Accommodating moral explanations:

Why it's harder than you might think

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Andrea Viggiano

August 2016

Accommodating moral explanations:

Why it's harder than you might think

Andrea Viggiano, Ph.D.

Cornell University 2016

This dissertation deals with one aspect of the quasi-realist program in meta-ethics: the attempt to accommodate our practice of discussing and accepting moral explanatory claims like 'Injustice causes revolutions' within a non-cognitivist framework that denies that moral properties are causally efficacious. I argue that such an attempt fails and that the problems it faces do not extend to analogous attempts by cognitivist philosophers who also reject the causal efficacy of moral properties. If you want to deny that moral properties are causally efficacious and at the same time avoid calling for a revision of our moral practice, you had better be a cognitivist.

I argue for this claim by distinguishing two non-cognitivist strategies to accommodate moral explanatory claims: one due to Simon Blackburn, the other to Allan Gibbard. Both strategies rely on the idea –widely held among meta-ethicists– that the moral depends on the non-moral: actions, people, etc. owe their moral features to their non-moral properties. I distinguish two ways of articulating this idea within a non-cognitivist framework –as a claim either in moral epistemology or about the logic of moral language– and of using it to develop an accommodationist strategy.

I then argue that both non-cognitivist strategies fail. Both ways of fleshing out the idea that the moral depends on the non-moral are inadequate: the only adequate way of articulating that idea

is as a claim in moral metaphysics, which assumes a cognitivist framework. Even granting that non-cognitivists can appeal to the idea that the moral depends on the non-moral, the strategies they build on it fail in any case to accommodate our practice with regard to moral explanatory claims: they are unable to vindicate such features of our practice as the way in which we reason about moral explanatory claims and use them when reasoning about our moral views more broadly, or the way in which we detect agreement and disagreement about moral explanatory claims.

When it comes to moral explanations, it does make a difference whether you are a cognitivist or a non-cognitivist: if you want to accommodate moral explanatory claims, you had better be a cognitivist.

BIOGRAPHICAL SKETCH

Andrea Viggiano received his B.A. degree in philosophy from the University of Florence, Italy in 2003 and his Ph.D. degree in analytic philosophy from the University of Bologna, Italy in 2007. Before joining the Ph.D. program in philosophy at Cornell University, he was a postdoctoral fellow at the University of Bologna in 2007-08.

For Pinuccio and Enzo.

We will keep building.

ACKNOWLEDGMENTS

I would like to thank Marzio Vacatello, Sergio Landucci, Paolo Leonardi, and my special committee at Cornell: Nick Sturgeon, Dick Boyd, and Karen Bennett.

I would also like to thank my fellow graduate students at the Sage School. The memory of the sound of Eric E.'s bagpipes will forever tie us all together in my heart.

TABLE OF CONTENTS

Chapter 1 - Introduction

1. <u>The issue</u>

Does honesty prevent politicians from being elected, or does it help them? Does justice make the just person's life happier or harder? Do decency and humanity in raising one's children make them grow up happy and self-confident or weak and more vulnerable to the difficulties they will later have to face? In our actual moral practice, we discuss –among others- questions like this, answering them one way or the other (sometimes, both ways). That is, we offer, discuss, and accept claims that, at least on the face of it, ascribe causal efficacy to moral properties. As a canonical form for claims of this sort, I'll use 'M-ness causes effects of type E', with 'M' a moral predicate. In discussing and accepting such claims, we seem to commit ourselves to the view that moral properties can be, and sometimes are, causally efficacious.

Are moral properties causally efficacious? Most meta-ethicists say no. Their reasons vary. There are no moral properties, say some, so *a fortiori* moral properties do not cause anything. There are moral properties, say others, but they are not the kind of properties that cause anything. Either way, moral properties are not causally efficacious. In the philosophical literature, the issue of the causal efficacy of moral properties is discussed under the heading 'moral explanations': these philosophers then reject moral explanations.

If they are right, should we then conclude that a serious mistake is involved in our practice of offering, discussing, and accepting what I'll call moral explanatory *claims*, claims of the form 'M-ness causes effects of type E'? Most of the same philosophers say no: no deep revision of our actual practice is called for by the rejection of moral explanations. Our discussing

and accepting moral explanatory claims can be accommodated, they maintain, within a metaphysical framework that denies causal efficacy to moral properties. The present work is about the prospects of this accommodationist package.

The package provides one example of an attitude present in different philosophical areas: combine a light metaphysics with a non-revisionary take on our actual practices. Within meta-ethics, this attitude is most clearly exemplified by non-cognitivists such as Simon Blackburn and Allan Gibbard, who have even coined a term −'quasi-realism'- for this project (the term is originally due to Blackburn). In this respect, what is peculiar to the debate on moral explanations is that the accommodationist camp extends to include also moral realists, i.e. those realists who think that moral properties are epiphenomenal (such as Robert Audi and Judith Thomson), thereby comprising philosophers with quite different metaphysical commitments, apart from their shared rejection of the causal efficacy of moral properties.

My view is that the accommodationist package is untenable and that we are faced with a stark choice: *either* accept that moral properties are causally efficacious *or* revise our moral practice in significant ways. In the present work, I'll focus on the non-cognitivist version of the package, which is the more carefully developed version available in the meta-ethical literature. In chapters 3-5, I'll argue that it fails: non-cognitivists cannot accommodate our practice with regard to moral explanatory claims. Along the way, I'll also draw some other -so far unnoticed- lessons about non-cognitivism. I'll have plenty to say about the realist, and more generally cognitivist, version of the package as well. For one thing, the non-cognitivist version is, as we shall see, designed to mimic the version available to the cognitivist. For another, I'll argue not just that the non-cognitivist version of the package fails, but that the problems I'll raise for it do not extend to the cognitivist version. So for both these reasons, we'll have to take a close look at

the cognitivist package in chapter 2, before turning to the non-cognitivist one in the following chapters. I won't, however, try to argue that the cognitivist version of the package fails too. That is an argument for another occasion. In the present work, my claim will be that if you like the accommodationist package, you had better be a cognitivist.

Before I begin to argue for that claim, a few preliminary remarks are in order about the accommodationist position and the arguments I'll offer in the present work.

2. <u>The accommodationist position</u>

Opponents of moral explanations usually deny that a deep revision of our actual practice is called for by their rejection of such explanations. As I have just noted, their position is analogous to that of other writers who, in other areas of philosophy, also claim that their metaphysical views do not have any radical revisionary consequences with regard to our actual practice. Phenomenalists and logical behaviorists, for instance, have typically claimed that our saying such things as 'My desk is to the left of my armchair' or 'He is in pain' can be accommodated within a metaphysical framework denying the existence of material objects or minds. Likewise, to take a more recent example, some philosophers who deny there are composite objects claim they can accommodate our saying such things as 'Here is a table'. And it is easy to think of other examples taken from other domains. In all such cases, two different options are available to the philosopher in question, with regard to the class of sentences she wants to accommodate.

One option is easier to see: it consists in claiming that such sentences do not actually mean what we might (perhaps naively) think they do. On the face of it, a moral explanatory claim of the form 'M-ness causes E' (where 'M' is a moral predicate) ascribes causal efficacy to the property of being M. The first option available to the opponent of moral explanations who

3

wants to go accommodationist is to claim that appearances are misleading here: she can offer an analysis of claims of this form on which they turn out not to ascribe causal power to the property of being M, or to any other moral property. Our practice of offering, discussing, and accepting moral explanatory claims can be accommodated, she can claim, because, once it is interpreted correctly, it embeds no mistakes and in no way conflicts with her view that moral properties do not cause anything. (Compare: the phenomenalist who wants to go accommodationist can claim that a sentence like 'My desk is to the left of my armchair' can be fully analyzed in terms of another mentioning no material objects but only, say, patterns of sense data, so that our practice of offering, discussing, and accepting such sentences in no way conflicts with her denial of the existence of material objects.)[1]

The second option available to the accommodationist philosopher is perhaps somewhat harder to see: it can be highlighted by means of still another example. Consider a sentence like 'The sun will rise at 5 a.m. tomorrow'. We still say things like this, but accommodating our practice with regard to such claims does not seem to require believing that they can be adequately analyzed in terms of sentences about the apparent position of the sun in the sky (or, even less plausibly, in terms of complex sentences on the Earth's rotating about its axis). Indeed, accommodating our practice with regard to such claims does not seem to require taking *any*

---

[1] This first accommodationist option claims that, once it is interpreted correctly, our practice with regard to the relevant class of claims (say, moral explanatory claims) does not conflict with the relevant metaphysical view (say, that moral properties are not causally efficacious). Since the claims I focus on in the present work are moral explanatory claims, it is important to note that this is not the same as claiming that, once interpreted correctly, the relevant claims turn out true (at least often enough). The difference is important because some of the opponents of moral explanations who want to go accommodationist are non-cognitivists who claim that, once interpreted correctly, moral explanatory claims are moral claims that ascribe causal efficacy to no moral property: these philosophers count as adopting the accommodationist option I am describing, but it is controversial whether they can claim that, once interpreted correctly, moral explanatory claims turn out true (I suppose this would not be a problem, *if* one assumed minimalism about truth; however, though contemporary non-cognitivists do tend to be sympathetic to minimalism, it is a good idea not to have one's taxonomy of accommodationist options rely on such an assumption, so I won't assume minimalism in drawing my taxonomy.) In areas of philosophy other than meta-ethics in which there is no position analogous to non-cognitivism, it usually does little harm to define this first accommodationist option as claiming that, once interpreted correctly, the relevant claims turn out true (at least often enough).

stand at all on the issue of what the semantic content of these claims exactly is: it seems to be enough to say that, in making such claims, we are 'onto' something, something that can be worth communicating, and that we do manage to communicate by uttering those claims, whether or not that is what such claims strictly speaking *mean*. The second option available to our accommodationist philosopher is modeled on these lines. The opponent of moral explanations who wants to go accommodationist can claim that, no matter how claims of the form 'M-ness causes E' should be analyzed, there is something potentially interesting that we are 'onto' when we make these claims, something that we manage to communicate by means of them, and that in no way conflicts with her view that moral properties do not cause anything: it is all right, she can say, that we use moral explanatory claims to convey this something. (Compare: the phenomenalist who wants to go accommodationist can claim that, no matter what the semantic content of a sentence like 'My desk is to the left of my armchair' is, when we make a claim like this we are 'onto' something: what we are 'onto' is not about material objects but rather, say, about patterns of sense data, and there is nothing wrong in our using sentences of this sort to convey this type of information.)[2]

It is useful to have a compact way to refer to the two options I have just described, so let us use the term 'payoff' to refer to what we are 'onto' and manage to communicate in making a certain claim, whether or not that is what the claim in question strictly speaking means. The two options available to the opponent of moral explanations who wants to accommodate our practice

---

[2] The example I use in the text to introduce this second accommodationist option was a standard example in 17th- and 18th-century philosophical discussions (one used, for instance, by both Berkeley and Leibniz). A common view about it was that our practice with regard to claims like 'The sun will rise at 5 a.m. tomorrow' can be accommodated because (1) in uttering such claims we are 'onto' and manage to communicate something that does not conflict with the truth of Copernicanism, even if (2) such claims are strictly speaking false. Since (1) is all that is required in order to achieve accommodation, I have dropped (2) and defined the second accommodationist option as remaining neutral on the issue of the semantic content of the relevant class of claims. (For a philosopher who thinks that claims like 'The sun will rise at 5 a.m. tomorrow' are literally true –and that our practice with regard to them can be accommodated using the first option described in the text- see Jackson [2007].)

with regard to moral explanatory claims can thus be described as follows: she can claim that the semantic content, or at least the payoff, of a moral explanatory claim of the form 'M-ness causes E' is that _____, where the blank needs to be filled in with a sentence[3] ascribing no causal power to the property of being M, or any other moral property. In the present work, I'll discuss different accommodationist proposals put forward along these lines by opponents of moral explanations.

Having just distinguished between two different accommodationist options, one dealing with the semantic content and the other with the payoff of moral explanatory claims, I'll take both of them into account in chapter 2 but, starting from chapter 3, I'll put the latter option to one side and consider accommodationist proposals framed in terms of the content of moral explanatory claims only. My reason for doing so is that, starting from chapter 3, I'll be considering accommodationist proposals put forward by non-cognitivists. How is that relevant? And why sort out accommodationist proposals –as I have been doing from the start- into cognitivist and non-cognitivist ones in the first place? After all, couldn't someone put forward one such proposal while remaining neutral on the cognitivism vs. non-cognitivism issue?

In principle, I do not see why that could not be done. As a matter of fact, however, the proposals discussed in the literature are put forward by philosophers who have a definite view about the cognitivism vs. non-cognitivism issue, and are developed so as to make them fit with such a view. As we shall see, attempts at accommodation all rely on a common idea, one that is widespread in meta-ethical discussions quite independently of the issue of moral explanations: briefly put, the idea that, for any moral feature M-ness, if something is M, it has some non-moral property that makes it M (I'll examine this idea closely in what follows). All such attempts can

---

[3] Strictly speaking, a sentential schema. In what follows, I'll omit to draw this distinction: I do not expect any confusion or misunderstanding to arise because of this.

be viewed as ways of putting this widespread idea to work so as to develop an accommodationist proposal out of it.

Now, the most natural way of fleshing out this idea, as we shall see, assumes that moral sentences can be true or false, so the accommodationist proposal based on this way of fleshing out that idea relies on cognitivism. But opponents of moral explanations who want to accommodate our practice with regard to moral explanatory claims are not all cognitivists: indeed, the most forceful proponents of the view that rejecting moral explanations does not require any deep revision of our practice are non-cognitivists. These philosophers claim they can provide a way of fleshing out the idea in question that does not assume that moral sentences can be true or false, or anything else that is not available to the non-cognitivist, and so that they too can rely on that idea and use it in order to develop an accommodationist proposal.

This explains why discussion of accommodationist proposals in the literature is carried out with an eye to the cognitivism vs. non-cognitivism issue, and why it is natural to sort out such proposals into cognitivist and non-cognitivist ones; this is what I'll do in what follows, discussing cognitivist proposals in chapter 2, and turning to non-cognitivist ones in chapters 3-5. As I said, in principle you need not be either a cognitivist or a non-cognitivist to find moral explanations objectionable and at the same time want to accommodate our practice with regard to moral explanatory claims: in this case, you can read the argument of the next four chapters as *offering* you a reason for taking sides. Either non-cognitivism or cognitivism is true; if non-cognitivism is true, then (I'll argue) accommodationist proposals fail; so if you want to go accommodationist, you had better be a cognitivist.

What does all this have to do with the distinction between proposals framed in terms of the semantic content of moral explanatory claims and proposals framed in terms of the payoff of

such claims? Suppose you are a non-cognitivist. Then if s is a sentence containing moral terminology, you owe us an answer to the question 'What is your story about the semantics of s?', one on which it is not the purpose of such a sentence to report the occurrence of a moral fact. Suppose now you reject moral explanations (as of course you do, since you believe there are no moral properties) but want to accommodate our practice with regard to moral explanatory claims: though in principle you have two options available, it is only natural for you to frame your accommodationist proposal in terms of the semantic content (as opposed to the payoff) of such claims, since you owe us a story about their semantic content anyway[4].

As I say, this is only natural for you to do, but you are certainly not logically committed to it by your position. It would not be inconsistent of you to claim that there is a story about the semantic content of moral explanatory claims on which it is not the purpose of such claims to report the occurrence of a moral fact, but that the semantic content individuated by such a story is different from what your proposal offers as their payoff. But then why not plug in the former, instead of the latter, into your accommodationist proposal, and frame this proposal in terms of the content of moral explanatory claims?

---

[4] Are things different if you are a cognitivist? Yes, they are. If you are a cognitivist, then faced with a sentence s containing moral terminology and with the question 'What is your story about the semantics of s?', there is an answer that is always available to you, namely that the semantic story to tell about s is the one we get if we take s at face value. So, for instance, if s is 'That action is wrong', then the semantic story to tell about s is (to put it briefly) that s ascribes the property of being wrong to the action in question; if s is 'Injustice causes revolutions', the semantic story is (briefly put) that s ascribes causal efficacy, with regard to the bringing about of revolutions, to the property of being unjust; and so on. Let us call this the 'face value' answer. Note that the face value answer is available to you even if you believe there are no moral explanations. And, if you reject moral explanations and want to accommodate our practice with regard to moral explanatory claims, you are under no pressure to frame your accommodationist proposal as one about the semantic content of moral explanatory claims. Of course you can, if you want to (thereby replacing the face value answer with your proposal, as your answer to the question 'What is your story about the semantics of moral explanatory claims?'), but you are under no pressure to: you can frame your accommodationist proposal in terms of the payoff of moral explanatory claims and, if asked about their semantic content, appeal to the face value answer (or point to the face value answer and your proposal and say that one of the two -you remain neutral about which- individuates the semantic content of such claims). That is, if you are a cognitivist, you can always rely on the face value answer as a fallback answer to the question 'What is your story about the semantics of s?', for any sentence s containing moral terminology, including a moral explanatory claim. But of course this is not so if you are a non-cognitivist, since according to non-cognitivism the face value answer is the wrong answer to that question.

Nor would it be inconsistent of you to refuse the burden of providing, for *each* sentence containing moral terminology, a story about its semantic content on which it is not the purpose of that sentence to report the occurrence of a moral fact: you could claim that, even if there is a type of such sentences whose semantic content is not given by a story along these lines (maybe they just have no semantic content, i.e. they turn out to be meaningless; or maybe their purpose is to report the occurrence of a moral fact, i.e. they turn out to be an exception to the non-cognitivist nature of moral language), this is no big problem for your non-cognitivism; and you could claim that this is precisely what happens with moral explanatory claims. And if you were *merely* a non-cognitivist, you might perhaps be right, provided that you have a non-cognitivist story to offer about most other types of moral sentences, and that moral explanatory claims are somewhat peripheral to our moral thought. But since you *also* want to accommodate our practice with regard to moral explanatory claims, I doubt you are in a position to think that such claims are peripheral to our moral thought: why, after all, would you bother putting forward a proposal in order to accommodate our practice with regard to them, unless you thought they are central enough to our moral thought that rejecting them all would amount to a serious revision of our actual practice?

Now, I do not mean to suggest that nothing could be said in reply to these questions, nor that the alternatives I have just mentioned exhaust all the relevant possibilities. However, these remarks show, I think, why the obvious way to go, for the non-cognitivist who wants to accommodate our practice with regard to moral explanatory claims, is to frame her proposal as one about the content (not merely the payoff) of such claims. And they help explain why, as a matter of fact, actual non-cognitivist proposals in the literature are explicitly put forward (or at least generally taken and discussed) as proposals about the content of moral explanatory claims.

In my discussion of non-cognitivist proposals in chapters 3-5, I'll therefore put the payoff option aside and assume that such proposals deal with the semantic content of moral explanatory claims. As I said, in chapter 2, when discussing cognitivist proposals, I'll continue to take account of both accommodationist options.

3. <u>A quick note on the arguments to follow</u>

In the present work, I said in section 1, I'll argue that the non-cognitivist version of the accommodationist package fails, and that the problems I'll raise for it do not extend to the cognitivist version of the package. All attempts at accommodation, I noted in section 2, rely on the common idea that, for any moral feature M-ness, if something is M, it has some non-moral property that makes it M. Some of the arguments I'll offer against the non-cognitivist who wants to go accommodationist are arguments to the effect that, *contra* what non-cognitivists claim, that idea *is not available* to them. Others are arguments to the effect that, even if that idea is available to them, non-cognitivists *cannot in any case use it* to develop an accommodationist proposal in the same way in which cognitivists can.

Now, as I pointed out in section 2, the idea in question is widespread in meta-ethical discussions, quite independently of the issue of moral explanations: meta-ethicists in general, and non-cognitivists in particular, appeal to it in various other debates as well. If my arguments in what follows work, they therefore have quite general implications: they point to a general problem with non-cognitivism, affecting the treatment of the various issues with regard to which non-cognitivists appeal to that idea. In the discussion to follow, I won't have room to explore these general implications of my arguments. I will, however, keep track of one such issue,

looking at how my arguments affect the non-cognitivist account of the content, not just of moral

explanatory claims, but of moral sentences in general.

And now it is time to turn to this widespread idea and take a closer look at it.

Chapter 2 – The cognitivist strategy

1. <u>The Assumption</u>

All accommodationist proposals put forward in the literature, be they cognitivist or non-cognitivist, rely on a common idea. This is an idea that is widely shared among writers in meta-ethics, quite independently of the issue of moral explanations; what friends of accommodationist proposals do is to put this idea to work in the context of discussions of moral explanations, in order to develop an accommodationist proposal out of it. Let me start by giving a rough and intuitive characterization of the idea in question, and of the way in which it might be thought to provide a useful tool to develop an accommodationist proposal, before presenting the idea, and the proposals built on it, in a more formal way.

In the previous chapter I formulated the idea in question as follows: for any moral feature M-ness, if something is M, it has some non-moral property that makes it M. This claim is meant to capture and give an account of the following feature of the way in which we conduct our moral discussions. Suppose someone says, of a certain action (person, institution, etc.), that it is wrong (good, unjust, etc.): it seems it is always appropriate for us to ask 'What is wrong about it (good about them, etc.)?'. Our interlocutor might well be unable to offer a detailed answer (or maybe even anything definite enough to count as an answer at all) to our question, still our question seems to be one she must (and, in actual moral discussions, does) admit as relevant[1]. There *must*, meta-ethicists say, be some property in virtue of which that action is wrong (that

---

[1] Strictly speaking, there are of course many situations in which it would not be appropriate to ask 'What is M about it?': for example, the answer may be obvious, raising the question may be impolite, or it may fail to make sense conversationally in some other way (consider 'What is wrong about the Holocaust?'). In the text, I am abstracting away from such standards of appropriateness and thinking of the standards more centrally in place in a conversation among philosophers or when someone is debating with him/herself.

person is good, etc.), some property that *makes* it a wrong action, whether or not the person making the judgment has any definite idea about what that property is; and, so goes the widespread idea, such property is a non-moral property: actions, people, institutions, etc. have the moral features they do in virtue of their non-moral properties[2]. Since this is a widely shared idea among writers in meta-ethics, it deserves a name of its own: I'll call it the Assumption.

Grant now the Assumption and consider a moral feature M-ness. Given the Assumption (and assuming that at least one thing is M), there is a non-empty set of (one or more) non-moral properties that make things M: the set of the M-making properties, as philosophers sometimes call them. Consider now a moral explanatory claim of the form 'M-ness causes E'. What someone who wants to put forward an accommodationist proposal needs is a sentence ascribing no causal power to the property of being M (or any other moral property) that she can propose as capturing the content, or payoff, of such a claim. The Assumption points to a route to such a sentence: just propose that the content, or payoff, of 'M-ness causes E' is that the M-making properties cause E. As we shall see, all accommodationist proposals (be they cognitivist or non-cognitivist) can be viewed as ways of implementing this insight[3].

---

[2] Since this idea is claimed to be available to both cognitivists and non-cognitivists, I avoid framing it as the view that actions, people, institutions, etc. have the moral *properties* they do in virtue of their non-moral properties. In what follows, I'll appropriate the term 'feature' for formulations that remain neutral with respect to the cognitivism vs. non-cognitivism issue (my choice of the term is entirely arbitrary). This explains why in the main text I formulate the idea in question as 'For any moral feature M-ness, …'; the standard non-cognitivist gloss on this is 'For any moral predicate 'M', …'.

