp. 47-94.

BOYD, R.N.: 1988, 'How to Be a Moral Realist', in G. Sayre-McCord (ed.), *Essays on Moral Realism*, Cornell University Press, Ithaca, N.Y., pp. 181-228.

BRINK, D.O.: 1984, 'Moral Realism and the Sceptical Arguments from Disagreement and Queerness', *Australasian Journal of Philosophy* **62**(2), pp. 111-125.

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

CHIRKOV, V.I. and R.M. RYAN: 2001, 'Parent and Teacher Autonomy-Support in Russian and U.S. Adolescents: Common Effects on Well-Being and Academic Motivation', *Journal of Cross-Cultural Psychology* **32**(5), pp. 618-635.

DARLEY, J.M. and C.D. BATSON: 1973, "'From Jerusalem to Jericho": a Study of Situational and Dispositional Variables in Helping Behavior', *Journal of Personality and Social Psychology* **27**(1), pp. 100-108.

DECI, E.L., R.M. RYAN, M. GAGNE', D.R. LEONE, J. USUNOV and B.P. KORNAZHEVA: 2001, 'Need Satisfaction, Motivation, and Well-Being in the Work Organizations of a Former Eastern Bloc Country: a Cross-Cultural Study of Self-Determination', *Personality and Social Psychology Bulletin* **27**(8), pp. 930-942.

GIBBARD, A.: 1990, *Wise Choices, Apt Feelings. A Theory of Normative Judgment*, Clarendon Press, Oxford.

GIBBARD, A.: 2002, 'Normative Explanations: Invoking Rationality to Explain Happenings', in J.L. Bermudez and A. Millar (eds.), *Reason and Nature. Essays in the Theory of Rationality*, Clarendon Press, Oxford, pp. 265-282.

GIBBARD, A.: 2003, *Thinking How to Live*, Harvard University Press, Cambridge, Mass.

HARE, R.M.: 1952, *The Language of Morals*, Clarendon Press, Oxford.

HARMAN, G.: 1977, *The Nature of Morality*, Oxford University Press, New York.

HARMAN, G.: 1986, 'Moral Explanations of Natural Facts – Can Moral Claims Be Tested against Moral Reality?', *The Southern Journal of Philosophy* **24** (Supplement), pp. 57-68.

- HARMAN, G.: 1996, 'Harman's Response to Thomson's Part II', in G. Harman and J.J. Thomson, *Moral Relativism and Moral Objectivity*, Blackwell, Oxford, pp. 157-187.
- HARMAN, G.: 1999, 'Moral Philosophy Meets Social Psychology', *Proceedings of the Aristotelian Society* **99**, pp. 315-331.
- MACKIE, J.L.: 1977, *Ethics: Inventing Right and Wrong*, Penguin, New York.
- MOORE, G.E.: 1903, *Principia Ethica*, Cambridge University Press, Cambridge.
- RAILTON, P.: 1986a, 'Moral Realism', *The Philosophical Review* **95**(2), pp. 163-207.
- RAILTON, P.: 1986b, 'Facts and Values', *Philosophical Topics* **14**(2), pp. 5-31.
- RAILTON, P.: 1989, 'Naturalism and Prescriptivity', *Social Philosophy and Policy* **7**(1), pp. 151-174.
- RAILTON, P.: 1998, 'Moral Explanation and Moral Objectivity', *Philosophy and Phenomenological Research* **58**(1), pp. 175-182.
- RAILTON, P.: 2003, 'Aesthetic Value, Moral Value, and the Ambitions of Naturalism', in P. Railton, *Facts, Values, and Norms. Essays toward a Morality of Consequence*, Cambridge University Press, Cambridge, pp. 85-130.
- RAWLS, J.: 1971, *A Theory of Justice*, Harvard University Press, Cambridge, Mass.
- REIS, H.T., K.M. SHELDON, S.L. GABLE, J. ROSCOE and R.M. RYAN: 2000, 'Daily Well-Being: the Role of Autonomy, Competence, and Relatedness', *Personality and Social Psychology Bulletin* **26**(4), pp. 419-435.
- ROSS, L. and R.E. NISBETT: 1991, *The Person and the Situation*, Temple University Press, Philadelphia.