[3] Accommodationist proposals along these lines have been put forward (among cognitivists) by Audi [1993] and Thomson [1998] and (among non-cognitivists) by Blackburn [1991a] and Gibbard [2003]. Of course, the suggestion that the content, or payoff, of 'M-ness causes E' is that the M-making properties cause E could in principle be correct even if the Assumption should be rejected, but in such a case it would face obvious problems as an *accommodationist* proposal put forward by an opponent of moral explanations. The Assumption guarantees that the set of M-making properties is non-empty (thereby guaranteeing that our practice of offering, discussing, and accepting moral explanatory claims is not like one of offering, discussing, and accepting claims about, say, the square roots of 2 larger than 3 or the squared circles with a radius at least one inch long) and contains only non-moral properties (thereby guaranteeing that our practice does not embed anything incompatible with the rejection of moral explanations). Now, as I say in the main text, the Assumption does not guarantee these things all by itself: one also needs the supposition that at least one thing is M. And one might wonder whether the need for this supposition creates problems for the accommodationist recipe sketched in the text. There are two different possible concerns to distinguish here: one quite general, the other more specific. The quite general concern is: what if nothing is M, for

As I say, the Assumption is a widely shared idea among writers in meta-ethics: indeed, if one looks at various debates in the literature (including the debate on moral explanations), it is fair to say that it is an idea often regarded not merely as correct, but as obviously correct and not in need of much defense. I myself am much more doubtful of it than most meta-ethicists seem to be and believe that, whether or not the Assumption is in the end correct, it cannot in any case be relied on without defense[4]. Of course, if there is something to these doubts, this might end up raising quite a serious problem for any accommodationist proposal: at a minimum, friends of these proposals should stop taking the Assumption for granted and begin to offer arguments in its defense, before proceeding to use it in order to build their proposals on it; and developing such a defense might actually prove quite hard to do. In my discussion in the present work, however, I'll bracket my worries about the Assumption, because the issues raised by them are distinct from those I want to raise about the accommodationist project. My claim in this work will be that this project faces serious problems *even if* one grants that actions, people, institutions, etc. owe their moral features to their non-moral properties.

---

any (or almost any) moral feature M-ness? In the context of the present discussion, this concern is one that neither party can raise. The present discussion is about whether rejecting moral explanations calls for a deep revision of our actual practice of offering, discussing, and accepting moral explanatory claims, where both parties assume that failing to accommodate this aspect of our actual practice is a cost (maybe not an unbearable one, but certainly a cost). But failing to accommodate this aspect of our actual practice could hardly be seen as a cost, if our whole moral practice were seriously misguided, as it would be if nothing were M, for any (or almost any) moral feature M-ness. A shared presupposition of the present discussion is thus that this quite general concern can be bracketed (which of course does not mean that such a presupposition cannot be called into question). The more specific concern is: what if nothing is M, for a certain specific moral feature M-ness (or only a few of them)? Would this create problems for the accommodationist recipe sketched in the text, even given the Assumption? I believe this is an interesting and complex question, about which various things could be said on behalf of both sides. To mention just one point: it might be argued, with some plausibility, that we rarely discuss moral explanatory claims involving a moral feature that nothing has, so that even if the recipe described in the text turned out not to be a good recipe for achieving accommodation with regard to such claims, this would not be a serious problem for the accommodationist project. On the other hand, don't we sometimes ask such questions as whether living in a perfectly just state would make our lives happier? Or isn't the issue that Socrates agrees to address in Plato's *Republic* of precisely this sort, i.e. whether the life of the fully just person is the happiest life for a human being? In what follows, I'll bracket this question and grant that no serious problems arise for the accommodationist project from consideration of moral features that nothing has.

[4] Of course, I am not alone: other meta-ethicists share these doubts. It is fair to say, though, that this is definitely a minority opinion. For discussions casting doubts on the Assumption, see McDowell [1979], Williams [1985], Griffin [1992], Sturgeon [2009], and the debate on moral particularism (see, for instance, Hooker & Little [2000]).

2. <u>The cognitivist reading of the Assumption</u>

I now want to flesh out the Assumption, and the accommodationist recipe based on it, in a more explicit and detailed way than I did in the previous section. In doing so, I take myself to be giving a (more explicitly worked out, under certain respects at least, and more detailed, but still) fair reconstruction of the accommodationist proposals put forward in the literature. Friends of such proposals do not usually pause to make explicit all the details I'll call attention to; my reason for pausing and making them explicit is that I believe interesting questions can be raised, and interesting lessons can be learnt, once we attend to them.

I'll proceed as follows. In the present chapter, I'll describe the most natural way of fleshing out the Assumption and of using it in order to develop an accommodationist proposal. As we shall see, this way of fleshing out the Assumption freely assumes that moral sentences can be true or false and is thus available only to a cognitivist. Starting from the next chapter, I'll describe the way in which non-cognitivists propose to flesh out the Assumption without assuming that moral sentences can be true or false (or anything else that is not available to them) and the accommodationist proposal they build on this way of taking the Assumption. As we shall see, there is more than one such way and one such proposal: the three chapters to follow will be devoted to assessing these non-cognitivist proposals. My assessment will be negative: the non-cognitivist proposals, I'll argue, fail to accommodate our actual practice. The problems they face, moreover, are avoided by the cognitivist proposals presented in this chapter.

Here is how I formulated the Assumption in the previous section: for any moral feature M-ness, if something is M, it has some non-moral property that makes it M. As I said, this is a rough and intuitive way of framing the Assumption: as it turns out, the picture implicit in the

Assumption is complex enough that it takes more than one thesis to articulate it[5]. What exactly is that picture? And what exactly is it for a property to be an M-making property? Suppose that, in order to have some fun, I get a stick and start beating my cat: what I do is (we can suppose) wrong. What does the Assumption say about my action? It says that it has a non-moral property, call it 'N-ness', that plays a certain role. What role is that? According to the Assumption, its being N makes my action a wrong action, i.e. the Assumption says that my action is N and N actions are wrong, and that it is because of these two things that my action is wrong: what makes it the case that my action is wrong is that it is N and N actions are wrong. Consider again the way in which we conduct our moral discussions. Suppose you were to tell me 'Stop doing that: it's wrong!' and I were to ask you 'What is wrong about it?': what would I be asking you for? I would (it seems right to say, and the Assumption does say) be asking you to point to a property with precisely the features just mentioned. Suppose you were to reply 'Really? You're causing pain to a cat for fun! This is what's wrong about it.': what you would be doing is to point to the non-moral property 'causing pain to a cat for fun' and say that my action has this property and actions with this property are wrong, and that this is what makes it true (as you have been telling me) that my action is wrong.

Suppose now that you are right. My action is wrong: what is its wrong-making property? Is it its causing pain to a cat for fun? So far as the Assumption is concerned, this is one

---

[5] Indeed (as I'll sometimes point out explicitly), some of the elements of that picture are not strictly speaking entailed by the rough formulation of the Assumption given in the previous section (at least under certain ways of reading of it), which is due to the fact that that formulation is indeed rough, so it is important to go beyond it: the picture in question is complex enough that it makes things clearer to start with a simplified sketch of the Assumption and then present the fuller picture in detail, but of course a simplified sketch is bound to leave some details out. Why, it might be asked, not stop at the simplified sketch? Why not take the formulation in the previous section to capture strictly speaking everything the Assumption is meant to say? Because the Assumption is a tool philosophers rely on in various debates in meta-ethics in order to put it to work in those debates and do things with it: in the debate on moral explanations, for example, the Assumption is used in order to develop an accommodationist proposal. And it is not clear that it can do the job it is meant to do unless it is read so as to include the details I'll make explicit: as will be clear from what follows, the accommodationist proposals I'll discuss rely on those details, and it is not easy to think of ways of developing an accommodationist proposal that do not. If someone claims they can build a plausible accommodationist proposal with less, I'll be glad to examine it once it has been put forward.

possibility. If nothing else can be said about my action in addition to what you have already told me, then my action's wrong-making property is its causing pain to a cat for fun. A different possibility, one that is also allowed for by the Assumption, is that more can be said about my action than what you have already told me: more, in particular, can be said about the moral judgment you appealed to in making your remark. Suppose that, in reply to your remark 'You're causing pain to a cat for fun!', I were to ask 'And what is wrong about *that*?'. What I would be doing is to ask about the moral judgment you appealed to: you appealed to 'Actions that cause pain to a cat for fun are wrong', I would be asking 'What is wrong about them?'. And it is a possibility allowed for by the Assumption that there is something to be said in reply to this: maybe it is that such actions cause pain to a (non-human) animal for fun; and maybe what is wrong about *these* actions in turn is that they cause pain to a sentient being (be it human or non-human) for fun. Or maybe the right story to tell here tracks a different path. No matter what the right path is, however, the overall structure of the story is the same. What makes it true that my action is wrong is that it has a certain non-moral property, N-ness (causing pain to a cat for fun, as we are supposing), and all N actions are wrong; but this is not the whole story about my action, since the following can be added to it: what makes it true that all N actions (of which mine is one) are wrong is that they have a certain non-moral property, N*-ness, and all N* actions are wrong. As I said, that this is the case about my action is a possibility allowed for by the Assumption, but the Assumption does say that, if there is something else that can be said about my action in addition to what is conveyed by your remark, if (as I put it) there is a path to be tracked beyond 'All actions that cause pain to a cat for fun are wrong', this path sooner or later reaches an end, i.e. there is a non-moral property (what I called 'N*-ness' above) such that my action is N* and all N* actions are wrong, and nothing else can be said about my action in

addition to *this much*, nothing can be said in reply to the question 'What is wrong about N* actions?' (except perhaps that N* actions are wrong *qua* N* actions), so that my action is ultimately made wrong by the fact that it is N* and all N* actions are wrong. If this is the case about my action, then my action's wrong-making property is its being N*.

Let us call my action 'a' and consider the sentence 'a is wrong', a sentence that (we are supposing) is true. What the Assumption says about this truth is thus the following. The truth in question is made true by another moral truth (plus an appropriate non-moral truth); this moral truth may or may not in turn be made true by a further moral truth (plus an appropriate non-moral truth) –and if it is, the same applies to this further moral truth, and so on-; either way, the truth in question is ultimately made true by a moral truth (plus an appropriate non-moral truth) that is not made true by other truths; the latter moral (and non-moral) truths are truths to the effect that all N's are wrong (and a is N), for some non-moral property N-ness; such non-moral property is a's wrong-making property.

Let me now introduce a little bit of terminology so that I can state what the Assumption says in a quite general and more compact way. Let <u>individual</u> moral sentences be sentences of the form 'x is M', where 'M' is a moral predicate and 'x' is a singular term; let <u>universal</u> moral sentences be sentences of the form 'All P's are M', where 'M' is a moral predicate and 'P' is a predicate; finally, let <u>categorical</u> moral sentences be moral sentences that are either individual or universal. When an individual / universal moral truth to the effect that x is M / all P's are M is ultimately made true by a moral truth to the effect that all N's are M, plus a non-moral truth to the effect that x is N / all P's are N, for some non-moral property N-ness, I'll say that the individual / universal moral truth is ultimately made true <u>in the relevant way</u>[6]. With this

---

[6] As I'll mention in a short while, this definition of 'being ultimately made true in the relevant way' is actually too narrow to be adequate for our purposes and will need to be broadened, but for the time being we can work with this

terminology in place, the picture implicit in the Assumption can be articulated by means of four different theses: I'll lay down the four theses and then make several clarificatory remarks about them. Here are the four theses:

i) All categorical moral truths are either <u>derived truths</u>, i.e. they are made true by other categorical moral truths (plus appropriate non-moral truths) or <u>brute truths</u>, i.e. they are not made true by other truths.

ii) All derived moral truths are ultimately made true by brute moral truths (plus appropriate non-moral truths).

iii) All derived moral truths are ultimately made true in only one way, i.e. the relevant way.

iv) All brute moral truths can be expressed by an appropriately strong universal moral sentence using a non-moral predicate denoting a non-moral property, i.e. by an appropriately strong sentence of the form 'All N's are M', for some non-moral predicate 'N' denoting a non-moral property[7].

simpler version. I should point out explicitly that, in formulating the Assumption, I'll throughout use the phrase 'being ultimately made true by' so that, in order for a moral truth m to be ultimately made true by a moral truth m', plus an appropriate non-moral truth n, it is required that m' not be made true by other truths, but it is *not* also required that n not be made true by other truths: the Assumption is, after all, a claim about ethics and it need not take a stand on how things are in the non-moral domain.

[7] In the present work, I'll distinguish between brute and derived truths only within the set of categorical moral truths, so I'll throughout feel free to use 'brute (derived) truth' or 'brute (derived) moral truth' as shorthand for 'brute (derived) categorical moral truth'. Theses i)-iv) should be read accordingly. I'll explain what I mean by 'appropriately strong universal moral sentence' in due course. For ease of exposition, in the present work I'll bracket the view, held by some writers in meta-ethics, that some moral properties are reducible to others (for example, that being right is the same as maximizing goodness) and frame i)-iv) as theses about *all* categorical moral truths. If you accept the view in question, then the Assumption commits you to accepting i)-iv) only with respect to the categorical moral truths that ascribe the non-reducible moral property (or properties) to which the others can be reduced; you are then free to start from such categorical moral truths and account for all the others by plugging in your favorite reductions. In addition to interesting versions of the view in question which you may or may not accept, there are somewhat less interesting versions of it: for example, that the property of being M* is reducible to that of being M, where being M is a moral property and the predicate 'M*' is introduced as shorthand for 'not M'. Once again, the Assumption does not commit you to accepting i)-iv) with respect to the categorical moral truths ascribing M*-ness.

19

According to i)-iv), categorical moral truths form a neat, structured set. Some of them (the brute ones) are not made true by other truths and are (together with non-moral truths) what ultimately makes all the others (the derived ones) true. The non-moral properties picked out by those, among the brute truths, that ascribe M-ness can (for any moral property M-ness) be thought of as the ultimate source of all the M-ness instantiated in the world: these properties are the M-making properties.

The four theses are neutral as to the number of brute moral truths there are for a given moral property M-ness: so far as i)-iv) go, there might be only one such truth, and so only one M-making property; or there might be more than one, and so more than one M-making property. According to ii), all derived moral truths are ultimately made true by brute moral truths (plus appropriate non-moral truths): given the possibility that there be more than one brute moral truth ascribing M-ness, it is possible that a derived moral truth be ultimately made true by a series of such brute moral truths, as opposed to a single such truth (plus an appropriate non-moral truth). For example, it might be that 'All P's are M' expresses a derived moral truth, but that there is no single M-making property that all the P's instantiate: rather, there is a list of M-making properties such that each P instantiates at least one property in the list. And as a matter of fact, for the purposes of articulating the picture implicit in the Assumption, the definition of 'being ultimately made true in the relevant way' given above should actually be broadened so as to include the possible case I have just described as well, in addition to the two cases previously described: that is, it should be broadened so as to include also the case in which a universal moral truth to the effect that all P's are M is ultimately made true by a series of different moral

---

Parallel remarks apply, *mutatis mutandis*, to the scope of the other sets of theses I'll present in sections 3 and 4 of this chapter and in chapters 3, 4, and 5.

truths to the effect that all N_1's are M, all N_2's are M, …, where N_1-ness, N_2-ness, … are different non-moral properties, plus the non-moral truth that each P is either N_1 or N_2 or … .

In addition to being neutral as to the number of brute moral truths there are for a given moral property M-ness, the theses are also compatible with the possibility that, starting from a derived moral truth and asking 'What makes it true?', there are different paths to track that end in different (series of) brute moral truths. So, for example, the theses are compatible with the possibility that an object x get (so to speak) multiple shots of M-ness, i.e. that it instantiate more than one M-making property: they are compatible with (say) the possibility that 'x is M' be ultimately made true both by the brute moral truth that all N_1's are M (plus the non-moral truth that x is N_1) and by the brute moral truth that all N_2's are M (plus the non-moral truth that x is N_2), where N_1-ness and N_2-ness are two different properties. Or they are compatible with (say) the possibility that 'All P's are M' be ultimately made true both by the brute moral truth that all N_1's are M (plus the non-moral truth that all P's are N_1) and by the two brute moral truths that all N_2's are M and all N_3's are M (plus the non-moral truth that each P is either N_2 or N_3), where N_2-ness and N_3-ness are two different properties. No matter how many paths can be tracked starting from a derived moral truth t, and no matter what their specific details turn out to be, there is something they all share in common: according to iii), the (series of) brute moral truth(s) each of them ends in, and the appropriate non-moral truth that together with it ultimately makes t true, are such that they ultimately make t true in the relevant way, i.e. they fit one of the three different cases distinguished above when defining 'being ultimately made true in the relevant way'[8].

---

[8] Friends of the Assumption do not usually devote much attention to sentences of the form 'All M's are M*', where both 'M' and 'M*' are moral predicates. Thus, in articulating the picture implicit in the Assumption, one option consists in simply ignoring such sentences and articulating that picture in such a way that it says nothing about them. I chose instead to articulate that picture in such a way that it does apply to such sentences (my definition of 'universal moral sentence' makes room for them), while at the same time preventing their inclusion in the picture from running afoul of its central tenets (though my reconstruction allows that sentences of this form can be true, none of the truths in question, if indeed there are any, is a brute moral truth and they are all, just like any other

The theses are formulated in terms of the relation 'being made true by'. Providing a detailed characterization of this relation would be a task as interesting as it is daunting, a task I cannot undertake here because it would make for the topic of a different work, one in metaphysics rather than meta-ethics (actually, a task I would not be competent to undertake anywhere). Luckily enough, providing such a detailed characterization is not necessary in order to elucidate the four theses above: gesturing at the relation in question so as to give a working idea of it is enough. And I have already done this, because in introducing the Assumption I have already pointed to a place where the relation at issue is meant to be at work. As I have already mentioned, the Assumption is meant to capture and give an account of the following feature of the way in which we conduct our moral discussions: when someone says something of the form 'x is M' (where 'M' is a moral predicate), it seems it is always appropriate for us to ask 'What is M about it?'. Now, if *we* were asked why we find it appropriate to raise this question, it would be hardly possible to say anything in reply that did not invoke the likes of 'Because there must be something about it that *makes* it M' or 'Because, if it is M, it must be M *in virtue of* its being something else' (and we do, as a matter of fact, say things of this sort in our moral discussions), i.e. without using language that makes (more or less explicit) reference to a relation of making it the case or making true. The relation in whose terms the theses above are formulated is the relation (whatever its exact nature turns out to be) that we seem to (at least implicitly) make reference to when, in our moral discussions, we raise that pervasive question 'What is M about

derived moral truth, ultimately made true in the relevant way by brute moral truths with the features specified in iv), plus appropriate non-moral truths). My reason for choosing this option is that I take it to be a strength of the picture implicit in the Assumption that it can be articulated so as to accommodate sentences of this form too. I mention this point only for clarity and completeness: the difference between the two options I have just described will not play any role in my discussion. I should perhaps note that, given my choice, when I say that derived moral truths are made true by other categorical moral truths plus appropriate non-moral truths, the non-moral truths in question can be truths to the effect that all M's are N, with M-ness a moral property. I myself do not find this objectionable; if someone does, they should feel free to replace 'non-moral truth' with 'truth ascribing a non-moral property'. Again, I mention this only for clarity and completeness, not because it will play any role in my discussion.

it?'. It seems to me a friend of the Assumption would be right in claiming that the relation 'being made true by' is one that, once it is pointed at this way, it is intuitive enough to recognize, and that we have a (no doubt partial, but still) clear enough grasp of this relation for us to be able to work with it and with the four theses above, which are framed in terms of it[9].

Even this rough characterization of the relation 'being made true by' is enough to bring out a feature of it that it is important to highlight, in discussing the four theses above. This is that the relation 'being made true by' is a straightforwardly metaphysical relation: in particular, it is a *different* relation from the logical relation 'following from'. More accurately (and more interestingly, since of course q can logically follow from p even if p is not true or q is not true, whereas q can only be made true by p if both p and q are true), q's being made true by p is not the same as p's being true and q's logically following from p. In order to see this, consider Dave: he is a philosophy graduate student, and he is a good person. I can tell you because I know Dave: he and I are in the same year in the philosophy graduate program. Actually, I can tell you, all the people in my year are good people. So the following two sentences are both true: 'Dave is in my year' and 'All the people in my year are good people'. And, of course, that Dave is a good person logically follows from these two sentences. Now ask yourself: do these two sentences make it true that Dave is a good person? Not at all. If you were to ask 'What is good about Dave?' and I were to reply 'He's in my year', you would probably think I have quite bizarre

---

[9] As I said above, I consider the project of providing a detailed characterization of the relation 'being made true by' invoked in i)-iv) a very interesting one. One of the questions to address, in pursuing it, would be whether, and if so how, this relation is connected to the relation(s) that contemporary metaphysicians discuss under the heading 'grounding relation(s)'; and one of the possible answers would of course be that the relation 'being made true by' invoked in i)-iv) just is (one of) the grounding relation(s). If correct, this answer would allow me to enrich my rough characterization of the relation 'being made true by' in the text by adding that it is (one of) the relation(s) that contemporary metaphysicians talk about in discussing about grounding (whatever the exact nature of *that* turns out to be). The reason why I do not do this is that I do not know whether this answer would be correct and I mean to remain neutral throughout the present work with regard to the question I have just pointed out, mainly because the issues involved are complex ones, and I believe I can say everything I want to say here without having to address them. For a recent and illuminating discussion of grounding (or, as she prefers to call them, building) relations, see Bennett [forthcoming].

moral views, or hold weird, nearly fanatic views about my year in the program (or, more likely, that I misunderstood your question): in any case, you would reject my answer. And you would be right in doing so[10]. If (*a*) 'x is M' (or 'All P's are M') is made true by 'All N's are M' and 'x is N' (or 'All P's are N'), then no doubt (*b*) the latter couple of sentences are true and the former logically follows from them; still, (*a*) does not amount merely to (*b*), since 'x is M' (or 'All P's are M') could logically follow from a couple of true sentences without being made true by them. Therefore the instantiation of the relation 'being made true by' reported by (*a*) amounts to something more than what is reported by (*b*), something that cannot be captured in terms of the logical relation 'following from'.

According to iv), all brute moral truths can be expressed by an appropriately strong universal moral sentence. I use 'strong sentence' as shorthand for 'sentence that supports counterfactuals'. Why a *strong* universal moral sentence? Because this is how we treat proposals about what counts as a brute moral truth in our moral discussions. In such discussions we do not, of course, speak in terms of 'brute moral truths', but we do put forward and discuss proposals about what ultimately, or at bottom, or in the final analysis (as we might put it) makes things M. For example, suppose you were to tell me that what at bottom makes my action of beating my cat wrong is that it causes pain to a sentient being for fun: in terms of the Assumption, and of the theses i)-iv) that articulate it, you would be proposing that 'causing pain to a sentient being for fun' is a wrong-making property, or that 'All actions that cause pain to a sentient being for fun are wrong' is a brute moral truth. One obvious way for me to object to your proposal would be to

---

[10] I am assuming (what I take to be fairly uncontroversial) that you would be right in rejecting my answer. However, even if it were to turn out that you are not right, the story about Dave still brings out the gap there is between the instantiation of the logical relation and that of the metaphysical one: even if it were to turn out (surprisingly) that my answer is correct, that it is correct is not *guaranteed* by the fact that 'Dave is in my year' and 'All the people in my year are good people' are both true and 'Dave is a good person' logically follows from them. Said otherwise, when we ask 'What is M about x?', it is not the case that an answer of the form 'x is N', for some 'N' such that both 'x is N' and 'All N's are M' are true, counts *ipso facto* as a correct answer.

produce an example of an action that, though causing pain to a sentient being for fun, fails to be wrong; and if I were to try to produce such a counterexample, I would likely not bother looking for an actual action with the desired characteristics: producing an example of a hypothetical such action would be enough. This appeal to hypothetical examples is commonplace in our moral discussions and shows that proposals about what counts as a brute moral truth are meant to cover possible hypothetical, counterfactual cases as well as actual ones: by requiring that a brute moral truth be expressible by a strong universal sentence, thesis iv) is meant to capture this aspect of the way in which we conduct our moral discussions[11].

What do I mean by 'appropriately strong', and so what counts as an *appropriately* strong universal moral sentence? Given two sentences both of which support counterfactuals, we sometimes want to compare them in terms of what might be called their degree of strength and say that one, so to speak, sticks its neck out more than the other does. Consider, for example, 'All the soldiers at place p and time t are wearing an iron helmet'[12] and 'All iron helmets melt at n degrees Celsius'. Suppose the former is true because orders have been given which guarantee that soldiers at the place and time in question all wear an iron helmet and the latter is true due to a law of nature: under these circumstances, both sentences support counterfactuals. Even if they both do, we might want to draw a difference between them and say that one is, so to speak, more robustly counterfactual-supporting than the other, or that even if they both hold not merely in the

---

[11] Though appeal to hypothetical examples is commonplace in moral discussions, some philosophers resist appealing to *far-fetched* hypothetical examples, for various reasons: though brute moral truths cover hypothetical as well as actual cases, they are not necessary and so do not cover *all* hypothetical cases, including far-fetched ones; though brute moral truths are necessary and so do cover all hypothetical cases, our intuitions about far-fetched cases are not reliable, so we should not use them to test proposals about what counts as a brute moral truth; though brute moral truths do cover all hypothetical cases, it is pretty hard to find out about them, so it is good enough if, in our moral discussions, we aim for something less, something that covers only realistic enough cases. As I said, philosophers who press these worries mean to object to the use of far-fetched hypothetical examples, not to the use of any hypothetical examples in moral discussions, so what I point out in the main text is compatible with taking their worries seriously.
[12] I borrow this example from Armstrong [1983], p. 47 (Armstrong uses it to make a different point).

actual world but also in hypothetical, counterfactual ones, it is in some sense easier for one to fail (for someone to counterfactually be a soldier at the place and time in question and fail to wear an iron helmet) than it is for the other (for something to counterfactually be an iron helmet and fail to melt at the temperature in question). Here is a rough-and-ready way of making this thought more precise: given two sentences, s and r, both of which support counterfactuals, s is stronger than r just in case –supposing them true- the changes one needs to make to the actual world in order to have a failure of s are more radical than those needed in order to have a failure of r (or no matter what changes one makes to the actual world, one never has a failure of s, whereas there are changes that lead to a failure of r). Using the standard tool of an ordering of possible worlds according to a similarity metric: s is stronger than r just in case –supposing them true- one needs to travel further away from the actual world in order to reach a possible world where s fails than one does in order to reach one where r fails (or no matter how far away one travels, one never reaches a possible world where s fails, whereas there are worlds where r fails)[13].

An exercise philosophers sometimes engage in consists precisely in assessing degrees of strength. Laws of nature, for example, are typically taken to support counterfactuals: some philosophers think that, though strong, they are weaker than, say, the laws of logic; others disagree and claim that the former are as strong as the latter, even that they are as strong as they can be, i.e. that they are necessary. Analogous interesting and hard questions can be raised about brute moral truths: are they as strong as, say, the laws of nature? If the latter are not necessary, are the former stronger than the latter? Are they necessary? As I say, these are interesting

---

[13] As I say in the main text, this characterization of the relation 'being stronger than' is rough-and-ready and is not meant to answer all the questions one would need to answer in order to provide a full account of it (to mention only one unanswered question: what if –as is sometimes claimed in discussions about counterfactuals- which similarity metric to adopt is a context-dependent matter?). A rough-and-ready characterization is all that is needed for my purposes here.

questions, and some friends of the Assumption do take a stand on them[14]. However, the Assumption and the four theses that articulate it can be endorsed, and used to develop an accommodationist proposal, even without taking a stand on such hard questions, so given my purposes here I shall not build any specific answer to them into thesis iv): I'll leave thesis iv) neutral with respect to these questions and use the symbol ' ■ ' for the degree of strength brute moral truths have, whatever exactly it turns out to be. In what follows, I'll sometimes need to keep track of sentences that support counterfactuals and have the degree of strength in question: I'll do so by prefacing them with ' ■ ' (sentences not prefaced with ' ■ ' may or may not be strong and, if strong, they may or may not be appropriately strong). The one thing I'll assume is (what will strike most as obvious, and is, so far as I can see, taken for granted by friends of the Assumption) that brute moral truths are at most as strong as logical truths, so that if it is logically true that p, then ■ p.

Thesis iv) explicitly mentions something I passed over in my rough formulation of the Assumption in section 1. That rough formulation went as follows: for any moral feature M-ness, if something is M, it has some non-moral property that makes it M; the only thing said about M-making properties was that they are non-moral properties. Thesis iv), on the other hand, says something more: according to it, M-making properties are non-moral properties *that can be denoted by non-moral predicates*[15]. The reason why in section 1 I framed the Assumption the way I did, instead of saying '… it has some non-moral property *which can be denoted by a non-moral predicate* that…', is that my purpose there was to give merely a rough and intuitive

---

[14] For an example of a friend of the Assumption who claims that (what I am here calling) brute moral truths are necessary, see Thomson [1990], introductory chapter.

[15] Thesis iv) does not say that M-making properties *are* denoted by non-moral predicates, i.e. that, for any M-making property, we already have a non-moral predicate denoting it, but it does say that M-making properties *can* be denoted by non-moral predicates, i.e. that, for any such property, we either already have or can come up with a non-moral predicate denoting it.

characterization of it, and of the way in which it might be used to develop an accommodationist proposal, and for that purpose the simpler formulation was enough. But, it might be asked, why not take that simpler formulation to capture everything the Assumption says about M-making properties? Why formulate thesis iv) so as to add the further claim that M-making properties can be denoted by non-moral predicates? After all, the reason why I am discussing the Assumption is that it is the tool opponents of moral explanations avail themselves of in order to develop accommodationist proposals. Now, the way this tool is meant to help them achieve that aim is the following: if one has the Assumption, one has the (non-empty set of) M-making properties, so one can propose that the content, or payoff, of 'M-ness causes E' is that the M-making properties cause E. Isn't it enough, for opponents of moral explanations to achieve their aim, that the M-making properties be non-moral properties, whether or not they can be denoted by non-moral predicates?

In formulating thesis iv) the way I did, I take it that I am being faithful to the intentions of opponents of moral explanations who rely on the Assumption in order to develop an accommodationist proposal (and, I would say, of friends of the Assumption generally), who do assume that M-making properties are not merely non-moral properties, but also properties that can be denoted by non-moral predicates. *Should* they assume this much? I think there are quite good reasons why they do.

Qua opponents of moral explanations, they deny that moral properties are causally efficacious. *Qua* accommodationist philosophers, they deny that any serious mistake is involved in our practice of offering, discussing, and accepting moral explanatory claims of the form 'M-ness causes E', on the ground that the content, or payoff, of a claim like this is captured by 'The M-making properties cause E'. True, such a sentence also has explanatory content, i.e. it ascribes

28

causal efficacy to certain properties[16]; the point, though, is that these properties, the M-making properties, are non-moral. So far, so good. However, coupled with the claim that (at least some of) the M-making properties cannot be denoted by any non-moral predicate, such a proposal results in an unstable position.

Consider what someone adopting this position would be claiming. Does the moral property of being M produce any causal effects? No, she claims. Is there anything seriously mistaken, then, with our practice of offering, discussing, and accepting claims of the form 'M-ness causes E'? No, she says, because the issues we get to discuss by engaging in such practice, and the claims we get to accept, are causal issues and claims that have nothing objectionable about them. By engaging in such practice we do manage to lock onto causal links involving certain causally efficacious properties, she grants: only, she adds, the causal efficacy at play is not that of the moral property of being M, but that of the M-making properties, which are non-moral properties. And what properties are they? Well, she says, we cannot say: they (at least some of them) cannot be denoted by any non-moral predicate (and of course, being non-moral properties, they cannot be denoted by any moral predicate either), hence they cannot be denoted by any predicate at all. Which, by the way, means that we cannot track their causal efficacy by conducting our discussions in purely non-moral terms: we cannot factor out the moral terminology embedded in our practice and replace such practice with one that keeps track of the causal links in question in purely non-moral terms, capturing them by means of purely non-moral explanatory claims.

So what she claims is: by engaging in the practice of discussing and accepting moral explanatory claims of the form 'M-ness causes E' we do manage to lock onto causal links

---

[16] As we shall see, on certain accommodationist proposals this is not the only thing such a sentence does, but on all of them this is one of the things it does: on all of them, the explanatory content is present (whether or not content of some different sort is also present).

involving certain causally efficacious properties; engaging in such practice is the only way in which we can do this, i.e. if we stick to purely non-moral terminology we cannot keep track of the causal links in question; still, the causal efficacy at play here is *not* that of the moral property of being M, but that of certain *other*, non-moral properties.

Someone adopting this position would have a hard time answering the following question: if, by discussing and accepting claims of the form 'M-ness causes E', we really manage to lock onto causal links that are there to be tracked (and that cannot be tracked using purely non-moral terminology), then why deny that the causally efficacious property at play is the moral property of being M? Conversely: if (as opponents of moral explanations claim) there really are good reasons to believe that moral properties are not causally efficacious, then why believe that, by engaging in our practice, we do manage to lock onto causal links that are there to be tracked? Why not conclude that there is something seriously mistaken with our practice of discussing and accepting claims of the form 'M-ness causes E'? The claim that (at least some of) the M-making properties cannot be denoted by any non-moral predicate thus creates a strong tension between the two tenets of an opponent of moral explanations who wants to go accommodationist.

These considerations explain why opponents of moral explanations who rely on the Assumption in order to develop their accommodationist proposals had better assume that M-making properties can be denoted by non-moral predicates, and thus why it is fair to formulate thesis iv) the way I did. And that opponents of moral explanations who go accommodationist are indeed (at least implicitly) sensitive to the considerations adduced above is, I think, shown by the fact that it is entirely standard for them to assume that we have, or can at least come up with, non-moral predicates denoting the non-moral properties whose causal efficacy, according to them, we manage to track by engaging in the practice of discussing and accepting moral

explanatory claims. These philosophers typically assume that we *can* use purely non-moral vocabulary from disciplines other than ethics (say, psychology, social theory, etc.) or maybe just from common sense discourse and (as I put it above) factor out the moral terminology embedded in our practice and replace such practice with one that keeps track of the causal links in question in purely non-moral terms, capturing them by means of purely non-moral explanatory claims[17].

Thesis iv) explicitly says that M-making properties can be denoted by non-moral predicates. It remains silent, however, on the logical complexity of such predicates: so far as iv) goes, the non-moral predicates in question can be of any logical complexity. In what follows, I'll often use 'N' (with or without a subscript or superscript) for non-moral predicates, including

---

[17] Though all agree there is a difference between (*a*) 'M-making properties are non-moral properties' and (*b*) 'M-making properties are non-moral properties that can be denoted by non-moral predicates', it has sometimes been suggested to me in conversation that the gap between the two is not one that it is particularly hard to bridge. I myself am less certain that the gap between (*a*) and (*b*) can be bridged so easily as is sometimes suggested; however, my point in the text is only that (*a*) and (*b*) are different and there are interesting and quite good reasons why opponents of moral explanations who want to go accommodationist had better articulate the Assumption along the lines of (*b*), whether or not this is something they can do without incurring any serious costs. There is, I believe, at least one other reason why opponents of moral explanations who rely on the Assumption in order to develop an accommodationist proposal assume that M-making properties can be denoted by non-moral predicates (though one that I am not sure would stand in the way of ridding the Assumption of such a claim while still being able to use it for accommodationist purposes). This other reason actually applies, I think, to friends of the Assumption quite generally, and is easier to see if one looks at things from a historical perspective. From such a perspective, it is useful to think of the Assumption as the combination of two other widely shared views. The first view I have in mind can be formulated (on the model of my rough formulation of the Assumption in section 1) as follows: for any moral feature M-ness, if something is M, it has some property which can be denoted by a non-moral predicate that makes it M (fleshing out this rough formulation along the lines of theses i)-iv) above would require solving some difficult problems, but the rough formulation is enough for my purposes here). The thing to note about this view is that the only thing it says about M-making properties is that they can be denoted by non-moral predicates: it remains silent on whether they are non-moral properties and is therefore compatible with the view that (there is only one M-making property and) the M-making property is identical to the property of being M. This view is thus compatible with (ethical) reductionism i.e. the view that moral properties are identical to properties that can be denoted by non-moral predicates. It is, I think, plausible to read Moore [1903] and Hare [1952] (among others) as starting from this view, presented as something they shared with their reductionist (or, as they actually called her –assuming that ethical naturalism entails reductionism-, naturalist) opponent, and then arguing that their opponent's account of it, according to which (there is only one M-making property and) the non-morally denotable M-making property just is the property of being M, is mistaken. The second view I have in mind is precisely the rejection of such an account and of reductionism quite generally: if a property can be denoted by a non-moral predicate, then it is a non-moral property. This is a view that (thanks again to the influence of Moore [1903] and Hare [1952], among others) was extremely popular for most of the 20th-century and is still very widely accepted. Taken together, these two views yield the Assumption, but note: the claim they yield is not merely that, for any moral feature M-ness, if something is M, it has some non-moral property that makes it M, but more specifically that, for any moral feature M-ness, if something is M, it has some non-moral property *which can be denoted by a non-moral predicate* that makes it M.

non-moral predicates denoting M-making properties, but this should not be taken to imply that the predicates in question are logically simple[18].

---

[18] This is a good point at which to mention something that a careful reader might have noted already. If 'N*' is a non-moral predicate denoting a non-moral property, so is also 'N*∧~N*'. Let us abbreviate the latter expression as 'N' and consider the sentence '■ All N's are M', with 'M' a moral predicate. This is an appropriately strong universal moral sentence using a non-moral predicate denoting a non-moral property. And, many would say, it is true. Does it express a brute moral truth? If it does, then being N is an M-making property. And so are, of course, an awful lot of other properties (just replace 'N*' with any other non-moral predicate denoting a non-moral property and you get other examples; moreover, once one knows the trick, it is easy enough to generate other sorts of examples). I take it that this is a counterintuitive and implausible result and that a friend of the Assumption will want to avoid it, so let us try saying that '■ All N's are M' expresses a derived moral truth. What, though, are the other categorical moral truths that, on i), make it true, together with appropriate non-moral truths? It seems that, if we say that '■ All N's are M' is true, then the right thing to say in reply to 'What makes it true?' is just '■ Nothing is N', and the latter sentence is not a categorical moral sentence. (It might be argued that, by means of some technical tricks, the truth expressed by the latter sentence could also be expressed by a sentence that fits my definition of 'categorical moral sentence'. Even granting this can be done, all the ways of doing it that I am aware of still generate problems, so far as I can see, not for i) but for iii). I'll spare the reader the technical details.) The upshot of all this is that, if we say that '■ All N's are M' is true and we do not want to say it expresses a brute moral truth, then we have to conclude it expresses a categorical moral truth of which i) (or iii)) is false. Though this shows there is a problem with articulating the picture implicit in the Assumption by means of i)-iv) as stated, the problem is one that is going to pop up, so far as I can see, on any plausible way of articulating that picture in some detail. I expect there are different possible ways of dealing with it; my favorite one goes as follows. Let us say that a categorical moral truth to the effect that all P's are M that is true only because it is made true by a truth to the effect that nothing is P is a degenerate categorical moral truth; then theses i)-iv) should be re-framed so as to apply not to all categorical moral truths, but only to non-degenerate categorical moral truths (this adds a further restriction on the scope of i)-iv), in addition to that pointed out in n. 7). This seems to me to capture what is intuitively the right thing to say about the difficulty at issue: in order to raise the difficulty, one needs to turn to cheap (so to speak) and not very interesting categorical moral truths, whereas the Assumption is meant to be a claim with an interesting philosophical content, one taking a stand on meatier and more interesting categorical moral truths. Having noted the problem that degenerate categorical moral truths raise for i)-iv) as stated, I'll ignore them throughout the present work and stick to the formulation of i)-iv) given in the main text. It will be noted that, in presenting my favorite way of dealing with this difficulty, I did not assume either that the problematic moral truth is strong, or (supposing it is) that it is appropriately strong, or that such truth involves a non-moral property that can be denoted by a non-moral predicate. This is so because, though I presented the difficulty by means of a moral truth of which these three things are true, the difficulty is entirely general and can be raised by means of any universal moral truth (of course, in the case of a universal moral truth that does not meet these three conditions, the reason why it cannot be a brute moral truth is that it does not satisfy iv)). It will also be noted that, on my way of dealing with this difficulty, it is not merely truths involving contradictory properties that count as degenerate. For instance, suppose that I never had and will never have any children: the sentence 'All my children are good people' (as uttered by me) expresses a degenerate categorical moral truth, even if the property of being a child of mine is not a contradictory property. This too, I think, is as it should be.

Let me quickly add two further points: though I am inclined to accept both, I advance them tentatively and the solution presented above does not rely on them. First point: in order to raise the difficulty, one needs to assume that, in saying that a categorical moral truth to the effect that all P's are M is made true by a truth to the effect that nothing is P, the relation 'being made true by' we appeal to is the same as the relation 'being made true by' in terms of which theses i)-iv) are formulated (absent this assumption, the reasoning above would fail to show that degenerate categorical moral truths provide a counterexample to i)-iv) as stated). I would be inclined to deny this assumption: the former relation does not, I would be inclined to say, carry the same metaphysical weight as the latter does. And the lesson I would be inclined to draw from this is that the sense of 'truth' in which degenerate categorical moral truths count as truths is not the same as that in which i)-iv) are theses about categorical moral truths, so that degenerate categorical moral truths do not actually provide a counterexample to i)-iv) even as stated, it

What exactly does thesis iv) add to theses ii) and iii)? Theses ii) and iii) already entail that the paths that can be tracked, starting from derived moral truths and asking, about each of them, 'What makes it true?', end in brute moral truths that are universal and involve non-moral properties. Thesis iv) adds that such brute moral truths are strong (with the appropriate degree of strength) and involve non-moral properties that can be denoted by non-moral predicates. Thesis iv), however, adds more than just this much, since it is a thesis about all brute moral truths, whereas whatever ii) and iii) entail about brute moral truths, they entail it only about the brute moral truths that lie at the end of the paths originating from derived moral truths. So, for example, iv) rules out (what is not ruled out by ii) and iii)), that there be individual brute moral truths. And ruling this out is essential to capturing the picture implicit in the Assumption. To see why, suppose there were an individual brute moral truth, say 'a is M': even if this were the only such truth among those ascribing M-ness, the non-moral properties picked out by the (other) brute truths ascribing M-ness could no longer be thought of (as I put it above) as the ultimate source of all the M-ness instantiated in the world, since there would be something, i.e. a, that is M but whose M-ness is not ultimately due to any of those non-moral properties, actually to any property at all. But the idea at the heart of the picture implicit in the Assumption is precisely that the instantiation of M-ness is *always* mediated (so to speak) by the instantiation of other

---

being an equivocation on 'truth' which makes it seem that they do. As I said, the solution presented above does not rely on this point, so whether or not one agrees with it, one knows how to deal with degenerate categorical moral truths. The second point is that, on my definition of 'degenerate categorical moral truth', in order for a categorical moral truth to the effect that all P's are M to count as degenerate it is not sufficient that nothing be P -and so that the former truth be made true by the latter-: it is further required that the former truth be true only because it is made true by the latter. This further requirement, I would be inclined to say, is not trivial. The reason is easier to see if one considers an example from outside ethics. Consider sentences of the form 'All frictionless objects with such-and-such features in such-and-such conditions behave in such-and-such a way'. It seems right to say that, on our current physics, some of these sentences are true, but I would be inclined to deny that the whole story about what makes them true is that there are no frictionless objects; and, absent specific reasons to the contrary, I am inclined to leave it as an open possibility that cases fitting this pattern occur within ethics as well (after all, don't we sometimes discuss of the moral value of actions/people/institutions/etc. with characteristics that no action/person/institution/etc. has?). Once again, I mention this point only for completeness: the solution presented above to the difficulty discussed in this footnote does not rely on it.

properties, or that things' instantiating M-ness is always ultimately due to their instantiating some other, non-moral properties, so that the set of such non-moral properties *can* be thought of as the ultimate source of all the M-ness instantiated in the world, and thus gets to count as the set of the M-making properties. Thesis iv) is essential to capturing this idea.

A rough and ready slogan to summarize theses i)-iv) could thus be 'No ethics without principles'. That is: if 'there is an ethics', i.e. if there is any categorical moral truth, then 'there are principles', i.e. there are categorical brute moral truths with the features that iv) and -if the categorical moral truths there are include derived ones- ii) and iii) assign to brute moral truths. One aspect that is usefully pointed out by this way of summarizing theses i)-iv) is that they amount to a conditional claim and as such are silent as to whether the antecedent is true or not. Theses i)-iv) are meant to be neutral with regard to whether there are any categorical moral truths: all they say is that *if* there are any, they have certain characteristics (categorical moral truths, if there are any, are either derived or brute; derived moral truths, if there are any, are ultimately made true by brute moral truths, plus appropriate non-moral truths; and so on), so that *if* there are any categorical moral truths, then there are brute moral truths with the features assigned to them by iv) and -if the categorical moral truths there are include derived ones- ii) and iii). Said otherwise: the point of i)-iv) is not to assert that 'there is an ethics', it is rather to lay out a certain condition and assert that, unless that condition is met (unless 'there are principles'), 'there is no ethics'. As a matter of fact, friends of the Assumption do typically believe that 'there is an ethics'[19]; this, however, is an additional claim that goes beyond the Assumption, and one could in principle accept the Assumption while rejecting this claim. Indeed, one could in principle rely on the Assumption in order to reject this claim, arguing as follows: if there are any

---

[19] And, as I have already remarked in n. 3, opponents of moral explanations who want to accommodate our practice with regard to moral explanatory claims are entitled to assume, in the context of the present discussion, that our moral practice is not, as a whole, seriously misguided, hence that 'there is an ethics'.

categorical moral truths, then there are brute moral truths with the features described in iv) and - depending on the categorical moral truths there are- ii) and iii); but 'there are no principles', i.e. there are no such moral truths (say, because there are no true appropriately strong universal moral sentences using non-moral predicates: they all have exceptions); so 'there is no ethics', i.e. there are no categorical moral truths, and nothing is right or wrong, good or bad, etc[20].

Having just highlighted one respect in which theses i)-iv) are weaker than one might have thought, let me end my series of clarificatory remarks about them by pointing to a different respect in which they are stronger than it might at first appear. The Assumption is meant to capture an aspect of the very nature of morality: the idea is meant to be that the nature of moral properties is such that, if there are categorical moral truths, they have the features described in i)-iv). In terms of the instantiation of moral properties: whether or not moral properties are instantiated, their nature is such that they can only be instantiated in such a way as to conform to the pattern described in i)-iv). Theses i)-iv) are thus meant to be more than a mere description of what is as a matter of fact the case: for instance, what thesis ii) says is not merely that derived moral truths (if there are any) are as a matter of fact all ultimately made true by brute moral truths, plus appropriate non-moral truths, but that derived moral truths (if there are any) cannot but be all ultimately made true by brute moral truths, plus appropriate non-moral truths; and the

---

[20] Moral particularists go the opposite way: as I mentioned in n. 4, they are among those raising doubts about the Assumption, since they argue that 'there are no principles', but typically take this to show that ethics does not need principles, not that 'there is no ethics'. Note that what particularists usually object to is the Assumption, not the appropriateness of our asking, about something claimed to be M, 'What is M about it?': that is, they usually find nothing objectionable about the feature of our moral practice that, as I noted above, the Assumption is meant to give an account of; what they find objectionable is accounting for that feature in terms of the Assumption, thereby making principles necessary for anything to be right or wrong, good or bad, etc. (I'll have occasion to refer to particularism at various points in what follows. In the present work, I'll define moral particularism as the denial of all appropriately strong universal moral sentences using non-moral predicates. Though this is not, strictly speaking, a fully accurate characterization of the particularist position, it is a good first approximation, which is all that is needed for my purposes here.)

same goes for the other theses. In terms of our slogan: there *can* be no ethics without principles, or if there is an ethics, there *must* be principles[21].