RYAN, R.M.: 1995, 'Psychological Needs and the Facilitation of Integrative Processes', *Journal of Personality* **63**(3), pp. 397-427.

RYAN, R.M. and E.L. DECI: 2000, 'Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being', *American Psychologist* **55**(1), pp. 68-78.

RYAN, R.M. and E.L. DECI: 2001, 'On Happiness and Human Potentials: a Review of Research on Hedonic and Eudaimonic Well-Being', *Annual Review of Psychology* **52**, pp. 141-166.

RYFF, C.D.: 1989, 'Happiness Is Everything, or Is It? Explorations on the Meaning of Psychological Well-Being', *Journal of Personality and Social Psychology* **57**(6), pp. 1069-1081.

RYFF, C.D.: 1995, 'Psychological Well-Being in Adult Life', *Current Directions in Psychological Science* **4**(4), pp. 99-104.

RYFF, C.D. and C.L.M. KEYES: 1995, 'The Structure of Psychological Well-Being Revisited', *Journal of Personality and Social Psychology* **69**(4), pp. 719-727.

SABINI, J. and M. SILVER: 2005, 'Lack of Character? Situationism Critiqued', *Ethics* **115**(3), pp. 535-562.

SHELDON, K.M., R.M. RYAN and H.T. REIS: 1996, 'What Makes for a Good Day? Competence and Autonomy in the Day and in the Person', *Personality and Social Psychology Bulletin* **22**(12), pp. 1270-1279.

SLOTE, M.A.: 1971, 'The Rationality of Aesthetic Value Judgments', *The Journal of Philosophy* **68**(22) pp. 821-839.

STURGEON, N.L.: 1986a, 'Harman on Moral Explanations of Natural Facts', *The Southern Journal of Philosophy* **24** (Supplement), pp. 69-78.

- STURGEON, N.L.: 1986b, 'What Difference Does It Make Whether Moral Realism Is True?', *The Southern Journal of Philosophy* **24** (Supplement), pp. 115-141.
- STURGEON, N.L.: 1988, 'Moral Explanations', in G. Sayre-McCord (ed.), *Essays on Moral Realism*, Cornell University Press, Ithaca, N.Y., pp. 229-255.
- STURGEON, N.L.: 1991, 'Contents and Causes: a Reply to Blackburn', *Philosophical Studies* **61**, pp. 19-37.
- STURGEON, N.L.: 1992, 'Nonmoral Explanations', in J.E. Tomberlin (ed.), *Philosophical Perspectives*, Vol. 6: *Ethics*, Ridgeview Publishing Co., Atascadero, Calif., pp. 97-117.
- STURGEON, N.L.: 1998, 'Thomson against Moral Explanations', *Philosophy and Phenomenological Research* **58**(1), pp. 199-206.
- STURGEON, N.L.: 2002, 'Ethical Intuitionism and Ethical Naturalism', in P. Stratton-Lake (ed.), *Ethical Intuitionism: Re-evaluations*, Clarendon Press, Oxford, pp. 184-211.
- STURGEON, N.L.: 2003, 'Moore on Ethical Naturalism', *Ethics* **113**(3), pp. 528-556.
- STURGEON, N.L.: 2006a, 'Ethical Naturalism', in D. Copp (ed.), *The Oxford Handbook of Ethical Theory*, Oxford University Press, New York, pp. 91-121.
- STURGEON, N.L.: 2006b, 'Moral Explanations Defended', in J. Dreier (ed.), *Contemporary Debates in Moral Theory*, Blackwell, Oxford, pp. 241-262.
- THOMSON, J.J.: 1996, 'Moral Objectivity', in G. Harman and J.J. Thomson, *Moral Relativism and Moral Objectivity*, Blackwell, Oxford, pp. 65-154.
- THOMSON, J.J.: 1998, 'Reply to Critics', *Philosophy and Phenomenological Research* **58**(1), pp. 215-222.
- WRIGHT, C.: 1992, *Truth and Objectivity*, Harvard University Press, Cambridge, Mass.

WRIGHT, C.: 1995, 'Truth in Ethics', *Ratio* **8**(3), pp. 209-226.

ZANGWILL, N.: 2006, 'Moral Epistemology and the Because Constraint', in J. Dreier (ed.), *Contemporary Debates in Moral Theory*, Blackwell, Oxford, pp. 263-281.