3. <u>The 'Go by the M-making properties' strategy</u>

Let us now take a closer look at the way in which theses i)-iv) can be used in order to develop an accommodationist proposal. Remember what the problem is that such a proposal is meant to solve. On the face of it, a moral explanatory claim of the form 'M-ness causes E' (where 'M' is a moral predicate) ascribes causal efficacy to the property of being M. What an opponent of moral explanations who wants to go accommodationist needs is a sentence ascribing no causal power to the property of being M (or any other moral property) that she can propose as capturing the content, or payoff, of such a claim. The shortest, most conservative way for her to go would be to construct such a sentence out of the original moral explanatory claim, by replacing the latter claim's apparent reference to the moral property of being M with a reference to non-moral

---

[21] This point, together with the point that the 'principles' in question are strong (with the appropriate degree of strength), *might* ease a possible difficulty I pointed out in n. 3 for the accommodationist project, arising from the need to assume (in addition to the Assumption) that at least one thing is M (where M-ness is the moral feature mentioned in a moral explanatory claim one aims at accommodating). It might be argued that, given the two points I have just mentioned, what one needs to assume (in addition to the Assumption) in order to make sure that the set of M-making properties is not empty, is merely that something can be M, not that something is M. Intuitively, the idea would be to observe that, on the Assumption, for any moral feature M-ness, if something were M, it would have some non-moral property that makes it M, and argue that, provided that something can be M, then even if nothing is M, the non-emptiness of the set of M-making properties is still guaranteed, at least under certain conditions (just think of the M-making properties as the non-moral properties such that, if something were M, it would be made M by at least one of them). Developing such an argument with some care would require addressing some interesting questions (for instance, about the scope of the counterfactuals involved). Supposing such an argument can be developed, the upshot for the proposal that the content, or payoff, of 'M-ness causes E' is that the M-making properties cause E, as applied to a case in which nothing is M but something can be M (and the relevant conditions are met), would be that, in discussing and accepting the moral explanatory claim, the causal claim we get to discuss and accept is one about the causal power of the non-moral properties that, if there were M things, would make them M. Presumably, since nothing is M, these non-moral properties are not instantiated, so in order to show that the proposal in question really counts as an accommodationist proposal, its friends would still want to tell us why no serious mistake is embedded in discussing the causal power of uninstantiated properties, but they might have plausible things to say about this (for instance, they can point out that discussing the causal power of uninstantiated properties does seem to be part of current empirical science, as when physicists tell us how different kinds of matter would behave if, say, frictionless or at a temperature of absolute zero). As I said in n. 3, I think this possible difficulty for the accommodationist project raises interesting and complex issues, but I won't explore it in what follows.

properties. If theses i)-iv) are true, then there is a plausible way of implementing precisely this strategy, one that could be summarized by the slogan 'Go by the M-making properties': replace the moral explanatory claim's apparent reference to the (moral) property of being M with a reference to the (non-moral) M-making properties. This should strike one as plausible because, on i)-iv), the M-making properties can be thought of (as I put it above) as the ultimate source of all the M-ness instantiated in the world, so if one is looking for properties to hold responsible for a job that, though apparently performed by M-ness, one is not willing to assign to it, the M-making properties are an obvious candidate.

In order to see what the resulting accommodationist proposal looks like in detail, let us start by making a simplifying assumption (we'll eventually have to drop this assumption, but for the time being it is easier to work with it). As I noted in the previous section, theses i)-iv) are neutral as to the number of brute moral truths there are for a given moral property M-ness. No matter what their number is, these truths can all be expressed by an appropriately strong sentence of the form 'All N's are M', for some non-moral predicate 'N' denoting a non-moral property. Let us assume that there is only one brute moral truth ascribing M-ness, hence that there exists *the* non-moral property that can be denoted by some non-moral predicate 'N' such that ■ all N's are M -for brevity's sake, I'll shorten this definite description to 'the *n*'-; consider now the non-moral predicate, whatever it will turn out to be, denoting such property: suppose it is 'N$_1$'[22].

---

[22] In the definite description 'the non-moral property that can be denoted by some non-moral predicate 'N' such that ■ all N's are M', 'N' is a bound variable. Under i)-iv), we can think of substantive moral inquiry as aiming at figuring out what non-moral predicate to substitute for 'N' in ' ■ all N's are M' so as to get a sentence expressing the only brute moral truth ascribing M-ness we are assuming there is, with one moral theory proposing 'N$_1$', i.e. that N$_1$-ness is the *n*, another proposing 'N$_2$', i.e. that N$_2$-ness is the *n*, and so on. For example: if M-ness is goodness, then one theory is that we should substitute 'N' with 'pleasant', i.e. that being pleasant is the *n*, a different one is that we should substitute 'N' with 'desired', i.e. that being desired is the *n*, and so on. The accommodationist strategy I am describing is neutral as to which of such substantive proposals is correct: in order to formulate it, I am supposing that the one proposing to substitute 'N' with 'N$_1$' is correct; if a different one is correct, then the relevant predicate should be replaced for 'N$_1$' throughout my formulations (and if the correct substantive theory is such that what to

There are three slightly different ways of developing the accommodationist strategy described above. The first way is the following:

(A) The content, or payoff, of 'M-ness causes E' is that N_1-ness causes E[23].

That is, replace the moral expression in the moral explanatory claim with the non-moral **noun referring to the n**: call this operation 'simple non-enriched singular substitution'. As intended, on (A) the sentence capturing the content, or payoff, of a moral explanatory claim also has explanatory content, i.e. it ascribes causal efficacy to certain properties, but the properties in question are non-moral properties.

One noteworthy feature of (A) is that what is true of the sentence it offers as capturing the content, or payoff, of a moral explanatory claim is not merely that (as intended) it ascribes causal efficacy to non-moral properties, but also that it lacks moral content altogether: on (A), the content, or payoff, of a moral explanatory claim is entirely non-moral. Suppose you find this disturbing: suppose you think that, no matter what else is true of it, the sentence capturing the content, or payoff, of a moral explanatory claim must have moral content. Here is a different way of developing the accommodationist strategy described above that might be more appealing to you:

---

substitute for 'N' is not a predicate we already have but one that needs to be introduced into our language, then that predicate –supposing we do introduce it into our language- should be replaced for 'N_1' throughout my formulations).
[23] Supposing that M-ness is goodness and being pleasant is the n: the content, or payoff, of 'Goodness causes E' is that being pleasant causes E. Since the proposals I'll describe are proposals about the content, or payoff, of a certain class of claims, they usually take a stand on how to deal with the negations of the relevant claims as well. In the case of (A), how to deal with the negation of a moral explanatory claim is pretty straightforward: just take the negation of the sentence capturing the content, or payoff, of the moral explanatory claim, i.e. the content, or payoff, of 'M-ness does not cause E' is that N_1-ness does not cause E.

(B) The content, or payoff, of 'M-ness causes (does not cause) E' is that N_1-ness causes (does not cause) E and N_1-ness is the n[24].

That is, replace the moral expression in the moral explanatory claim (or its negation) with the non-moral **noun referring to the n and** add a clause expressing the moral relevance of the property in question: call this operation 'simple enriched singular substitution'. On (B), the sentence capturing the content, or payoff, of a moral explanatory claim does have moral content. The important point, for the purposes of an opponent of moral explanations who wants to go accommodationist, is that this content does not mesh (so to speak) in an undesired way with the explanatory content that sentence also has. On (B), the content, or payoff, of 'M-ness causes E' can be factored into two quite distinct components: an explanatory, entirely non-moral component, and a separate purely moral, non-explanatory component. On (B), just as much as on (A), no ascription of causal efficacy is made to any moral property.

Still a different way of developing the accommodationist strategy described above goes as follows:

(C) The content, or payoff, of 'M-ness causes E' is that the n causes E[25].

That is, replace the moral expression in the moral explanatory claim with the **definite description 'the n'**: call this operation 'simple general substitution'. The fact that the replacing

---

[24] In terms of my previous example: the content, or payoff, of 'Goodness causes (does not cause) E' is that being pleasant causes (does not cause) E and being pleasant is the non-moral property that can be denoted by some non-moral predicate 'N' such that ■ all N's are good. In framing (B) this way, I am assuming that the definite description in the sentence 'N_1-ness is the n' is treated *à la* Russell. This assumption is merely for ease of exposition (so that I can avoid having to write down the Russellian *analysans* every time I need to state (B)) and it is not meant to play any substantive argumentative role.

[25] Using my previous example: the content, or payoff, of 'Goodness causes E' is that the non-moral property that can be denoted by some non-moral predicate 'N' such that ■ all N's are good causes E.

expression still contains moral terminology is of course no problem for the purposes of an opponent of moral explanations who wants to go accommodationist: this fact does mean that on (C), just as much as on (B), the content, or payoff, of a moral explanatory claim has a moral component; however, just as on (B), this component does not mesh in an undesired way with the explanatory component. This is even easier to see if one paraphrases the sentence offered by (C) as capturing the content, or payoff, of 'M-ness causes E' as follows: there is a property that *both* is the non-moral property that can be denoted by some non-moral predicate 'N' such that ∎ all N's are M *and* causes E. On (C), just as much as on (A) or (B), no ascription of causal efficacy is made to any moral property[26]. The sentences offered by (B) and (C) as capturing the content, or payoff, of 'M-ness causes E' differ from one another in the same way as 'Obama is black and Obama is the president of the U.S.' and 'The president of the U.S. is black' differ from one another[27]. One way to point out the difference is by noting that the latter sentence expresses what is, on anyone's view, a general proposition, whereas the former does not. A different way comes out if we suppose that brute moral truths could have been different from what they actually are, so that even if (as we are supposing) N_1-ness is the *n*, a different non-moral property, say N_2-ness, could have been the *n*. Then the sentence offered by (C) as capturing the content, or payoff,

---

[26] As above, in framing (C) as I did I am assuming that the definite description in the sentence 'The *n* causes E' is treated *à la* Russell. Once again, this assumption is merely for ease of exposition and it is not meant to play any substantive argumentative role (for brevity's sake, I'll sometimes shorten the Russellian *analysans* framing it, as I have just done in the main text, as an existential claim that still makes use of the definite description in question). Note that if one starts from the sentence 'The *n* causes E' as capturing the content, or payoff, of a moral explanatory claim and plugs in a reading of the definite description that is closer to Strawson than to Russell, one gets (A) (I am simplifying Strawson's actual views a bit, bracketing complications that are irrelevant to my purposes here). (B) too can be seen as the result of starting from 'The *n* causes E' as capturing the content, or payoff, of a moral explanatory claim and plugging in still a different reading of the definite description, one somewhat in-between Russell's and Strawson's readings. Quite generally, in thinking about accommodationist proposals it is useful to keep debates about the theory of definite descriptions in mind (in doing that, one should also keep in mind that the referent or, in Russellian terms, denotation of 'the *n*' is not, as is usually the case with examples of definite descriptions used in those debates, an individual, but a property). How does (C) deal with the negation of a moral explanatory claim? I'll leave (C) neutral as to whether the content, or payoff, of 'M-ness does not cause E' is captured by the external or the internal negation of 'The *n* causes E': someone endorsing (C) can go either way.

[27] Where 'the president of the U.S.' is, just like 'the *n*', treated *à la* Russell.

of a moral explanatory claim is entailed by, but does not entail, the sentence offered by (B), and so it could in principle be true even if the latter sentence is false: had N_2-ness been the n, the sentence offered by (B) would have been false, but the sentence offered by (C) could still have been true, depending on whether N_2-ness did or did not cause E[28].

For purposes that will become clear starting from the next chapter, it is useful to think of accommodationist proposals as particular cases of a more general recipe that can be applied to other moral sentences as well, in addition to moral explanatory claims. To illustrate, let us consider individual moral sentences, i.e. sentences of the form 'x is M', where 'M' is a moral predicate and 'x' is a singular term. Suppose you want to replace an individual moral sentence's apparent reference to the moral property of being M with a reference to non-moral properties[29]: once again, given i)-iv), it should strike you as a plausible strategy to 'go by the M-making properties', since such properties can be thought of as the ultimate source of all the M-ness instantiated in the world. Then you can revise in obvious ways the definitions given above of 'simple non-enriched singular substitution', 'simple enriched singular substitution', and 'simple general substitution' so as to define these operations for the case of individual moral sentences as well, and get the following three proposals:

(a) The content, or payoff, of 'x is M' is that x is N_1.

(b) The content, or payoff, of 'x is (is not) M' is that x is (is not) N_1 and N_1-ness is the n.

---

[28] This second way of pointing out the difference between (B) and (C) will not work if brute moral truths are necessary (a possibility I am leaving open). If they are, then the sentence offered by (C) does entail the one offered by (B): there is no possible world in which the former is true but the latter is not. The other way of pointing out the difference still works: for comparison, think of '1 is an odd number and 1 is the positive number that is its own square' and 'The positive number that is its own square is an odd number' (where 'the positive number that is its own square' is treated *à la* Russell).
[29] Why would you want to do that? Never mind why for now: we'll look at a reason why you might in the next chapter (see section 3.2).

(c) The content, or payoff, of 'x is M' is that x has the n[30].

It is time now to drop our simplifying assumption and consider the case in which there is more than just one brute moral truth ascribing M-ness, hence more than just one non-moral property that can be denoted by some non-moral predicate 'N' such that ■ all N's are M. For simplicity, I'll consider the case in which there are two such truths: it will be easy to extrapolate from what I say to cases in which there are more than two such truths. Let us then use 'the n's' as shorthand for 'the non-moral properties that can be denoted by some non-moral predicate 'N' such that ■ all N's are M' and suppose that the n's are two; consider now the non-moral predicates, whatever they will turn out to be, denoting such properties: suppose they are 'N$_1$' and 'N$_2$'. One way of starting from the definitions given above and defining new operations (call them 'multiple non-enriched singular substitution', 'multiple enriched singular substitution', and 'multiple general substitution') for the case in which the n's are two is obvious enough. If we start with individual moral sentences, we get the following three proposals:

(a') The content, or payoff, of 'x is M' is that x is N$_1$ or x is N$_2$.

(b') The content, or payoff, of 'x is M' is that (x is N$_1$ or x is N$_2$) and N$_1$-ness and N$_2$-ness are the n's; the content, or payoff, of 'x is not M' is that x is not N$_1$ and x is not N$_2$ and N$_1$-ness and N$_2$-ness are the n's.

(c') The content, or payoff, of 'x is M' is that x has (at least) one of the n's.

---

[30] As above, I am assuming that the definite description in the sentences 'N$_1$-ness is the n' and 'x has the n' is treated *à la* Russell and I am leaving (c) neutral as to whether the content, or payoff, of 'x is not M' is captured by the external or the internal negation of 'x has the n'; (a) deals with negations in the same straightforward way as (A) does (as do the forthcoming (A'), (a'), (A''), and (a'')).

Turning now to moral explanatory claims, what we get is:

(A') The content, or payoff, of 'M-ness causes E' is that N_1-ness causes E and N_2-ness causes E.

(B') The content, or payoff, of 'M-ness causes E' is that N_1-ness causes E and N_2-ness causes E and N_1-ness and N_2-ness are the *n*'s; the content, or payoff, of 'M-ness does not cause E' is that (N_1-ness does not cause E or N_2-ness does not cause E) and N_1-ness and N_2-ness are the *n*'s.

(C') The content, or payoff, of 'M-ness causes E' is that (each of) the *n*'s cause(s) E[31].

It is standard among philosophers to assume that there is also another way of extending the 'Go by the M-making properties' strategy to the case in which the *n*'s are two, one that appeals to the disjunctive property built out of the *n*'s and so 'goes', strictly speaking, not 'by the

---

[31] Three remarks. One: note the difference between individual moral sentences and moral explanatory claims, which shows up in each of the three different types of proposal. In the case of an individual moral sentence, trying to capture its content, or payoff, by saying that x instantiates both of the *n*'s would be intuitively too strong: the appropriate thing to say is that x instantiates (at least) one of them. After all, that x instantiates only one of the *n*'s, but not the other, is a reason for accepting, rather than rejecting, 'x is M'. In the case of a moral explanatory claim, on the other hand, trying to capture its content, or payoff, by saying that (at least) one of the *n*'s causes E would be intuitively too weak: the appropriate thing to say here is that both of the *n*'s cause E. After all, that only one of the *n*'s causes E, while the other does not, sounds like a reason for rejecting, rather than accepting, 'M-ness causes E' (consider for instance the claim 'Injustice causes revolutions': that only one of the forms of injustice causes revolutions, while the other does not, sounds like a reason for rejecting, rather than accepting, that claim). One of the things that can turn out to be wrong with a moral explanatory claim is precisely that it (so to speak) goes too general. Two: just as 'the *n*' in (B), (C), (b), and (c), 'the *n*'s' in (B'), (C'), (b'), and (c') (and in the forthcoming (B''), (C''), (b''), and (c'')) is meant to be read along Russellian lines. For example, the sentence 'The *n*'s cause E' in (C') is shorthand for 'There is a set of at least two properties such that: (1) each of the properties in the set is a non-moral property that can be denoted by some non-moral predicate 'N' such that ■ all N's are M; (2) for any property P, if P is a non-moral property that can be denoted by some non-moral predicate 'N' such that ■ all N's are M, then P is identical to one of the properties in the set; and (3) each of the properties in the set causes E'. Three: just as I did with (C) and (c), I am leaving (C') and (c') (and the forthcoming (C'') and (c'')) neutral as to whether to deal with negations in terms of external or internal negation. (Given the number of quantifiers occurring in 'The *n*'s cause E', what do I mean by its internal negation? Take the long sentence I have just written down that 'The *n*'s cause E' is shorthand for: what I mean is the sentence one gets by replacing (3) with its *external* negation, i.e. with 'not each of the properties in the set causes E', *not* with 'none of the properties in the set causes E'. Analogously with the internal negation of 'x has one of the *n*'s'.)

43

M-making properties' but by their disjunction (call the resulting operations '<u>disjunctive non-enriched singular substitution</u>', '<u>disjunctive enriched singular substitution</u>', and '<u>disjunctive general substitution</u>'). Starting again with individual moral sentences and turning then to moral explanatory claims, we get the following proposals:

(a'') The content, or payoff, of 'x is M' is that x has the disjunctive property of being $N_1 \lor N_2$.

(b'') The content, or payoff, of 'x is (is not) M' is that x has (does not have) the disjunctive property of being $N_1 \lor N_2$ and N_1-ness and N_2-ness are the n's.

(c'') The content, or payoff, of 'x is M' is that x has the disjunctive property built out of the n's.

(A'') The content, or payoff, of 'M-ness causes E' is that the disjunctive property of being $N_1 \lor N_2$ causes E.

(B'') The content, or payoff, of 'M-ness causes (does not cause) E' is that the disjunctive property of being $N_1 \lor N_2$ causes (does not cause) E and N_1-ness and N_2-ness are the n's.

(C'') The content, or payoff, of 'M-ness causes E' is that the disjunctive property built out of the n's causes E.

For ease of exposition, it will be useful to speak of the three operations of simple substitution defined above as limiting cases of the three corresponding operations of both multiple and disjunctive substitution. Speaking this way, let us take stock. No matter how many brute moral truths there are that ascribe M-ness, there are three slightly different ways of

developing an accommodationist proposal for moral explanatory claims out of theses i)-iv): the sentence to be put forward as capturing the content, or payoff, of a moral explanatory claim can be constructed out of such a claim *via* either general or singular substitution, which can in turn be either enriched or non-enriched; each of the three resulting proposals can be fleshed out in two different ways, depending on whether the substitution performed is multiple or disjunctive (with these two ways of fleshing out each proposal converging on the same result in the case in which there is only one brute moral truth ascribing M-ness)[32].

'No matter how many brute moral truths there are that ascribe M-ness', I have just said, even though I considered only the cases in which the number of such truths is either one or higher than one: what about the case in which that number is zero? On any of the above proposals, the content, or payoff, of 'M-ness causes E' is specified *via* reference to the *n*'s, so if there turn out to be no brute moral truths ascribing M-ness, the upshot is (depending on whether the proposal relies on singular or general substitution) either that 'M-ness causes E' has no content, or payoff, or that it, or the sentence capturing its payoff, is false: in any case, there turns out to be something wrong with it. Said otherwise (and borrowing some terminology from debates about the theory of definite descriptions): on any of the proposals in question, 'M-ness causes E', or the sentence capturing its payoff, *presupposes or entails* that there is at least one non-moral property that can be denoted by some non-moral predicate 'N' such that ∎ all N's are

---

[32] Accommodationist proposals along these lines have been put forward by Audi [1993] and Thomson [1998]. Cognitivist opponents of moral explanations are usually less emphatic than non-cognitivist ones about their ability to accommodate our practice with regard to moral explanatory claims and it can, at least in some cases, be somewhat hard to pinpoint the type of proposal they mean to endorse. Specifically, there can be disagreement about whether it is a proposal of type A or one of type C that they mean to put forward (for the record, I believe the latter is the one that best captures their intentions). No matter who endorses exactly what, it is in any case useful to distinguish and lay out all the different possibilities available to a cognitivist opponent of moral explanations who wants to go accommodationist. This also explains why I mention proposals of type B as well, which strike me as an interesting alternative, and the option of fleshing out each type of proposal in terms of disjunctive substitution, which is something neither Audi nor Thomson mentions (as will become clear as we proceed, having a general framework that includes the latter two points will also prove useful for the discussion to follow in the next few chapters).

M, so that, if there are no such properties, there turns out to be something wrong with 'M-ness causes E'. That this is the upshot of the above proposals in the case in which there are no brute moral truths ascribing M-ness should of course come as no surprise, given that the idea behind them is that the content, or payoff, of 'M-ness causes E' can be captured by 'going by the M-making properties': no wonder that, if there are no such properties, the upshot of that idea is that there is something wrong with 'M-ness causes E'[33].

The accommodationist proposals just described can be seen as the application to moral explanatory claims of a general recipe that can be applied to other moral sentences as well, as I have illustrated by means of individual moral sentences. As I have already remarked about moral explanatory claims, and as is true in the case of individual moral sentences as well, this general recipe specifies the content, or payoff, of the relevant moral sentence *via* reference to the *n*'s: on any of the above proposals, the sentence capturing the content, or payoff, of the relevant moral sentence makes reference to the *n*'s, either because it uses the description 'the *n*'s' or because it uses nouns referring to, or predicates denoting, the *n*'s (or their disjunction). Since 'the *n*'s ' is shorthand for 'the non-moral properties that can be denoted by some non-moral predicate 'N' such that ■ all N's are M', this general recipe can be seen as a procedure to specify the content, or payoff, of different classes of moral sentences by means of a specific class of such sentences, whose content, or payoff, is assumed as given, i.e. the class of appropriately strong universal moral sentences that can be formulated using a non-moral predicate denoting a non-moral

---

[33] Note that this does mean that the proposals fail to be *accommodationist* proposals, if there turn out to be no brute moral truths. Said otherwise, one *possible* argument against opponents of moral explanations who put forward these proposals in order to accommodate our practice with regard to moral explanatory claims goes as follows: these proposals fail to accommodate our practice, unless there are brute moral truths; moral particularism is true, i.e. there are no true appropriately strong universal moral sentences using non-moral predicates, so there are no brute moral truths; therefore, these proposals fail to accommodate our practice. One of the premises of this argument, however, is that particularism is true, and this is a premise that is quite difficult to establish; in any case, an adequate discussion of it would make for the topic of a different work. I will thus not explore this argument in what follows.

property[34]. What the procedure does is to start from this class of moral sentences and specify the content, or payoff, of other classes of moral sentences *via* reference to the non-moral properties picked out by the true ones among the former[35]. It is easy enough to see that this general recipe can be applied more broadly than I have illustrated so far. Let moral sentences <u>of the relevant type</u> be moral sentences that are either moral explanatory claims or categorical moral sentences[36]. Here is how one can apply the recipe in question to specify the content, or payoff, of all moral sentences of the relevant type starting from a subset of them, whose content, or payoff, is assumed as given:

1)   All moral sentences (of the relevant type) are either <u>built sentences</u>, i.e. their content, or payoff, is specified by means of other moral sentences (of the relevant type) or <u>basic sentences</u>, i.e. their content, or payoff, is not specified by means of other moral sentences (of the relevant type).

2)   The content, or payoff, of all built moral sentences is specified by means of basic moral sentences.

3)   The set of basic moral sentences is the set of appropriately strong universal moral sentences that can be formulated using a non-moral predicate denoting a non-moral property.

---

[34] I should remind the reader (see n. 15) that the modal expression 'sentences that *can* be formulated' covers both sentences that require, for their formulation, a non-moral predicate we already have and sentences that require a non-moral predicate we do not already have but can come up with.

[35] Or by those, among the former, that are such that the sentences capturing their payoff are true: for simplicity's sake, I'll omit this clause in what follows.

[36] This category includes all the moral sentences that are relevant to my discussion in the present work: this explains my choice of the label; it also explains why I'll sometimes put the qualification 'of the relevant type' within parentheses or drop it altogether. Of course, this is not meant to deny that there are lots of other types of moral sentences that are philosophically interesting, in addition to moral sentences of the relevant type.

4)   The content, or payoff, of all built moral sentences is specified in terms of

        (multiple / disjunctive) non-enriched singular / enriched singular / general

        substitution, *via* reference to the non-moral properties that can be denoted by

        some non-moral predicate 'N' such that ■ all N's are M[37].

Theses 1)-4) articulate in some detail the strategy summarized by the slogan 'Go by the M-making properties' and show how to apply it to all moral sentences of the relevant type (except those singled out by thesis 3)). As I noted above, given theses i)-iv) this is a strategy that should strike you as a plausible way to go, if your aim is to replace a moral sentence's apparent reference to the moral property of being M with a reference to non-moral properties. The application of this general recipe to the particular case of moral explanatory claims yields the accommodationist proposals I described above; the reasons why it is useful to think of these proposals as applications of the general recipe articulated by 1)-4) will become clear starting from the next chapter.

---

[37] In the present work, I'll distinguish between basic and built sentences only within the set of moral sentences (of the relevant type), so I'll throughout feel free to use 'basic / built sentence' as shorthand for 'basic / built moral sentence (of the relevant type)'. Three remarks. One: it is easy enough to define the various substitution operations for the case of universal moral sentences, so I'll spare the reader the technical details. I'll mention only the following point: the contrast pointed out in n. 31 between moral explanatory claims and individual moral sentences extends to all categorical moral sentences. Trying to capture the content, or payoff, of a universal moral sentence of the form 'All P's are M' by saying that every P instantiates each of the *n*'s would be intuitively too strong: the appropriate thing to say is that every P instantiates (at least) one of the *n*'s. Two: strictly speaking, the clause '*via* reference to…' in 4) is redundant, since it is part of the definitions of the various substitution operations that they replace the apparent reference to M-ness with a reference to the *n*'s. I add it with an eye to the chapters to follow, where I'll broaden the definitions of the substitution operations so as to include operations that replace the apparent reference to M-ness with a reference to non-moral properties other than the *n*'s: read with this broader notion of a substitution operation in mind, the clause in question is not redundant anymore. Three: a careful reader will note the following difference between 1)-4) and the theses I presented in 2.2. Whereas thesis i) defines brute truths as ones that are not made true at all, thesis 1) does not define basic sentences as ones whose content, or payoff, is not specified at all. The reason is that a friend of 1)-4) will likely have *some* account of the content, or payoff, of basic sentences to offer: she will likely have something to tell us, in order to explain what the content, or payoff, of such sentences is. An analogous difference will emerge in 2.4 between 1)-4) and the theses I'll present there.

It is instructive to look in detail at how 1)-4) are related to i)-iv). One helpful way to do this is by comparing and contrasting 1)-4) with one traditional empiricist line of reply to skepticism. In reply to the skeptical challenge that we cannot, for any sentence s of a certain type (say, about material objects), know that s is true because the evidence we can gather in its support (say, about sense data) is never enough to draw that conclusion, one standard empiricist move is to claim that s means the same as the sentence e formulating our evidence (say, that sentences about material objects are translatable into ones about sense data), so that we can, *contra* the skeptic's claim, conclude from our evidence that s is true. In reply to the skeptic, what this empiricist line does is to account for s's meaning in terms of our evidence for it. One can think of 1)-4) as doing something of the same kind, with one important difference: since the empiricist line just described is a reply to skepticism, the materials it relies on are epistemological (i.e. our evidence for s); the materials 1)-4) rely on, on the other hand, are straightforwardly metaphysical. As we saw, theses i)-iv) can be summarized by the slogan 'No ethics without principles': if there is any categorical moral truth, then there are categorical brute moral truths with the features that iv) and -depending on the categorical moral truths there are- ii) and iii) assign to brute moral truths. This is a straightforwardly metaphysical claim: i)-iv) put forward a metaphysical picture that describes how moral reality (if there is one) is structured. Consider now a categorical moral sentence s lacking the features mentioned in iv), hence one that, if true, would express a derived moral truth: what 1)-4) do is to account for the content, or payoff, of s in terms of what, according to i)-iv), would ultimately make s true[38].

---

[38] This is a slightly simplified but illuminating way of describing what 1)-4) do. To see why it is slightly simplified, suppose there are exactly two brute moral truths ascribing M-ness, i.e. ' ■ All N_1's are M' and ' ■ All N_2's are M', and consider 'x is M': on i)-iv), what ultimately makes it true that x is M, supposing it is true, is the brute moral truth that ■ all N_1's are M, plus the non-moral truth that x is N_1, or the brute moral truth that ■ all N_2's are M, plus the non-moral truth that x is N_2, but no matter whether you use (multiple or disjunctive) singular or general substitution to specify the content, or payoff, of 'x is M', its content, or payoff, is not that (x is N_1 and ■ all N_1's are

49

One consequence of this emerges if we look again at the empiricist reply to the skeptic. The empiricist and the skeptic agree on the following claim: our evidence in support of claims about material objects is evidence concerning our sense data. This is an epistemological claim about the nature of our evidence in support of claims of a certain type. On the empiricist proposal, this epistemological claim gets built into the very meaning of the relevant sentences: if we accept that proposal, then we can tell that the epistemological claim in question is true just by looking at the meaning of the relevant sentences. Theses 1)-4) have an analogous consequence: they end up building the metaphysical claim 'No ethics without principles' into the very content, or payoff, of categorical moral sentences. On 1)-4), a categorical moral sentence, or the sentence capturing its payoff, presupposes or entails that there is at least one non-moral property that can be denoted by some non-moral predicate 'N' such that ■ all N's are M, or that 'there are principles'[39]. Said otherwise: if we accept 1)-4), then we can tell, just by looking at the content, or payoff, of a categorical moral sentence, that if it, or the sentence capturing its payoff, is true, then 'there are principles'. On 1)-4), the metaphysical claim 'No ethics without principles' thus becomes something we can read off the content, or payoff, of categorical moral sentences[40].

M) or (x is N_2 and ■ all N_2's are M). Though the formulation in the text is admittedly simplified, there is an obvious sense in which it illuminates the general shape of each of the different ways of specifying the content, or payoff, of a categorical moral sentence mentioned by 4), which is why I use it.

[39] 1)-4) guarantee that this is true of categorical moral sentences lacking the features mentioned in iv), i.e. that are not appropriately strong universal moral sentences using a non-moral predicate denoting a non-moral property; the reason why this is true of categorical moral sentences that do have these features is obvious.

[40] Something stronger is true on 1)-4), given that they provide an account of the very content, or payoff, of categorical moral sentences, namely that a categorical moral sentence, or the sentence capturing its payoff, *cannot* be true unless 'there are principles'. This matches one of the features of i)-iv) pointed out above, i.e. that they are meant to capture an aspect of the very nature of moral properties, so that what they say is not merely 'If there is an ethics, then there are principles', but 'There *can* be no ethics without principles'. A careful reader will have noted that what, on 1)-4), we can read off the content, or payoff, of categorical moral sentences is strictly speaking only a part of the full claim 'No ethics without principles' put forward by i)-iv). This is so because what it takes to count as a 'principle' is being a categorical brute moral truth with the features assigned to brute moral truths by iv), and possibly ii) and iii), whereas what on 1)-4) is presupposed or entailed by a categorical moral sentence, or by the sentence capturing its payoff, is just that there is at least one categorical moral truth with the features mentioned in iv) (it is not presupposed or entailed that such truth be a brute moral truth or that it have the features assigned to brute moral truths by ii) and iii)). For ease of exposition, I am bracketing this complication in the main text.

What about moral explanatory claims? Theses i)-iv) are silent about them, so one cannot say that 1)-4) account for the content, or payoff, of a moral explanatory claim in terms of what, according to i)-iv), would ultimately make such a claim true. However, on 1)-4) it is just as true of a moral explanatory claim, as it is of a categorical moral sentence, that it, or the sentence capturing its payoff, presupposes or entails that there is at least one non-moral property that can be denoted by some non-moral predicate 'N' such that ■ all N's are M, or that 'there are principles'. On 1)-4), the latter claim is true of *any* moral sentence (of the relevant type). In short: theses i)-iv) put forward a metaphysical claim according to which if there is any categorical moral truth, then 'there are principles'; 1)-4) build this metaphysical claim into the content, or payoff, of categorical moral sentences, and extend it to moral explanatory claims by building it into the content, or payoff, of such claims as well.

I said above that 1)-4) specify the content, or payoff, of all moral sentences of the relevant type starting from a subset of them (the basic ones), whose content, or payoff, is assumed as given. The subset in question is defined by 3) as the set of appropriately strong universal moral sentences that can be formulated using a non-moral predicate denoting a non-moral property. However, what 4) looks at (so to speak), in order to specify the content, or payoff, of all the other moral sentences (the built ones), are only the *true* ones among the sentences in the set defined by 3): the way 4) specifies the content, or payoff, of built sentences is *via* reference to the non-moral properties picked out by the true ones among the sentences in that set. So, it might be asked, why not have 3) define the set of basic sentences as the set of *true* appropriately strong universal moral sentences that can be formulated using a non-moral predicate denoting a non-moral property, and include the false ones among such sentences in the

set of built sentences[41]? Because the content, and logical relations, of a sentence s (or of the sentence capturing s's payoff) do not change depending on whether s (or the sentence capturing its payoff) is true or false. But this is exactly the result one would get, if one were to define the set of basic sentences as suggested: on the proposed suggestion, an appropriately strong universal moral sentence that can be formulated using a non-moral predicate denoting a non-moral property would count as basic, and so have its content, or payoff, specified a certain way (however exactly the content, or payoff, of basic sentences is specified), if it (or the sentence capturing its payoff) were true; it would count as built, and so have its content, or payoff, specified a different way (i.e. as described by 4)), if it (or the sentence capturing its payoff) were false. This is, I take it, an unacceptable result, which explains why one cannot define the set of basic sentences as including only the true ones among the appropriately strong universal moral sentences that can be formulated using a non-moral predicate denoting a non-moral property.

It is finally time now to recap the main thread of the present chapter. In order to accommodate our practice with regard to moral explanatory claims, opponents of moral explanations put forward proposals aimed at capturing the content, or payoff, of 'M-ness causes E' by means of a sentence that ascribes no causal power to the property of being M (or any other moral property). All accommodationist proposals put forward in the literature rely on a common idea, which I labeled 'the Assumption'. After presenting in a rough and intuitive way both what the Assumption says and how it might be thought to provide a useful tool to develop an accommodationist proposal (section 1), I proceeded to give a much more explicit and detailed characterization of the Assumption (section 2) and of the ways in which it can be used to develop

---

[41] Strictly speaking (see n. 35): since 4) specifies the content, or payoff, of built sentences *via* reference to the non-moral properties picked out by those, among the sentences in the set defined by 3), that are either true or such that the sentences capturing their payoff are true, why not have 3) define the set of basic sentences as including only those, among the appropriately strong universal moral sentences that can be formulated using a non-moral predicate denoting a non-moral property, that are either true or such that the sentences capturing their payoff are true?

an accommodationist proposal (section 3). Along the way, I pointed out that the accommodationist proposals built on the Assumption can be seen as the application to the particular case of moral explanatory claims of a quite general recipe that can be used to specify the content, or payoff, of all moral sentences (of the relevant type). One noteworthy feature of my discussion has been that, starting from section 2, I have freely assumed that moral sentences can be true or false: this means that the way of fleshing out the Assumption presented in this chapter is available only to a cognitivist, and so are of course also the accommodationist proposals built on it. Starting from the next chapter, I'll begin to examine the ways in which non-cognitivists have attempted to flesh out the Assumption without assuming that moral sentences can be true or false (or anything else that is not available to them) and the accommodationist proposals they have built on these ways of taking the Assumption. The discussion of non-cognitivist attempts at mimicking the accommodationist proposals available to a cognitivist will occupy us for the rest of the present work.

Before turning to it, however, it seems only fair, after dealing a bit with moral metaphysics in section 2 and the philosophy of moral language in section 3, to end this chapter with some remarks on moral epistemology.

4. <u>A possible view in moral epistemology</u>

There is a certain view in moral epistemology that it is for various reasons useful to keep in mind throughout my discussion in the present work. In this section I'll lay it down and describe it in some detail. One reason why it is useful to keep such a view in mind is that it is useful to compare and contrast it with the metaphysical picture put forward by theses i)-iv). Another reason is that I'll describe another view in moral epistemology in the next chapter and it will be

useful to keep the view in this section in mind as a term of comparison for the view I'll describe in the next chapter.

Like i)-iv), the epistemological view in question deals with categorical moral truths. It too can be articulated by means of four different theses:

- All categorical moral truths we know are either <u>inferred truths</u>, i.e. they are inferred from other categorical moral truths (plus appropriate non-moral truths) we know or <u>direct truths</u>, i.e. they are not inferred at all.

- All inferred moral truths are ultimately inferred from direct moral truths (plus appropriate non-moral truths we know).

- All inferred moral truths are ultimately inferred in only one way, i.e. *via* subsumption.

- All direct moral truths can be expressed by an appropriately strong universal moral sentence using a non-moral predicate denoting a non-moral property[42].

The epistemological view articulated by the four theses above has several noteworthy features. It is, in the first place, a foundationalist view. As I understand foundationalism about knowledge, it claims (*a*) that everything we know is either inferred from other things we know or not inferred at all and (*b*) that everything we know inferentially is ultimately inferred from things

---

[42] In the present work, I'll distinguish between direct and inferred truths only within the set of categorical moral truths, so I'll throughout feel free to use 'direct (inferred) truth' or 'direct (inferred) moral truth' as shorthand for 'direct (inferred) categorical moral truth'. I should point out explicitly that the distinction between direct and inferred truths is one within *known* categorical moral truths, so direct and inferred truths are all known truths. I should also point out explicitly that, in formulating the epistemological view in the text, I'll use the phrase 'being ultimately inferred from' so that, in order for a known moral truth m to be ultimately inferred from a known moral truth m', plus a known appropriate non-moral truth n, it is required that m' be known without inference, but it is *not* also required that n be known without inference: the epistemological view I am describing is a view in moral epistemology and it need not take a stand on how things are in the non-moral domain. The four theses in the text should be read accordingly. I'll explain what I mean by 'subsumption' in due course.

we know without inference. On the four theses above, foundationalism turns out to be true of our knowledge of categorical moral truths[43]. Given foundationalism, if t is a categorical moral truth we know that is not direct, then t must be inferred from *other things* we know (and ultimately from things we know without inference). As a matter of fact, the four theses above say something more specific than this: they impose a stricter constraint on what t can be inferred from (and so also on what it can ultimately be inferred from). According to them, t is inferred from *other categorical moral truths* (plus appropriate non-moral truths) we know. This points to the second noteworthy feature of the epistemological view I am describing: it honors the so-called doctrine of the autonomy of ethics, i.e. the claim that there are no inferences leading to a moral conclusion that start from entirely non-moral premises. The third noteworthy feature I want to highlight is that the epistemological view in question imposes an even stricter constraint on what t can ultimately be inferred from. According to the third of the four theses listed above, t is ultimately inferred only *via* subsumption. Intuitively, what I mean by this is that t is ultimately subsumed under some universal moral truth (or truths) involving a non-moral property (or non-moral properties). More formally: I'll say that a known individual / universal moral truth to the effect that x is M / all P's are M is ultimately inferred *via subsumption* when it is ultimately inferred from a series of known moral truths to the effect that all N_1's are M, all N_2's are M, …, where N_1-ness, N_2-ness, … are different non-moral properties, plus a known non-moral truth to the effect that x is N_1 or x is N_2 or … / all P's are N_1 or all P's are N_2 or … / each P is either N_1

---

[43] I borrow the formulation of foundationalism in the text from Sturgeon [2002]. My thinking about moral epistemology in general, and specifically about the epistemological view I am describing (a close cousin of which is discussed by Sturgeon), owes much to this article. Strictly speaking, the epistemological view in question meets condition (*b*) only partially: regarding categorical moral truths we know that are not direct, the view says they are ultimately inferred from moral truths, plus appropriate non-moral truths, we know, and while the former are ones we know without inference, the view remains neutral about whether this is so for the latter as well (see n. 42). Since the foundationalist structure is preserved so far as the moral branch of the inferential tree is concerned, it is fair to regard the view in question as a foundationalist view in moral epistemology: according to it, if there is anything non-foundationalist about our knowledge of categorical moral truths, it is entirely due to how our knowledge of non-moral truths works.

or N_2 or …[44]. The epistemological view I am describing does not rule out that there be more than one inferential path leading to t, but it does say that, no matter how many such paths there are and no matter what their specific details turn out to be, they all start from a (series of) known moral truth(s), plus a known non-moral truth, that are of a form that fits the definition just given of 'being ultimately inferred *via* subsumption'. The last of the four theses listed above, moreover, makes a claim about all direct moral truths, thereby adding a further constraint on the moral truths t can ultimately be inferred from: in addition to being universal and involving non-moral properties, such truths are strong (with the appropriate degree of strength) and involve non-moral properties that can be denoted by non-moral predicates[45].

As I said, my aim in presenting this epistemological view is not to discuss it, but to use it as a helpful foil against which to compare and contrast other views I'll be discussing, so I shall not provide reasons for ascribing it to any particular writer in meta-ethics. It is arguable that the four theses above capture the epistemology endorsed by classical intuitionists: more precisely, by

---

[44] Alternatively: plus a known non-moral truth to the effect that x has the disjunctive property of being $N_1 \lor N_2 \lor …$ / all P's have the disjunctive property of being $N_1 \lor N_2 \lor …$. The definition of 'being ultimately inferred *via* subsumption' is meant to include the limiting case in which the series of known moral truths in question is made up of only one truth.

[45] The definition of 'being ultimately inferred *via* subsumption' matches in obvious ways that of 'being ultimately made true in the relevant way', with one difference worth pointing out: a categorical moral truth can ultimately be inferred *via* subsumption from more than just what ultimately makes it true in the relevant way. To see why, suppose that there are exactly three brute moral truths ascribing M-ness, i.e. ' ■ All N_1's are M', ' ■ All N_2's are M', and ' ■ All N_3's are M'; that one knows all three of them without inference; and that x is N_1 but it is neither N_2 nor N_3. Consider two possible cases: one finds out that x is N_1 and concludes it is M; one finds out that x has one of the first two, but not the third, of the three non-moral properties in question (or that it has one or the other of those three properties) and concludes it is M. On the definition of 'being ultimately inferred *via* subsumption', the individual moral truth that x is M counts as ultimately inferred *via* subsumption both in the first case -from the moral truth that ■ all N_1's are M, plus the non-moral truth that x is N_1-, and in the second case -from the two (three) moral truths that ■ all N_1's are M, ■ all N_2's are M (and ■ all N_3's are M), plus the non-moral truth that x is N_1 or x is N_2 (or x is N_3)-. On the definition of 'being ultimately made true in the relevant way', on the other hand, the individual moral truth that x is M counts as ultimately made true in the relevant way by the moral truth that ■ all N_1's are M, plus the non-moral truth that x is N_1, but it does not also count as ultimately made true in the relevant way by the two (three) moral truths that ■ all N_1's are M, ■ all N_2's are M (and ■ all N_3's are M), plus the non-moral truth that x is N_1 or x is N_2 (or x is N_3). When doing epistemology and talking about how things are inferred by epistemic agents, ignorance is a possible factor to consider and the view I am describing allows for some degree of non-moral ignorance; when doing metaphysics and talking about how things are made true in reality, ignorance is not an issue.

those among them who (like, for instance, G.E. Moore) took the content of intuitive moral knowledge to be general rather than particular (in my terminology, who took intuitive moral knowledge to be of universal rather than individual moral truths). The epistemological view in question is also one of the premises in one standard argument for moral skepticism, given that it (just like theses i)-iv) with regard to the existence of categorical moral truths) is meant to be neutral with regard to whether we have any knowledge of categorical moral truths: the argument relies on this epistemological view -hence on the claim that, if there are any categorical moral truths we know, there are universal moral truths we know without inference- and argues that, since we do not have knowledge of the latter sort, we have no knowledge of categorical moral truths at all[46].

The analogies between i)-iv) and the four theses above are obvious: the two sets of theses are structurally identical[47]. Whereas their shape is the same, the materials they work with are different. Theses i)-iv) deal with categorical moral truths and are framed in terms of the metaphysical relation 'being made true by': they put forward a metaphysical picture that describes how moral reality (if there is one) is structured. The four theses above, on the other hand, deal with our knowledge of categorical moral truths and are framed in terms of the epistemological relation 'being inferred from': they articulate an epistemological view describing how moral knowledge (if there is such a thing) works. Just as I did with the relation 'being made true by' in section 2, I should comment explicitly on certain features of the epistemological relation 'being inferred from' in whose terms the four theses above are framed. Since those theses articulate an epistemological view, the notion of inference at play here is not merely psychological. For example, the claim made by the first thesis that, if a categorical moral truth is

---

[46] See Sturgeon [2002] for closely related claims about the attribution of a view closely related to the one I am describing in the text.
[47] Except for the difference pointed out in n. 45.

inferred at all, it is inferred from other categorical moral truths (plus appropriate non-moral truths), is not a psychological claim about what sort of premises we, as a matter of fact, draw our conclusions from. It is rather a claim about what sort of premises we draw our conclusions from *insofar as* we avail ourselves only of inferences of a certain good sort (which we do only to a limited extent), namely those that preserve the epistemic value of their premises: those that allow us, if we start from things we know, to acquire new knowledge[48]. What inferences fall under this heading is a large issue in epistemology: a possible position is that they include only logically valid inferences; another (more widely held) position is that other inferences, in addition to logically valid ones, qualify as well. This helps bring out the second feature of the relation 'being inferred from' I want to emphasize, namely that it is a different relation from the logical relation 'following from': q's being inferred from p is not the same as q's logically following from p. For one thing, whether or not the second of the two positions just mentioned is correct, q can logically follow from p but fail to be inferred from it simply because the logical relation goes unnoticed, so the inference is not drawn: being inferred is not, after all, the same as being inferable. Moreover, if (as many believe) that position is correct, it is possible for q to be inferred from p even if q does not logically follow from p.

As I noted in section 2, theses i)-iv) represent the most natural way of fleshing out the Assumption. Friends of the Assumption, read this way, may or may not go on to endorse the epistemological view articulated by the four theses above. Whereas i)-iv) put forward a metaphysical picture of how moral reality is structured, the epistemological view I have been

---

[48] Strictly speaking, even if the four theses above were read as theses about the inferences we as a matter of fact draw, this might not end up altering their meaning, given that they are formulated as theses about the categorical moral truths *we know* (for example, the claim made by the first thesis is that, if a categorical moral truth *we know* is inferred at all, it is inferred from …). However, since they articulate an epistemological view, it is only natural to frame them in terms of an epistemological, rather than merely psychological, relation of being inferred from. Moreover, the other view in moral epistemology I'll present in the next chapter is framed in terms of the epistemological relation 'being inferred from' and it is useful, for purposes of comparison, to have the view I am describing here framed the same way.

describing says that our knowledge of such reality works by tracking, so to speak, that metaphysical structure. This is a further, distinct claim than that made by i)-iv), one that someone accepting i)-iv) is not thereby committed to: even if reality is arranged a certain way, our knowledge of it may or may not be arranged the same way. Friends of i)-iv) may claim it is, at least in the case of moral knowledge, and thus endorse the epistemological view presented in this section; they may as well claim it is not and thus reject this view: their epistemology is not settled by their acceptance of the Assumption.

Comparing and contrasting it with theses i)-iv) was one of two reasons I mentioned at the beginning of this section for introducing the epistemological view articulated by the four theses above. The second reason was using it as a term of comparison for the other view in moral epistemology I would discuss in the next chapter: to this other epistemological view we must now turn.

Chapter 3 – The first non-cognitivist strategy, and why it fails

1. <u>Hare's reading of the Assumption</u>

The accommodationist proposals described in chapter 2 all rely on theses i)-iv), which represent the most natural way of fleshing out the Assumption. This way of fleshing out the Assumption freely assumes that moral sentences can be true or false, so i)-iv), and the accommodationist proposals built on them, are available only to a cognitivist. Opponents of moral explanations who want to accommodate our practice with regard to moral explanatory claims are not, however, all cognitivists; indeed, the most forceful proponents of the view that rejecting moral explanations does not require any deep revision of our practice are non-cognitivists. These non-cognitivist philosophers claim they can provide a way of fleshing out the Assumption that does not assume that moral sentences can be true or false (or anything else that is not available to them) and build, on this different way of fleshing out the Assumption, accommodationist proposals that mimic the proposals available to a cognitivist that I described in chapter 2. Non-cognitivists have put forward two different ways of fleshing out the Assumption without assuming that moral sentences can be true or false: in this chapter, I'll discuss the first one and the accommodationist proposals built on it; starting from the next chapter, I'll turn to the second one and to the accommodationist proposal built on it. The upshot of my discussion in the present chapter will be that the accommodationist proposals built on the first non-cognitivist reading of the Assumption fail: they are false and ultimately unable to accommodate our practice with regard to moral explanatory claims. The problems they face, moreover, are specific to them: they are not faced by the accommodationist proposals available to a cognitivist. If built on the reading of the

Assumption we'll discuss in this chapter, the non-cognitivist attempt to mimic the cognitivist proves unsuccessful[1].

Both attempts at providing a non-cognitivist reading of the Assumption flesh out the idea that actions, people, institutions, etc. owe their moral features to their non-moral properties in terms of the idea that use of moral language commits one to general moral principles; they differ on the nature and extent of the commitment. In a nutshell: on the first of them, use of moral language commits you to having general moral principles from which you infer the other moral beliefs you have; on the second, use of moral language commits you to holding general moral principles that entail your other moral beliefs, if you develop your views so as to become fully decided on moral and non-moral matters.

The first non-cognitivist reading of the Assumption is due to Hare [1952], its use in order to develop an accommodationist proposal to Blackburn [1991a]. As we saw in chapter 2, theses i)-iv) put forward a straightforwardly metaphysical picture describing how categorical moral truths are arranged in terms of the metaphysical relation 'being made true by'. Hare's strategy in providing a non-cognitivist reading of the Assumption is to preserve the shape of that

---

[1] Don't contemporary non-cognitivists tend to be sympathetic to minimalism about truth and so insist they too can say that moral sentences can be true or false? Though contemporary non-cognitivists do usually insist they too can say that moral sentences can be true or false, they also insist there still remain important metaphysical and epistemological differences between their position and, say, that of the realist (after all, this is why, in their view, we should choose non-cognitivism over realism): for example, they insist that the notion of moral truth they can make room for lacks the allegedly extravagant flavor that the realist's notion, with its metaphysical commitments, has. As we saw in the previous chapter, however, theses i)-iv) deal precisely in moral metaphysics: they put forward a straightforwardly metaphysical picture that describes how moral reality is structured. Even if they insist they can make room for moral truth, non-cognitivists will thus want to reject such a picture and propose a different way of fleshing out the Assumption. There is a line to be drawn between the metaphysical and epistemological commitments of non-cognitivism and those of other meta-ethical positions, and this makes theses i)-iv) not available to non-cognitivists: these two points are not controversial. What is controversial, given non-cognitivists' insistence that they can make room for moral truth, is that the line at issue can be drawn in terms of the question whether moral sentences can be true or false. Since this is the traditional way of drawing that line, in the present work I'll continue to adopt it. *If* minimalism about truth turns out to be correct, and non-cognitivists actually can earn their right to a robust enough notion of moral truth (a question on which, as they themselves recognize, they bear the burden of proof, and upon which my discussion in this and the next few chapters will have some bearing), then that line will need to be drawn a different way, but my point in the text will not be affected. The views about non-cognitivism and moral truth expressed here owe much to Sturgeon [1986b].

metaphysical picture and replace the materials it is built out of (categorical moral truths, the relation 'being made true by') with ones available to the non-cognitivist. In order to see how he implements this strategy, we need to look at (one of the central aspects of) his meta-ethical theory.

The function of moral (and, more generally, evaluative) language is, according to Hare, to guide choices and actions: moral language is one of the tools we have to frame answers to questions of the form 'What shall I do?'. This kind of questions can only be answered, says Hare, by means of an imperative sentence: you do not answer my question 'What shall I do?' until you tell me 'Do this!'. Moral language can be used to frame answers to questions of this sort because moral sentences entail imperatives: you can answer my question 'What shall I do?' by saying 'You ought to do this' because 'You ought to do this' entails, on Hare's view, 'Do this!'. What then is the difference between an imperative sentence and a moral sentence? What else are you telling me in addition to 'Do this!' if, in answer to my question 'What shall I do?', you tell me 'You ought to do this'? You are, according to Hare, appealing to a general principle or standard, thereby committing yourself to a whole range of other imperative sentences, in addition to the one you are offering me in answer to my question on this particular occasion.

Suppose for example I win a little sum at the lottery and ask you 'Shall I keep this money or shall I give it to Judy?'. If you answer 'Give it to Judy!', you are offering me guidance only with respect to the particular situation we are discussing, namely mine with regard to this money and Judy now. If instead you answer 'You ought to give it to Judy', you are, on Hare's view, appealing to a general principle that you think applies to this particular situation: you might, say, think I borrowed money from Judy in the past, and appeal to the principle that one ought to return money one has borrowed from someone else as soon as one can do so; or you might think

my friend Judy is in need of money whereas I am not, and believe that friends ought to help each other in difficult times when they can do so at little cost to themselves. No matter what particular principle you are invoking, by appealing to a principle you are committing yourself not just to the imperative 'Give it to Judy!' addressed to me on this specific occasion, but also to a whole range of other imperatives: for instance (if the principle you are appealing to is the first of the two mentioned above), to the imperative 'Give it to Marta!' addressed to someone who has borrowed money from Marta and is in a position to return it. If you answer my question by saying 'You ought to give it to Judy', you are thus offering guidance not just with respect to the particular situation we are discussing, but with respect to a whole *kind* of cases (of which my current situation is an example) in which one might face the question 'What shall I do?'. We often want to be able to offer this sort of guidance: moral language is the tool that enables us to do it. Whereas imperative sentences allow us to offer guidance that is limited only to particular choices and actions, moral (and, more generally, evaluative) language is, according to Hare, the tool we have to offer guidance with respect to *kinds* of choices and actions.

Since this is the function of moral language, it is reflected in the logic of moral predicates, i.e. in the rules according to which competent users of moral language use its predicates. It is a feature of the logic of moral predicates that, for any moral predicate 'M', any user S of moral language, and any object x, if S believes that x is M then, for some property N-ness, S is committed to accepting a general principle or standard to the effect that all N's are M and to believing that that principle or standard applies to x, hence that x is N[2]. If someone were

---

[2] Hare usually frames his claims as claims about moral language (for instance, about what one is committed to in asserting a certain moral sentence). The way he makes use of those claims in his arguments suggests, however, that he takes them to apply to moral thought just as much as to moral language (for instance, that he holds parallel claims about what one is committed to in holding a certain moral belief); and a meta-ethical theory needs an account of moral thought anyway. In presenting Hare's views, I'll thus feel free to extend what he says explicitly about moral language to moral thought as well and speak, for instance, of someone holding a certain moral belief as a user of moral language and moral predicates or as a speaker. I should note here that whether Hare's views, and more

to propose that x is M, but there were no general principle s/he took to apply to x and yield this verdict about it, s/he would, according to Hare, be abusing moral language. This claim about the logic of moral predicates captures the first of the two non-cognitivist readings of the Assumption: let us take a close look at it.

What I said above is 'if S believes that x is M then, for some property N-ness …'. Hare is clear that the property in question must be a non-moral property, which should of course come as no surprise, given that the claim we are looking at captures Hare's way of fleshing out the idea that actions, people, institutions, etc. owe their moral features to their non-moral properties. If S holds an individual moral belief, S must thus infer it, according to Hare, from a general principle and an appropriate non-moral belief S holds. He also suggests that, though it can at times be hard to formulate the standards that guide our application of moral predicates, we either already have, or can at least introduce into our language, a non-moral predicate denoting the non-moral property in question; and in any case, since the claim we are looking at is a tool non-cognitivist opponents of moral explanations use in order to develop accommodationist proposals, the considerations mentioned in 2.2 apply here too. I'll thus take the claim to be that if S believes that x is M then, for some non-moral property that can be denoted by a non-moral predicate 'N' …[3].

generally non-cognitivism, can plausibly be extended from the domain of moral language to that of moral thought is a live issue; however, since my interest here is not in non-cognitivism as such, but in non-cognitivist attempts at accommodating our practice with regard to moral explanatory claims, I'll bracket this issue and assume that non-cognitivists face no problems here (for a recent discussion of whether non-cognitivism should be formulated as a view about moral language or thought, see Schroeder [2008] and Alwood [2010]). In formulating Hare's views, I am going to need a term, neutral with respect to the choice between cognitivism and non-cognitivism, to refer to the counterpart, in the domain of thought, of the act of asserting a moral sentence: any choice here is bound to be a stipulation, my choice is to appropriate the term 'belief' for this purpose, thereby allowing non-cognitivists to speak of someone's holding a moral belief, even if they deny there can be moral truth or falsehood. Though, given my stipulation, non-cognitivists can speak of people's having moral beliefs, there remain interesting questions to answer about the status of such beliefs under non-cognitivism (I'll mention one of them in n. 8); in what follows, however, I'll bracket such questions and assume that non-cognitivists can answer them in a satisfactory way.

[3] Hare also recognizes that our standards can at times be somewhat vague; still, he insists, if it is a *moral* (or, more generally, evaluative) belief that we have about an object, we must be appealing to a (possibly vague) general

For reasons analogous to those mentioned in 2.2, Hare takes moral standards to cover not just actual cases but also possible counterfactual ones: in my terminology, he takes them to be expressible by a strong universal moral sentence. What exactly is the degree of strength at play here? Since a non-cognitivist can endorse Hare's claim and use it to develop an accommodationist proposal without taking a stand on this issue, I'll bracket it and, as I did in chapter 2, use the symbol ' ■ ' for the degree of strength in question, whatever exactly it turns out to be[4].

Since Hare's point is that one cannot use a moral sentence, unlike an imperative sentence, to guide choices and actions without thereby committing oneself to a general principle or standard, he usually frames the claim we are looking at *via* reference to individual moral sentences (or, as I did above, to individual moral beliefs: if S believes *that x is M* then …). If your point is that use of moral language commits one to general principles or standards, the most striking way for you to put it is by saying that this commitment is built already into the use of individual moral sentences. It is clear, however, that Hare takes his claim to apply more broadly. Consider, for instance, the universal moral belief that all the members of the Smith family are good people. Hare takes his claim to apply to such a belief as well: if someone believes that all the members of the Smith family are good people, s/he must accept a general principle (possibly

---

standard: as he puts it writing about the predicate 'good' (whose function in his view is to commend), "[w]henever we commend, we have in mind something about the object commended which is the reason for our commendation" (Hare [1952], p. 130). In what follows, I'll bracket the question of how exactly to make room for vague standards (should we say that, if S believes that x is M, there is a vague property … ? Can properties be vague? Or should we instead say that, if S believes that x is M, there is a property … , but S is unable to give a full characterization of that property and so to fully specify the standard she is applying?) and assume this can be done in a satisfactory way.

[4] There might be interesting questions to address for a non-cognitivist who wants to provide even a rough-and-ready characterization of the relation 'being stronger than' along the lines of the one I offered in chapter 2 (which appealed to the notion of truth). In what follows, I'll assume they do not raise any serious problem for the non-cognitivist. As I say in the main text, Hare takes moral standards to be expressible by a strong universal moral sentence; a further claim Hare makes is that such a sentence must use a non-moral predicate containing no singular terms (so that, say, ' ■ Any action that promotes my own happiness is right' won't do, though ' ■ Any action that promotes the agent's own happiness is right' will). Though this is one of the most characteristic claims of Hare's meta-ethical theory, it is one that a non-cognitivist can reject while still being able to rely on Hare's reading of the Assumption in order to develop an accommodationist proposal, so in presenting Hare's views I'll drop this claim.

a series of general principles) s/he takes to apply to the Smiths (possibly different principles to different family members) and yield this verdict about them[5]. Quite generally, Hare's claim is meant to apply to *all* categorical moral beliefs (with one exception I'll mention shortly): if S holds any categorical moral belief (except for those belonging to a class to be mentioned shortly), S must infer it from a general principle (possibly a series of general principles) and an appropriate non-moral belief S holds. The one exception is represented, according to Hare, by the universal moral beliefs that constitute one's standards: whereas all other categorical moral beliefs S holds must be inferred from beliefs belonging to this class (plus appropriate non-moral beliefs), such beliefs are the only ones that S need not, and indeed in Hare's view cannot, infer.

Hare's way of fleshing out the Assumption that actions, people, institutions, etc. owe their moral features to their non-moral properties can thus be articulated by means of the four following theses:

i.H)   All categorical moral beliefs we have are either <u>inferred beliefs</u>, i.e. they are inferred from other categorical moral beliefs (plus appropriate non-moral beliefs) we have or <u>direct beliefs</u>, i.e. they are not inferred at all.

ii.H)   All inferred moral beliefs are ultimately inferred from direct moral beliefs (plus appropriate non-moral beliefs).

iii.H) All inferred moral beliefs are ultimately inferred in only one way, i.e. *via* subsumption.

---

[5] Strictly speaking, this is so on the assumption that, if someone believes that all the members of the Smith family are good people, this is not itself one of his/her general moral principles; if (surprisingly) it is, the case falls within the class of exceptions to Hare's claim I am about to mention in the main text.

iv.H) All direct moral beliefs can be expressed by an appropriately strong universal moral sentence using a non-moral predicate denoting a non-moral property[6].

What I mean by 'subsumption' here is the same as what I meant in 2.4: the definition offered there applies only to known categorical moral truths, but it can easily be extended to categorical moral beliefs as well. So I'll say that an individual / universal moral belief to the effect that x is M / all P's are M is ultimately inferred _via subsumption_ when it is ultimately inferred from a series of moral beliefs to the effect that all N_1's are M, all N_2's are M, …, where N_1-ness, N_2-ness, … are different non-moral properties, plus a non-moral belief to the effect that x is N_1 or x is N_2 or … / all P's are N_1 or all P's are N_2 or … / each P is either N_1 or N_2 or …[7]. The notion of

---

[6] In the present work, I'll distinguish between direct and inferred beliefs only within the set of categorical moral beliefs, so I'll throughout feel free to use 'direct (inferred) belief' or 'direct (inferred) moral belief' as shorthand for 'direct (inferred) categorical moral belief'. In formulating i.H)-iv.H), I'll use the phrase 'being ultimately inferred from' (in conformity with my usage in formulating the epistemological view described in 2.4) so that, in order for a moral belief m to be ultimately inferred from a moral belief m', plus an appropriate non-moral belief n, it is required that m' be held without inference, but it is _not_ also required that n be held without inference: Hare's claim is a claim about ethics and it need not take a stand on how things are in the non-moral domain. Theses i.H)-iv.H) should be read accordingly. A careful reader will note that i.H)-iv.H) draw a distinction, between what a moral belief is inferred from and what it is ultimately inferred from, that is not explicitly drawn by Hare. The reason for this is that Hare (at least on a sympathetic reading of him) need not deny that S can infer a categorical moral belief from categorical moral beliefs other than general principles (plus an appropriate non-moral belief) she holds: for instance, he need not deny that S can infer 'All the members of the Smith family are good people' from 'Mr. Smith is a good person', 'Mrs. Smith is a good person', and 'Jack Smith is a good person' (plus 'Mr. and Mrs. Smith and Jack Smith are all the members of the Smith family'). All he needs to claim is that, if S does so, she must infer the latter categorical moral beliefs from further categorical moral beliefs and that this way we eventually get to general principles S holds, so that 'All the members of the Smith family are good people' is ultimately inferred from general principles (plus an appropriate non-moral belief) S holds: for instance, from ' ▪ All law-abiding adult citizens are good people' and ' ▪ All young children who obey their parents are good people' (plus 'Each member of the Smith family is either a law-abiding adult citizen or a young child who obeys his/her parents').

[7] Alternatively: plus a non-moral belief to the effect that x has the disjunctive property of being $N_1 \lor N_2 \lor \ldots$ / all P's have the disjunctive property of being $N_1 \lor N_2 \lor \ldots$. The definition of 'being ultimately inferred _via_ subsumption' is meant to include the limiting case in which the series of moral beliefs in question is made up of only one belief. I should point out explicitly that i.H)-iv.H) do not rule out that there be more than one inferential path through which S gets to an inferred moral belief she holds; they do, however, say that, no matter how many such paths there are and no matter what their specific details turn out to be, they all start from a (series of) moral belief(s), plus a non-moral belief, that are of a form that fits the definition just given of 'being ultimately inferred _via_ subsumption'. Once again, a careful reader will note that i.H)-iv.H) make room for something Hare does not explicitly make room for. Suppose, for instance, that S accepts three different general principles, i.e. ' ▪ All N_1's are M', ' ▪ All N_2's are M', and ' ▪ All N_3's are M', and consider two cases: S believes that x is N_1 and concludes it is M; S believes that x has one of the first two, but not the third, of the three non-moral properties in question (or that it has one or the other of

67

inference at play here is also the same as that at play in 2.4: the relation 'being inferred from' in whose terms i.H)-iv.H) are framed is an epistemological relation. As pointed out in 2.4, the relation in question is thus a different relation from the logical relation 'following from'. It is also not a merely psychological relation: for example, the claim made by i.H) that, if a categorical moral belief is inferred at all, it is inferred from other categorical moral beliefs (plus appropriate non-moral beliefs), is not a psychological claim about what sort of premises we, as a matter of fact, draw our conclusions from. It is rather a claim about what sort of premises we draw our conclusions from *insofar as* we avail ourselves only of inferences of a certain good sort (which we do only to a limited extent), it being precisely the task of epistemology to determine what inferences fall under this category[8].

---

those three properties) and concludes it is M. Hare usually writes as though, if S believes that x is M, S must take a specific general principle to apply to x, i.e. S must be in a situation of the first type; but he (at least on a sympathetic reading of him) need not deny that, if S believes that x is M, S's non-moral beliefs about x can be less specific than that and S can be in a situation of the second type. The definition of 'being ultimately inferred *via* subsumption' makes room for the latter case (in addition to the former), allowing S's belief that x is M to be ultimately inferred from the two (three) moral beliefs that ■ all N_1's are M, ■ all N_2's are M (and ■ all N_3's are M), plus the non-moral belief that x is N_1 or x is N_2 (or x is N_3).

[8] Can a non-cognitivist distinguish between inferences of a good and of a bad sort, epistemologically speaking? How is she going to draw such a distinction? In 2.4, when describing the epistemological view presented there, I drew it by saying that good inferences are those that preserve the epistemic value of their premises: those that allow us, if we start from things we know, to acquire new knowledge. Non-cognitivists cannot adopt this way of drawing the distinction, since they deny there can be moral truth, hence moral knowledge; and it is not clear they can mimic it by constructing analogs of the notions of moral truth and knowledge, since it is not clear they can earn their right to notions that are robust enough to enable them to explain what is good about the inferences in question, or why it should matter that we use them, as opposed to others, when developing our moral belief systems. If it is not the purpose of our moral belief systems to represent the way moral reality is arranged, why should it matter, for instance, that we use logically valid inferences (granting that, on non-cognitivism, moral beliefs can stand in logical relations to one another)? The same question arises about the other inferences (if there are any) that, though not logically valid, still count as epistemologically good ones. I believe these are interesting questions to raise about non-cognitivism. Though I am here bracketing some complications (for example, does it make a difference if the epistemological notion we focus on is justification rather than knowledge?), and though non-cognitivists have taken steps to address these worries by trying to offer a story about the function of moral language and thought that, using only non-cognitivist materials, allows them to draw a distinction between epistemologically good and bad inferences (as we have seen, Hare himself offers such a story to motivate the epistemological claims we are looking at, and he is well-known for arguing that moral beliefs can, even on non-cognitivism, stand in logical relations to one another), I believe there are still questions non-cognitivists need to answer here. In what follows, however, I'll bracket this issue and assume that non-cognitivists can draw a distinction between epistemologically good and bad inferences and that they can agree with everyone else about what inferences fall under these headings (for example, they can say that logically valid inferences are epistemologically good ones). For a writer raising some of the same worries as

The epistemological view formulated by theses i.H)-iv.H) is thus identical to the view described in 2.4, except for one crucial difference: whereas the view presented in 2.4 is about our *knowledge* of categorical moral *truths*, i.H)-iv.H) are about our categorical moral *beliefs*. The view formulated by i.H)-iv.H) can therefore be thought of as the non-cognitivist analog of the epistemological view described in 2.4, with which it shares the three features I pointed out in the previous chapter: it is a foundationalist view, it honors the doctrine of the autonomy of ethics, and it imposes a strict constraint on the kinds of beliefs that a categorical moral belief can ultimately be inferred from[9]. This view captures the first of the two non-cognitivist readings of the Assumption. As we saw in the previous chapter, the most natural way of fleshing out the Assumption is by means of theses i)-iv), i.e. by means of a straightforwardly metaphysical picture describing how categorical moral truths are arranged in terms of the metaphysical relation 'being made true by'. In order to provide a reading of the Assumption that does not assume that moral sentences can be true or false (or anything else that is not available to her), one option for the non-cognitivist is, as I put it above, to preserve the shape of that picture and replace the materials it is built out of, turning it into an epistemological picture describing how our categorical moral beliefs are arranged in terms of the epistemological relation 'being inferred from': this is what theses i.H)-iv.H) do. And this is in a sense a natural option to take for the non-cognitivist: if you are looking for a way to preserve the idea that a thing's moral features are due to its non-moral properties while denying there can be moral truths, a natural thought is to flesh out that idea in epistemological terms. Remember the feature of our moral practice that the

_____

those mentioned in this footnote, though not in the context of a discussion of non-cognitivist epistemology, see Sturgeon [2009], pp. 85-87.

[9] In the previous chapter I formulated foundationalism as a view about knowledge; at the same time, I formulated it as a view about how knowledge works, one that remains neutral with regard to whether we actually have any knowledge. So a view that remains neutral with regard to whether we have any knowledge of categorical moral truths, like the view formulated by i.H)-iv.H) (or, as I pointed out in the previous chapter, the epistemological view described in 2.4) can still count as a foundationalist view, and philosophers who deny we have any moral knowledge, like non-cognitivists, can still endorse a foundationalist view.

Assumption is meant to capture and give an account of: when someone says something of the form 'x is M' (where 'M' is a moral predicate), it seems it is always appropriate to ask 'What is M about it?'. If what accounts for this is not the way in which categorical moral truths are arranged, it is a natural thought to suggest that it is the way in which our categorical moral beliefs are arranged[10].

This brings out a difference between cognitivist and non-cognitivist friends of the Assumption that is worth pointing out. If you are a cognitivist, accepting the Assumption amounts to accepting a certain metaphysical thesis. As we saw in 2.4, there is an epistemological view that has the same shape as that metaphysical thesis, but it is a view you are not committed to: as I put it in 2.4, your epistemology is not settled by your acceptance of the Assumption. On the other hand, if you are a non-cognitivist (and adopt Hare's reading of the Assumption), your epistemology *is* settled by your acceptance of the Assumption, because the Assumption (on your reading of it) just is an epistemological thesis.

My view is that this difference has important consequences when it comes to assessing the proposals about moral explanatory claims that opponents of moral explanations put forward in order to accommodate our practice with regard to such claims: non-cognitivist accommodationist proposals face specific problems, ones that are not faced by cognitivist proposals, because of it. Before I can begin to argue for this claim, we need to look at the

---

[10] This option will seem even more natural for the non-cognitivist if one looks at things from a historical perspective. Though, as I pointed out in chapter 2, the metaphysical view articulated by i)-iv) and the epistemological view described in 2.4 are distinct views and someone holding the first is not committed to the second, it is arguable that early cognitivist friends of the Assumption (like Moore) did not keep them clearly distinct and held both. When non-cognitivists came on the scene and looked for a way to preserve the Assumption, they could only keep the epistemological component of this cognitivist package, or rather its non-cognitivist analog, i.e. theses i.H)-iv.H).

accommodationist proposals that non-cognitivists have built on Hare's way of fleshing out the Assumption[11].

2. The 'Go by the speaker's standards' strategy

Just as in the case of cognitivist proposals, non-cognitivist accommodationist proposals built on Hare's way of fleshing out the Assumption can be seen as the application to moral explanatory claims of a more general recipe that can be applied to other moral sentences as well. It is instructive, in studying those proposals, to look at the general recipe.

Consider individual moral sentences, i.e. sentences of the form 'x is M', where 'M' is a moral predicate and 'x' is a singular term. An interesting feature of Hare's meta-ethical theory is that Hare recognizes, indeed emphasizes, that, despite the differences that in his view exist between moral and non-moral sentences, some of the jobs performed by a moral sentence are the same as those performed by a non-moral one. In particular, Hare stresses that we often use sentences of the form 'x is M' to convey non-moral information about x: to use one of his examples, if two parsons are talking about Marta, one can tell the other 'Marta is a good girl' in order to convey the information that Marta goes to church and obeys a certain set of Biblical injunctions. This feature of the way in which we use individual moral sentences is easy to

---

[11] In the text I ascribe the non-cognitivist reading of the Assumption articulated by i.H)-iv.H) to Hare because he both provides a systematic defense of it and (as we are about to see) relies on it to develop a proposal about the semantic content of certain moral sentences. However, it is worth pointing out that the epistemological view formulated by i.H)-iv.H) has been present in the non-cognitivist tradition since its beginning. It is one of two (incompatible) views in moral epistemology, both of which are present in Ayer [1936], ch. 6. Ayer's official position is that we cannot rationally debate about moral questions. In the course of his argument, however, Ayer ends up defending a less radical position (apparently without noticing the inconsistency), according to which there is some limited room for rational discussion in ethics and we can, provided certain conditions are met, rationally debate about moral questions. How can we do that? According to Ayer, only by relying on background moral assumptions we cannot reason about: more specifically, only if we start from (what I am here calling) universal moral beliefs that we assume without inferring them from anything else, and then check whether, in light of the non-moral facts, other categorical moral beliefs can eventually be subsumed under them. This less radical position of Ayer's arguably amounts to the epistemological view formulated by i.H)-iv.H). (The presence of these two incompatible positions in Ayer's discussion was pointed out to me by Nick Sturgeon.)

account for, for someone holding that moral predicates can be analyzed in terms of non-moral ones, whereas it raises a *prima facie* difficulty for a non-cognitivist, since non-cognitivists deny that moral predicates can be analyzed this way. Hare's strategy in dealing with this difficulty is to propose that the semantic content of a sentence of the form 'x is M' can (as friends of such analyses would have it) be given by a sentence that replaces the moral predicate with a non-moral one, even if the non-moral predicate in question is not the one in whose terms the moral one can be analyzed (since there is no such non-moral predicate). Given i.H)-iv.H), there is a plausible way of implementing precisely this strategy, one that could be summarized by the slogan 'Go by the speaker's standards'.

In order to see what Hare's proposal looks like in detail, let us consider a user S of moral language who has a fully worked out theory about what things are M, for a given moral predicate 'M': in particular, S has made up her mind about what general principles (in Hare's terminology) or direct moral beliefs (in my own) to accept. Theses i.H)-iv.H) are neutral as to the number of direct moral beliefs S can have, for a given moral predicate 'M'. No matter what their number is, these beliefs can all be expressed by an appropriately strong sentence of the form 'All N's are M', for some non-moral predicate 'N' denoting a non-moral property. Let us make a simplifying assumption and assume that there is only one direct moral belief S holds, hence that there exists *the* non-moral property that can be denoted by some non-moral predicate 'N' such that S believes that ■ all N's are M –for brevity's sake, I'll shorten this definite description to 'the n_S'-; consider now the non-moral predicate, whatever it happens to be, denoting such property:

suppose it is 'N_1'[12]. There are two different ways of implementing the strategy mentioned above. The first way is the following:

(a.H) The content of 'x is M' (as uttered by S) is that x is N_1[13].

This way of specifying the content of individual moral sentences is appropriate, according to Hare, for instances of their use in which they do not perform their typical function, i.e. guiding choices and actions. Individual moral sentences can come to be used this way when one's standards become so rigid and, so to speak, ossified that, in using such sentences, one is not expressing the relevant non-cognitive attitude (commendation, condemnation, etc., depending on what 'M' is) toward x anymore, but is merely reporting it has certain non-moral properties[14].

---

[12] Hare's proposal is neutral as to what is the non-moral predicate denoting the n_S, i.e. as to what is the only direct moral belief we are assuming S holds: for example, if 'M' is 'good', it is neutral as to whether the only direct moral belief S holds is that ■ all pleasant things are good or that ■ all desired things are good or something else. In order to formulate Hare's proposal, I am supposing that the non-moral predicate in question is 'N_1'; if it is a different predicate, then that predicate should be replaced for 'N_1' throughout my formulations (and if the non-moral predicate in question is not one we already have but one that needs to be introduced into our language, then that predicate –supposing we do introduce it into our language- should be replaced for 'N_1' throughout my formulations).

[13] For the reasons given in 1.2, in my discussion of non-cognitivist accommodationist proposals I'll put the payoff option aside and frame such proposals in terms of the content of moral explanatory claims only. I'll do the same with proposals about categorical moral sentences, since my interest in them is as terms of comparison for proposals about moral explanatory claims; and in any case, non-cognitivist proposals about categorical moral sentences are typically put forward as proposals about the content of such claims. Like the cognitivist proposals described in chapter 2, non-cognitivist proposals usually take a stand on how to deal with the negations of the relevant claims. In the case of (a.H), how to deal with the negation of an individual moral sentence is pretty straightforward: just take the negation of the sentence capturing the content of the individual moral sentence, i.e. the content of 'x is not M' (as uttered by S) is that x is not N_1.

[14] Variations on (a.H) account for what Hare calls 'inverted commas' and 'conventional' uses of individual moral sentences. Here is an example of the inverted commas use: I say 'x is your duty' and do not express my approval of your doing x, meaning simply that x belongs to a certain kind of actions, i.e. the kind of actions that a certain, usually well understood, group of people would think it is your duty to perform. And here is an example of the conventional use: I say 'x is your duty' and do not express my approval of your doing x, meaning simply that x belongs to a certain kind of actions, i.e. the kind of actions that people would generally think it is your duty to perform (that is, in calling x your duty, I am merely paying lip service to the generally accepted moral standard). Such uses of individual moral sentences can, according to Hare, be accounted for along the lines of (a.H), provided that the non-moral predicate one replaces 'M' with is the one denoting the property picked out by the direct moral belief held, respectively, by the relevant group of people and by people generally, rather than by the speaker.

When used this way, individual moral sentences do not perform their typical function, so it is not surprising that, used this way, they lack any specifically moral, i.e. non-cognitive, content: on (a.H), the content of an individual moral sentence is entirely non-moral. This use is parasitic on the central, more typical use of individual moral sentences, in which the non-cognitive element is present. For instances of this more typical use, Hare proposes to implement the strategy mentioned above in a slightly different way:

(b.H) The content of 'x is (is not) M' (as uttered by S) is that x is (is not) N_1 and ■ all N_1's are M.

On (b.H) an individual moral sentence does have moral content, captured by the additional clause ' ■ all N_1's are M', whose function is, on non-cognitivism, to express the relevant non-cognitive attitude (commendation, condemnation, etc., depending on what 'M' is) toward N_1's[15].

---

[15] Two remarks. One: shouldn't such additional clause be stronger than ' ■ all N_1's are M'? Shouldn't it reflect the fact that S holds one *and only one* direct moral belief, so that she sees x as M if *and only if* she sees it as N_1? Yes. Note that, given i.H)-iv.H), one can strengthen the additional clause in either of two different ways. One is by having it be ' ■ all N_1's are M *and* ■ all M's are N_1' (equivalently: ' ■ all N_1's are M *and* ■ everything that is not N_1 is not M'). The other is by having it be ' ■ all N_1's are M *and* not (■ all N_2's are M) and not (■ all N_3's are M) and …'. These two ways of strengthening the additional clause are, as I say, different: they are not logically equivalent. But, given i.H)-iv.H), both of them reflect the fact that S sees x as M if and only if she sees it as N_1. Why, then, if the additional clause needs to be strengthened in one of these two ways, do I frame it the way I do in the main text? Because I want to make things simpler for the reader and easier for my opponent. There is a problem for Hare lurking here, one that I believe cannot be solved but that, for the purposes of the present work, I am going to grant he can solve, so I am going to bracket it in order to keep things simple. Here is the problem. As I have just pointed out in the main text, the additional clause is meant to be accounted for in terms of the core non-cognitivist idea that the function of a moral sentence is to express a non-cognitive attitude. Now, while it might not be too hard to see how to apply that idea to ' ■ all N_1's are M', it is not so easy to see how to apply it either to ' ■ all M's are N_1' (equivalently: ' ■ everything that is not N_1 is not M') or to 'not (■ all N_2's are M)' (and to 'not (■ all N_3's are M)', etc.): am I, in uttering either sentence, expressing any non-cognitive attitude? Toward what? This is one way to raise a standard worry about non-cognitivism, namely that it may have trouble dealing with negation. The problem for Hare is that, given this worry, it might be hard for him to have an additional clause that both is strong enough and can be accounted for in terms of the core non-cognitivist idea. Hare does not have much to say about this problem or about the underlying worry concerning negation and, as I have already remarked, I believe this is a problem he cannot solve. However, since the issues raised by the worry about negation are distinct from those I want to raise here, in the present work I won't press this problem and I'll grant Hare that he can solve it, i.e. that he can strengthen

It is the presence of this non-cognitive component that, according to what might be called the Master Argument for non-cognitivism[16], makes Hare's account of individual moral sentences preferable to the account adopted by the cognitivist mentioned above, i.e. the one who believes that moral predicates can be analyzed in terms of non-moral ones.

According to the Master Argument, the non-cognitivist succeeds where the cognitivist fails, namely in accounting for moral agreement and disagreement even among speakers with widely different moral standards. The cognitivist is unable to account for this because, if a moral predicate 'M' can be analyzed in terms of a non-moral one, the non-moral predicate in question will have to be determined (according to the Master Argument) by the standards a speaker uses in applying 'M'. Consider then two users of moral language, S_1 and S_2, and suppose (again, to make things simple) that each of them employs only one standard, but that they employ different

---

the additional clause in one of the two ways mentioned above and account for it in terms of the core non-cognitivist idea. Indeed, to keep things simple, I'll phrase the additional clause as ' ■ all N_1's are M' and drop the rest, taking it as understood (I'll remind the reader of it when necessary). In addition to the problem related to negation, there is also another problem for Hare lurking here, one that I *will* press in what follows; I won't be in a position to do so, however, until we get to chapter 5. Two: on (b.H) the content of an individual moral sentence has two components, a non-moral and a moral one, which Hare calls the 'descriptive meaning' and the 'evaluative meaning' of the sentence, respectively. Hare's view is that the descriptive meaning of an individual moral sentence is secondary (at least in the case of sentences using what we would now call 'thin' moral predicates, like 'good'). This does not mean, however, that on Hare's view descriptive meaning is not meaning: it is, so a full specification of the content of the sentence has to mention it. Hare introduces the notion of descriptive meaning in order to account for our use of individual moral sentences to convey non-moral information (see Hare [1952], pp. 112-118): he considers (pp. 117-118) the objection that the non-moral information in question is something the hearer infers from the speaker's utterance of the sentence, together with her knowledge of the speaker's standards, rather than part of the meaning of the sentence, but he eventually rejects it (Hare does start out by suggesting that the objection is largely verbal, so that his objector is not disagreeing with him on any issue of substance but merely suggesting to use the word 'meaning' in a different way, but even so he eventually rejects the objector's suggestion, "deem[ing] it best to adopt the term 'descriptive meaning'"; moreover, Hare provides reasons for the conclusion that his choice is overall "best", and those reasons sound like something more than merely a justification for using the word 'meaning' one way rather than the other: they sound like a justification for a motivated extension of the notion of meaning, so that there might be more than a merely verbal disagreement between him and his objector). Hare then goes on to treat descriptive meaning as one of the components of the meaning of an individual moral sentence throughout his book (for just one illustration, with regard to the parson example I mentioned above, see p. 146). Finally, he explicitly recognizes (p. 121) the existence of individual moral sentences (those using what we would now call 'thick' moral predicates) whose descriptive meaning is primary. The view that descriptive meaning is secondary means, not that descriptive meaning is not meaning, but (as Hare stresses: see pp. 118-119 and pp. 148-150) that it is neither the only component of the meaning of individual moral sentences nor the most interesting, in that an important class of semantic phenomena can be explained only by an appeal to the other component of their meaning, as I am about to illustrate in the main text.

[16] I borrow this useful phrase from Sturgeon [1995a].

ones: S_1's only standard is that ■ all N_1's are M, while S_2's only standard is that ■ all N_2's are M; suppose further that S_1 says 'x is M' while S_2 says 'x is not M'. According to the Master Argument, the cognitivist is committed to saying that S_1 and S_2 are not disagreeing with one another, rather they are talking past one another, since on her view the content of the first remark is that x is N_1, while the content of the second is that x is not N_2. And this is a serious problem for the cognitivist, since intuitively S_1 and S_2 are disagreeing with one another. On the other hand, the non-cognitivist has the resources to vindicate our intuition that S_1 and S_2 are disagreeing with one another, says the Master Argument. On (b.H), the content of S_1's remark is that x is N_1 and ■ all N_1's are M, while the content of S_2's remark is that x is not N_2 and ■ all N_2's are M. Looking at the non-moral component of what they say does not help locate their disagreement; however (so goes the Master Argument), the moral component of what they say, which captures the only standard employed, respectively, by S_1 and by S_2, explains how they can be disagreeing with one another, showing that their disagreement is non-cognitive in nature: S_1 and S_2 are expressing the relevant non-cognitive attitude toward (say, they are commending) different things[17].

---

[17] This is one of the points where I should remind the reader (see n. 15) that the clause expressing the moral component of what S_1 and S_2 say needs to be strengthened so as to reflect the fact that each of them holds one *and only one* standard. I said earlier this can be done in either of two different ways; interestingly, given Hare's reliance on the Master Argument, it turns out that he can use only one of them, i.e. the second of the two I distinguished above. On that way of strengthening the clause in question, the content of S_1's remark is that x is N_1 and ■ all N_1's are M *and* not (■ all N_2's are M) and not (■ all N_3's are M) and … , while the content of S_2's remark is that x is not N_2 and ■ all N_2's are M *and* not (■ all N_1's are M) and not (■ all N_3's are M) and … : if the clause in question is strengthened this way, then the moral component of what S_1 and S_2 say does explain how they can be disagreeing with one another. This would not be so, if the clause in question were strengthened in the first of the two ways distinguished above, so as to have the content of S_1's remark be that x is N_1 and ■ all N_1's are M *and* ■ all M's are N_1 and the content of S_2's remark be that x is not N_2 and ■ all N_2's are M *and* ■ all M's are N_2. From now on, I'll thus assume that the way to strengthen the clause in question is the second of the two I distinguished above. A few points about my presentation of the Master Argument are worth noting. As we have just seen, what is meant to allow Hare to vindicate the intuition that S_1 and S_2 are disagreeing with one another is that, on (b.H), one is rejecting a general principle the other holds. This means that, unless Hare can account for what it is to reject a general principle, he is not in a position to claim he can vindicate the intuition that S_1 and S_2 are disagreeing with one another. As a matter of fact, this is one of the standard criticisms raised against Hare's use of the Master Argument, motivated by

As we saw, there is a class of uses of individual moral sentences for which (a.H), hence something very much like the cognitivist's account, is appropriate, according to Hare; these, however, are not the central, typical uses of individual moral sentences, and they are parasitic on the typical use of such sentences, for which the cognitivist's account is inadequate. The cognitivist's mistake is, on Hare's view, to think that, since the content of an individual moral sentence can, for some of the uses of that sentence, be specified simply by replacing the moral predicate with a non-moral one, it can be so specified for all the uses of that sentence, hence that the moral predicate in the sentence can be analyzed in terms of the non-moral predicate in question. Hare's account of individual moral sentences avoids this mistake. Given the Master Argument, this account does a better job of explaining the phenomenon of moral agreement and disagreement than does that adopted by the cognitivist who believes that moral predicates can be analyzed in terms of non-moral ones. At the same time, this account allows Hare to explain the phenomenon mentioned above, i.e. our use of individual moral sentences in order to convey non-moral information, just as well as the cognitivist: both on (a.H) and on (b.H), a sentence of the form 'x is M' ascribes a non-moral property to x, so it is no surprise that we can use it to convey non-moral information about x.

So far, in presenting Hare's account of individual moral sentences, I have been making a simplifying assumption. It is time now to drop it and consider a user S of moral language who

the worry that non-cognitivism may have trouble dealing with negation. As I noted in n. 15, though I find this worry compelling, in the present work I'll grant that non-cognitivists can answer it, since the issues raised by it are distinct from those I want to raise here, so I'll bracket this standard criticism of Hare's argument. Another standard criticism is that the cognitivist Hare is criticizing is not committed to the claims he ascribes to her, at least if she adopts a more sophisticated account of reference than the one Hare's argument seems to be relying on. Since the issues it raises are not directly related to the topic of the present work, I'll bracket this criticism as well (for the record, I find this other criticism compelling too). There is a criticism of Hare's argument that has so far gone unnoticed and that is related to my topic in the present work: I'll turn to it in section 3.5. For Hare's presentation of the Master Argument, see Hare [1952], p. 49, where he argues that the cognitivist he is criticizing is committed to the implausible claim that S_1 and S_2 are talking past one another, and pp. 148-150, where he argues that the disagreement between S_1 and S_2 is accounted for by the moral component (the evaluative meaning, as he calls it) of what they say.

has made up her mind about what direct moral beliefs to accept, for a given moral predicate 'M',

and holds more than one such belief, so that there is more than one non-moral property that can

be denoted by some non-moral predicate 'N' such that S believes that ■ all N's are M. For

simplicity, I'll consider the case in which S has two direct moral beliefs: it will be easy to

extrapolate from what I say to cases in which S has more than two such beliefs. Let us then use

'the n_S's' as shorthand for 'the non-moral properties that can be denoted by some non-moral

predicate 'N' such that S believes that ■ all N's are M' and suppose that the n_S's are two;

consider now the non-moral predicates, whatever they happen to be, denoting such properties:

suppose they are 'N_1' and 'N_2'. One way to extend Hare's account of individual moral sentences

to the case in which the n_S's are two is the following:

(a.H') The content of 'x is M' (as uttered by S) is that x is N_1 or x is N_2.

(b.H') The content of 'x is M' (as uttered by S) is that (x is N_1 or x is N_2) and ■ all N_1's

are M and ■ all N_2's are M; the content of 'x is not M' (as uttered by S) is that x is not N_1

and x is not N_2 and ■ all N_1's are M and ■ all N_2's are M[18].

---

[18] Three remarks. One, about (a.H') and the forthcoming (a.H''): both deal with negations in the same
straightforward way as (a.H) does. Two, about (b.H') and the forthcoming (b.H''): the same considerations
mentioned in n. 15 and n. 17 apply here as well. The clause ' ■ all N_1's are M and ■ all N_2's are M' (in all of its
occurrences) should read ' ■ all N_1's are M and ■ all N_2's are M *and* not (■ all N_3's are M) and not (■ all N_4's are
M) and …', but in order to make things simpler for the reader and easier for my opponent I'll frame it the way I
have just done in the main text and drop the rest, taking it as understood. Three, about both (a.H') and (b.H'): Hare
is not explicit about how exactly to extend his account of individual moral sentences to the case in which a user of
moral language holds more than one direct moral belief (or, in his terminology, general principle), and one might
suggest that what he has in mind is not quite what I ascribe to him in the text. More specifically, one might suggest
that the content of 'x is M' (as uttered by S) should be determined only by the general principle S takes to apply to x.
To illustrate using (a.H') (analogous considerations apply to (b.H'), as does my reply to the suggestion in question):
if S believes that x is N_1, the content of 'x is M' (as uttered by S) is that x is N_1 (not, as I have it, that x is N_1 or x is
N_2), while if S believes that x is N_2, the content of 'x is M' (as uttered by S) is that x is N_2. The proposed suggestion,
however, leaves important questions unanswered and faces a serious problem. As for the questions, note that this
suggestion assigns a content to 'x is M' (as uttered by S) only on the assumption that S takes a specific general
principle, among those she accepts, to apply to x, and thus believes that x is M. What, on the proposed suggestion, is
the content of 'x is M' (as uttered by S), if S either doubts whether or denies that any of those general principles

It is standard among philosophers to assume that there is also another way to extend the 'Go by the speaker's standards' strategy to the case in which the n_S's are two, one that appeals to the disjunctive property built out of the n_S's and so 'goes', strictly speaking, not 'by the speaker's standards' but by their disjunction:

(a.H'') The content of 'x is M' (as uttered by S) is that x has the disjunctive property of being $N_1 \vee N_2$.

(b.H'') The content of 'x is (is not) M' (as uttered by S) is that x has (does not have) the disjunctive property of being $N_1 \vee N_2$ and ■ all N_1's are M and ■ all N_2's are M.

In presenting Hare's account of individual moral sentences, I have only been considering a user S of moral language who has made up her mind about what direct moral beliefs to accept, for a given moral predicate 'M': what about users of moral language who have not made up their minds about this, i.e. almost all of us? Note that there are at least two respects under which most of us differ from S: on the one hand, our standards are often somewhat vague (I believe *something* like ' ■ All N_1's are M', though the non-moral property in question might not be quite N_1-ness, it might be some other property in its vicinity); on the other, even when our standards

---

applies to x, and thus either doubts whether or denies that x is M? And what is the content of 'x is M' (as uttered by S), if S believes that x is M because she takes one or the other of those general principles, though no specific one, to apply to x (a possibility that, as I pointed out in n. 7, Hare need not rule out)? Finally, what exactly is the content of 'x is M' (as uttered by S), if S takes both of those general principles to apply to x? As for the problem, suppose S initially believes that x is N_1 but not N_2, then changes her mind and comes to believe that x is N_2 but not N_1: throughout, we want to say, there is a belief that S preserves, one she can express by saying 'x is M', but the proposed suggestion does not allow us to say this. Though on the suggestion in question there is, throughout the process, a belief held by S that she can express by saying 'x is M', this is a different belief at different times, hence not one that S preserves across her change of mind. For these reasons, I propose to read Hare in a different way than does the suggestion in question.

are clear-cut, we are often open to the possibility that they not be, so to speak, exhaustive (I believe that ■ all N_1's are M and ■ all N_2's are M, but am uncertain whether that is all or there are other items on the list). As pointed out in n. 3, Hare recognizes the first point, but he is not explicit about how exactly to move from assigning content to individual moral sentences as uttered by S to assigning content to such sentences as uttered by speakers like most of us[19]. Though the gap between S and most of us is worth noting, none of the objections I'll raise against the proposals built on Hare's way of fleshing out the Assumption will rely on the existence of such a gap: I'll argue that these proposals fail *even if* one ignores this gap and assumes we have all made up our minds about what direct moral beliefs to accept.

In the previous chapter I described three different types of cognitivist proposals and called the substitution operations they rely on 'non-enriched singular substitution', 'enriched singular substitution' and 'general substitution', respectively; in each case, the substitution performed can be simple, multiple or disjunctive. The analogies between the non-cognitivist proposals I have described in this section and the first two types of cognitivist proposals described in the previous chapter are obvious, so let us use 'simple / multiple / disjunctive non-enriched singular substitution' and 'simple / multiple / disjunctive enriched singular substitution' as names also for the substitution operations that the proposals described in this section rely on.

Since the analogies are obvious, let me comment briefly on some differences between the operations of singular substitution relied on, respectively, by the cognitivist and the non-cognitivist proposals. The operations in question differ with regard to the non-moral properties in whose terms they are defined: the operations of singular substitution employed by the cognitivist are defined in terms of the *n*'s, i.e. the non-moral properties that can be denoted by some non-

---

[19] Sticking to the first point and to the possibilities mentioned in n. 3: one might try saying that 'x is M', as uttered by a speaker with vague direct moral beliefs, has a vague content; or one might try saying it has a definite content, though one that the speaker is unable to fully specify.

moral predicate 'N' such that ■ all N's are M, whereas those employed by the non-cognitivist are defined in terms of the n_S's, i.e. the non-moral properties that can be denoted by some non-moral predicate 'N' such that S believes that ■ all N's are M. This difference is due to the different ways of fleshing out the Assumption that the proposals in question are built on: the cognitivist proposals are built on i)-iv), i.e. on a metaphysical thesis, so the substitution operations they rely on are defined in terms of the non-moral properties picked out by the brute moral truths there are; the non-cognitivist proposals, on the other hand, are built on i.H)-iv.H), i.e. on an epistemological thesis, so the substitution operations they rely on are defined in terms of the non-moral properties picked out by the direct moral beliefs held by S.

This brings out a further difference worth noting: both the cognitivist and the non-cognitivist proposals assign a content (or, in the case of the former, payoff) to individual moral sentences, but in the case of the non-cognitivist proposals the assignment is relativized to the speaker, as shown by the phrase 'as uttered by S' figuring in the formulation of each of the proposals described in this section.

A third difference emerges if we focus on proposals of type B, which rely on enriched singular substitution (I'll illustrate the difference using simple substitution: the same remarks apply to multiple and disjunctive substitution as well). This difference will be important in chapter 5; for now, I'll just point it out to the reader. On type B proposals, the sentence capturing the content (or, in the case of cognitivist proposals, payoff) of an individual moral sentence contains a clause designed to have moral content, which in the case of cognitivist proposals is 'N_1-ness is the n' (supposing the predicate denoting the n is 'N_1'). In the case of non-cognitivist proposals, such a clause is not 'N_1-ness is the n_S' (even supposing the predicate denoting the n_S is 'N_1'), because 'N_1-ness is the n_S' has no moral content: it is a non-moral remark about what S

believes, given that the phrase ' ■ all N's are M' in 'the n_S' occurs within the scope of 'S believes that'. The way non-cognitivist proposals are formulated is thus by having the clause in question be ' ■ all N_1's are M'.

Since I am using the name 'singular substitution' for the operations relied on by both cognitivist and non-cognitivist proposals, I'll add the phrase '*via* reference to the *n*'s (the n_S's)' when I need to refer only to the operations of singular substitution relied on by the cognitivist (non-cognitivist) proposals. For ease of exposition I will, as I did in the previous chapter, speak of the operations of simple substitution as a limiting case of the corresponding operations of both multiple and disjunctive substitution.

So far I have been dealing only with individual moral sentences, but it is easy enough to see that the 'Go by the speaker's standards' strategy can be applied more broadly, to yield an account of all categorical moral sentences. Here is how one can apply this strategy to specify the content of all categorical moral sentences, starting from a subset of them whose content is assumed as given:

1.H)   All categorical moral sentences are either <u>built sentences</u>, i.e. their content is specified by means of other categorical moral sentences or <u>basic sentences</u>, i.e. their content is not specified by means of other categorical moral sentences.

2.H)   The content of all built moral sentences is specified by means of basic moral sentences.

3.H)   The set of basic moral sentences is the set of appropriately strong universal moral sentences that can be formulated using a non-moral predicate denoting a non-moral property.

4.H)　The content of all built moral sentences is specified, for any user S of moral language, in terms of (multiple / disjunctive) non-enriched singular / enriched singular substitution, *via* reference to the non-moral properties that can be denoted by some non-moral predicate 'N' such that S believes that ■ all N's are M[20].

I said above that theses 1.H)-4.H) provide a recipe to specify the content of all categorical moral sentences, starting from a subset of them whose content is assumed as given. The subset in question is that defined by thesis 3.H): the set of appropriately strong universal moral sentences that can be formulated using a non-moral predicate denoting a non-moral property[21]. On non-cognitivism, the content of these categorical moral sentences is specified in terms of the core non-cognitivist idea: in uttering a sentence of the form ' ■ All N's are M', a speaker S is expressing the relevant non-cognitive attitude (commendation, condemnation, etc., depending on what 'M' is) toward N's. Starting from these categorical moral sentences, 1.H)-4.H) provide a

---

[20] It is easy enough, starting from the case of individual moral sentences, to define the operations of non-enriched and enriched singular substitution *via* reference to the n_S's for the case of universal moral sentences, so I'll spare the reader the technical details. Hare does not discuss moral explanatory claims, so in articulating the 'Go by the speaker's standards' strategy I am for now focusing only on categorical moral sentences. In a short while, I'll begin to discuss non-cognitivist accommodationist proposals that extend this strategy to moral explanatory claims: if one takes these proposals into account, one can replace 'moral sentences (of the relevant type)' for 'categorical moral sentences' throughout 1.H)-4.H) and obtain a set of theses that articulate the 'Go by the speaker's standards' strategy for all moral sentences (of the relevant type) and thus match theses 1)-4) presented in chapter 2. Indeed, after the replacement in question, the first three theses in the two sets are identical (except for the mention of the payoff option in 1)-3)) and the only difference lies in the fourth thesis, i.e. in the substitution operations in terms of which the content (or, in the case of 4), payoff) of built moral sentences is specified.

[21] I should remind the reader that the modal expression 'sentences that *can* be formulated' covers both sentence that require, for their formulation, a non-moral predicate we already have and sentences that require a non-moral predicate we do not already have but can come up with.

recipe that specifies the content of all the others, as uttered by S, *via* reference to the non-moral

properties picked out by those, among the former, that S accepts[22].

Theses 1.H)-4.H) articulate in some detail the strategy summarized by the slogan 'Go by

the speaker's standards'. As I noted above, given theses i.H)-iv.H) this is a strategy that should

strike you as a plausible way to go, if your aim is to replace the moral predicate occurring in a

categorical moral sentence with a non-moral one, while at the same time denying that the moral

predicate in question can be analyzed in terms of the non-moral one. Remember the problem in

response to which Hare sets out to achieve this aim: explaining our use of categorical moral

sentences to convey non-moral information[23]. This feature of the way in which we use

categorical moral sentences presents Hare with a problem because it is a feature that seems to

call for a cognitivist treatment: in particular, as I noted above, it is a feature that is easy to

account for, for the cognitivist who believes that moral predicates can be analyzed in terms of

non-moral ones. Hare's problem is not the only problem of this kind that non-cognitivists face,

since there are other features of our moral practice that also seem to call for a cognitivist

---

[22] Two remarks. One: the core non-cognitivist idea can at most work as an account of what a speaker S is doing in uttering a sentence of the form ' ■ All N's are M' on the assumption that S believes that ■ all N's are M. More needs to be said in order to go from here to a specification of the content of such a sentence: for example, what is the content of ' ■ All N's are M', as uttered by S, if S doubts whether or denies that ■ all N's are M? This is one way to raise the Frege-Geach problem for non-cognitivism (see Geach [1965]). Here I only have room to mention this problem without pursuing it, since it would make for the topic of a different work. I'll have occasion to refer to the Frege-Geach problem again in what follows. Two: the set of basic moral sentences is defined by 3.H) as the set of appropriately strong universal moral sentences that can be formulated using a non-moral predicate denoting a non-moral property; however, the content of built moral sentences, as uttered by a speaker S, is specified *via* reference to the non-moral properties picked out only by those, among the sentences in that set, that S accepts. So, it might be asked, why not have 3.H) define the set of basic moral sentences as made up only of those, among the appropriately strong universal moral sentences that can be formulated using a non-moral predicate denoting a non-moral property, that S accepts and include those, among such sentences, that S does not accept in the set of built moral sentences? Because the content, and logical relations, of a sentence, as uttered by S, do not change depending on whether S does or does not accept it, but this is exactly the result one would get, if one were to define the set of basic moral sentences as suggested (the reasoning showing this is the same as that near the end of 2.3, where I considered an analogous question about theses 1)-4)).

[23] When I introduced this problem at the beginning of this section, I illustrated it only with reference to individual moral sentences, but this is clearly not the only class of categorical moral sentences we use in order to convey non-moral information: to stick to the same example I used at the beginning of this section, if two parsons are talking about the Smith family, one can tell the other 'All the members of the Smith family are good people' in order to convey the information that they all go to church and obey a certain set of Biblical injunctions.

treatment; in the non-cognitivist tradition, Hare's strategy of 'going by the speaker's standards' has been applied more broadly than I have illustrated so far, to address these other problems as well. Faced with features of our use of moral expressions that seem to call for a cognitivist treatment, the standard non-cognitivist strategy has been to claim that, although the function of moral expressions is to express non-cognitive attitudes and so they do not admit of a cognitivist treatment, they can be replaced, in sentences uttered by a speaker S, with non-moral expressions determined by S's standards, without being analyzable in terms of these non-moral expressions. In the non-cognitivist tradition, this has been called, following Hare, the doctrine of the 'descriptive meaning' of moral expressions. One application of this general recipe has been to our main topic in the present work, i.e. moral explanatory claims[24].

Consider claims of the form 'M-ness causes E', where 'M' is a moral predicate. It is a feature of our moral practice that we offer, discuss, and accept such claims. This feature seems to call for a cognitivist treatment: in particular, it seems to call for a referent for the noun 'M-ness'; it thus presents you with a problem, if you are a non-cognitivist who wants to accommodate our practice with regard to moral explanatory claims. A way out of this problem would be for you to propose that the moral noun 'M-ness' in 'M-ness causes E' can be replaced with the non-moral noun 'N-ness', for some non-moral predicate 'N', even if the moral predicate 'M' cannot be analyzed in terms of the non-moral predicate in question. Once again, given i.H)-iv.H) it should strike you as a plausible strategy to 'go by the speaker's standards'. Then, starting from the

---

[24] In the text I ascribe the account of categorical moral sentences articulated by 1.H)-4.H), together with the strategy of 'going by the speaker's standards' and the related doctrine of the descriptive meaning of moral expressions, to Hare, but it might claimed that these ought to be traced back to Stevenson [1944] instead. Though much can be said on behalf of this claim, there are a few reasons for ascribing these views to Hare. One reason is that the account of categorical moral sentences in question provides at most only the second of Stevenson's two "patterns of analysis", as he calls them, or ways of specifying the content of such sentences, and the strategy followed by Stevenson's first pattern cannot be described as one of 'going by the speaker's standards'. Another reason is that Hare provides a systematic defense of the reading of the Assumption articulated by i.H)-iv.H), which is what makes the account of categorical moral sentences in question, and the strategy of 'going by the speaker's standards', plausible. I'll mention a third reason in section 3.6.

definitions given above, you can define the operations of non-enriched and enriched singular substitution *via* reference to the n_S's for the case of moral explanatory claims and put forward a proposal about the content of such claims that allows you to solve your problem. As I mentioned above, this is the strategy adopted by Blackburn [1991a], who puts forward just such an accommodationist proposal.

Blackburn's suggestion is that the non-cognitivist can accommodate our practice with regard to moral explanatory claims by taking such claims to "point[…] downwards to the properties upon which the moral verdict depends" (p. 12). A speaker S who says something of the form 'M-ness causes E' will, says Blackburn, have her own standards for applying the moral predicate 'M', standards which pick out certain non-moral properties; the non-cognitivist should propose that what 'M-ness causes E' does, as uttered by S, is to say, of those non-moral properties, that *they* cause E[25]. In my terminology, Blackburn's suggestion is that the non-cognitivist can accommodate our practice with regard to moral explanatory claims by means of an accommodationist proposal that relies on non-enriched singular substitution. Here then is Blackburn's proposal for the case of a user S of moral language who has made up her mind about what direct moral beliefs to accept, for a given moral predicate 'M', and holds only one such belief (supposing that the predicate denoting the n_S is 'N$_1$'):

---

[25] In presenting his proposal, Blackburn says (pp. 12-13) it is analogous to F. Jackson and P. Pettit's theory of 'program explanation' about explanatory claims invoking higher order properties, such as those offered by the special sciences and common sense (see Jackson & Pettit [1990]). Despite some undeniable similarities between the two proposals, the comparison is, I think, misleading in a way that it is important to point out, in order to avoid misunderstanding Blackburn's proposal. It is *not* part of Jackson and Pettit's view that explanatory claims invoking higher order properties, as uttered by a given speaker, ascribe causal efficacy to the lower order properties *that speaker takes* to underlie the higher order properties in question, whereas Blackburn's proposal is precisely that the non-moral properties that moral explanatory claims, as uttered by a given speaker, ascribe causal efficacy to are those picked out by the speaker's moral standards, as shown by the way in which Blackburn illustrates his proposal right after comparing it to Jackson and Pettit's theory: "If you tell me that injustice caused the revolution, I understand that there is some property which you take to give rise to injustice, and that caused the revolution. I must, in my own assessment, separate the truth of the causal story you are pointing towards, from my own verdict on whether it amounts to injustice. Knowing your devotion to Rawls I may have a good idea of the first, and I may agree […]. Perhaps I would not myself call that unjust, but I can assent to the explanation without endorsing the verdict you pass on the feature" (p. 13).

(A.H) The content of 'M-ness causes E' (as uttered by S) is that N_1-ness causes E.

And here are two different ways of extending Blackburn's proposal to the case of a user S of moral language who has made up her mind about what direct moral beliefs to accept, for a given moral predicate 'M', and holds two such beliefs (supposing that the predicates denoting the n_S's are 'N_1' and 'N_2') -it is easy to extrapolate from this case to cases in which S has made up her mind and holds more than two such beliefs-:

(A.H') The content of 'M-ness causes E' (as uttered by S) is that N_1-ness causes E and N_2-ness causes E;

(A.H'') The content of 'M-ness causes E' (as uttered by S) is that the disjunctive property of being $N_1 \lor N_2$ causes E[26].

---

[26] Though the theses in the text formulate Blackburn's proposal, I keep the letter 'H' in my names for them because they are built on Hare's way of fleshing out the Assumption. I should mention explicitly that Blackburn does not say he derives his proposal from Hare: my point is that his proposal can be seen as the application to moral explanatory claims of a more general, standard non-cognitivist strategy, one that can be traced back to Hare's account of categorical moral sentences (it is worth mentioning here that Blackburn [1991a] is a reply to Sturgeon [1986b], where Sturgeon puts forward, on behalf of the non-cognitivist and in order to criticize it, a proposal very close to Blackburn's: when he introduces that proposal, Sturgeon appeals explicitly to the non-cognitivist doctrine of the descriptive meaning of moral expressions). (A.H), (A.H'), and (A.H'') all deal with negations in the same straightforward way as (a.H) does. Blackburn is not explicit about how exactly to extend his proposal to the case of a user of moral language who has made up her mind about what direct moral beliefs to accept, for a given moral predicate 'M', and holds more than one such belief and I do not claim that either (A.H') or (A.H'') is explicitly endorsed by him; rather, my aim here is to distinguish and lay out the different options he has available to extend his proposal to cover such a case. As a matter of fact, Blackburn [1991a] puts forward not one but three accommodationist proposals, the second of which is the one I present in the text. However, Blackburn himself points out that his first proposal fails for some cases; and his third proposal (which he calls a "more speculative strategy" which "we do not have to rely upon" –p. 13-) is somewhat vague and I am not sure I see how it differs, at least for our purposes, from the second. For a writer who finds a difference here, see Gibbard [2003], ch. 10: after putting forward an accommodationist proposal that differs from Blackburn's second proposal, Gibbard claims (p. 207) it is equivalent to Blackburn's third proposal. I'll discuss Gibbard's proposal starting from the next chapter.

Blackburn's proposal is the first non-cognitivist accommodationist proposal that we need to discuss. It results from the application to moral explanatory claims of a standard non-cognitivist strategy, what I called the 'Go by the speaker's standards' strategy, by means of which non-cognitivists attempt to mimic the cognitivist 'Go by the M-making properties' strategy, and more specifically the proposals of type A and type B available to a cognitivist, that I described in chapter 2. In the next three sections, I'll give three reasons for thinking that Blackburn's proposal fails, and draw some more general lessons about both the 'Go by the speaker's standards' strategy and non-cognitivism.

I'll proceed as follows. Two of my objections to Blackburn's proposal have their source in theses i.H)-iv.H): in sections 3 and 4, I'll point out two different problems with those theses and argue that Blackburn's proposal, which is built on them, inherits them (the same, I'll argue, holds for Hare's account of categorical moral sentences). Then, in section 5, I'll argue there is a further problem with Blackburn's proposal, one that arises even if one grants that theses i.H)-iv.H) are true (once again, I'll argue that the same holds for Hare's account).

3. <u>First objection to the 'Go by the speaker's standards' strategy</u>

Both Blackburn's proposal about moral explanatory claims and Hare's account of categorical moral sentences are built on Hare's way of fleshing out the Assumption, articulated by theses i.H)-iv.H). But theses i.H)-iv.H) are false. There are two ways to see that they are false: one, which I'll explore in this section, is by looking at their consequences with regard to the way in which we develop our moral views; the other, which I'll explore in the next section, is by looking at what they require in order for us to hold moral views at all.

Suppose Mary has a fully worked out theory about what things are morally obligatory: in particular, she has made up her mind about what direct moral beliefs to accept. There are only two things Mary thinks are obligatory, namely keeping one's promises and returning help to those who have helped one: in the terminology of theses i.H)-iv.H), she has two direct moral beliefs, i.e. that ■ any action that consists in keeping one's promises is obligatory and that ■ any action that consists in returning help to those who have helped one is obligatory. Mary is walking home, holding her bag of groceries and thinking of what happened last week, when Mary's neighbor saved her daughter from a burglar about to shoot her. Approaching the crosswalk by her house, she sees her neighbor's son about to cross the street; he is texting, so he does not see the red light or the big truck that is coming. There is nobody else around and yelling will not help, as the boy is wearing headphones, so Mary drops the bag she is holding and runs toward the boy, pulling him off the street just in time to save his life. The bottle of French wine she has just bought is lost, but Mary has no doubts: 'It was my duty to save him', she thinks while she turns her eyes from the bag on the street to the boy, who looks fine, though a bit frightened, and gazes at her with a grateful smile. To her surprise, however, Mary realizes she was wrong: the boy is not her neighbor's son. As a matter of fact, he is a perfect stranger to her: she does not know him and has never had anything to do either with him or with anyone connected to him. Quite generally, her action of saving his life cannot be subsumed under either of her two direct moral beliefs. Mary thus needs to revise some of her moral beliefs: what are her options?

Intuitively, there are at least two ways Mary can go. She could, on the one hand, stick to her moral theory and, on that basis, revise her belief about her action of saving the boy's life. Mary has so far been thinking that such an action was obligatory, but what she has just learnt could lead her to change her mind about this. It could lead her to conclude that her action was

maybe a very nice thing to do, but it was not morally obligatory. Alternatively, she could stick to her belief about her action and, on that basis, revise her moral theory. There are only two things Mary has so far thought are obligatory, keeping one's promises and returning help to those who have helped one, but her recent experiences could end up changing this. They could lead her to conclude that her action of saving the boy's life was indeed obligatory and thus to revise her moral theory, for instance by adding a third item on her list: say, providing help to someone in need, at least when one can do so at little cost to oneself. Which of these two options is more reasonable, whether to revise her theory or her belief about her action, will depend on further details of Mary's evidential situation, but both of these options are available to her, after she learns that her action does not consist either in keeping one of her promises or in returning help to someone who has helped her.

The problem with theses i.H)-iv.H) is that they deny this. If they are true, there is only one way Mary can go: stick to her moral theory and revise her belief about her action. On i.H)-iv.H), Mary can only hold the individual moral belief that that action is obligatory by inferring it from other beliefs she has, and ultimately by inferring it *via* subsumption from her direct moral beliefs, plus appropriate non-moral beliefs. But Mary has just rejected the non-moral beliefs she would need in order to infer it this way: her action of saving the boy's life does not, she has just learnt, consist either in keeping one of her promises or in returning help to someone who has helped her. So given i.H)-iv.H) Mary has no option but to give that belief up. If, after learning what she has just learnt, Mary sticks to it, she develops her moral views in a way that is, epistemologically speaking, a bad way of doing it.

Both of the options I described above, I am saying, are available to Mary, but i.H)-iv.H) only allow for the first of them. In reply, friends of i.H)-iv.H) can deny there is anything else to

allow for; or they can agree there is and claim that they can, after all, make room for it. Starting with the former line of reply, friends of i.H)-iv.H) can deny that someone like Mary is likely to revise his/her theory in a case like the one I presented above. Mary, after all, has a fully worked out theory about what actions are morally obligatory. Presumably, she has thought long and hard about it and has come to accept it after careful reflection. In the case we are considering, she is thus unlikely to give up her carefully thought out theory on the basis of a belief about a particular action. She will rather stick to her guns: she will keep her moral theory and, on that basis, revise her belief about her action.

In response, I would distinguish between two claims. One claim is that someone like Mary is unlikely to revise his/her theory in a case like the one we are considering. My own views on the matter are much less definite than this. It seems to me that, when forced to choose between a theory and a belief about a particular case, we sometimes go one way, sometimes the other: a lot, I would say, depends on the details. Suppose, however, that friends of i.H)-iv.H) are right. Indeed, suppose that someone like Mary *will never* revise his/her theory in a case like the one we are considering. A different claim is that Mary *can* –it is epistemologically legitimate for her to- revise her theory in my example. *This* is the claim my argument relies on. My argument is neutral as to which option Mary ends up choosing: as I said above, which option to choose will depend on further details of Mary's evidential situation; my point is rather that both of the options I described above are available to her, even after learning what she has just learnt. Said otherwise, the fact that her moral theory cannot account for her belief about her action does not, all by itself, determine that such a belief has to go. Even supposing that friends of i.H)-iv.H) are right, indeed even supposing that Mary will not revise her theory, this is because her theory cannot account for that belief *and* her theory is so good: it is so carefully thought out. The fact

(supposing it a fact) that Mary will not choose the second option I described above does not mean that that option is not available to her: it is thus a problem for i.H)-iv.H) that they do not allow for it.

Or do they? Turning now to the second line of reply mentioned earlier, what if Mary changes her direct moral beliefs, say by adding a new one? Suppose it goes like this. After learning that her action does not consist either in keeping one of her promises or in returning help to someone who has helped her, Mary changes her moral theory, adding one more direct moral belief to the two she has held so far: say, that ■ any action that consists in helping someone in need at little cost to oneself is obligatory. Having revised her theory this way, Mary now can, consistently with i.H)-iv.H), hold the belief that her action was obligatory, since she once again has a direct moral belief, and an appropriate non-moral belief, from which she can ultimately infer that belief *via* subsumption. So Mary sticks to her belief about her action and revises her theory: doesn't this show that i.H)-iv.H) do allow for the second option I described above?

No, and the way to see this is by asking why Mary changes her theory: why does she add one more direct moral belief to the two she has held so far? If *this* is what happens, then it is not her belief that her action was obligatory that motivates Mary's revision of her theory: she does not revise her theory *by relying on* that belief. For the reasons given above, on i.H)-iv.H) that belief is one she cannot hold unless she infers it from direct moral beliefs different from those she has, so it cannot be what Mary relies on in changing the direct moral beliefs she has: though Mary can hold that belief once she has reached her new theory, she cannot rely on it in order to get there. To put it as I did when describing the second option: if this is what happens, it is not the case that Mary sticks to her belief about her action and, *on that basis*, revises her moral

theory. Indeed, if this is what happens, Mary revises her theory for apparently no reason at all: on this account of the change in her moral beliefs, Mary is still developing her moral views in an epistemologically objectionable way. On i.H)-iv.H), this is what follows if Mary sticks to her belief that her action was obligatory. But this consequence is false, and so are i.H)-iv.H).

Friends of i.H)-iv.H) might try a different route to make room for the second option I described above. They might suggest that what happens, if Mary sticks to her belief about her action and revises her theory by adding a third item on her list of obligatory things, is the following. Mary already held the direct moral belief that, say, ■ any action that consists in helping someone in need at little cost to oneself is obligatory, but this is a belief she held only tacitly or implicitly: the process of theory revision she goes through consists in a more explicit articulation and fuller systematization, precipitated by her recent experiences, of a tacit belief she already held. There is nothing here to trouble friends of i.H)-iv.H): on those theses, the belief that her action is obligatory is one Mary cannot hold unless she infers it from her direct moral beliefs, but Mary does have a direct moral belief from which she infers that belief, albeit one she has so far held only tacitly. They can agree that Mary revises her theory by relying on her belief about her action, noting only that, in accordance with i.H)-iv.H), the evidential foothold that belief provides her with in this process ultimately derives from one of the direct moral beliefs held -so far only tacitly- by her.

In order to see why this reply fails, it is worth remembering my opponent's overall position. My opponent's aim is to rely on theses i.H)-iv.H) in order to put forward proposals about the content of moral sentences. Her idea is that, given i.H)-iv.H), it should strike you as plausible that the content of a moral sentence, as uttered by a given speaker, can be specified *via* reference to the non-moral properties picked out by the direct moral beliefs held by that speaker.

93

It is worth pointing out here that, in the case of a speaker who holds some direct moral beliefs only tacitly, this should strike you as a plausible idea only if the direct moral beliefs taken into account in order to specify the content of a moral sentence, as uttered by that speaker, include her tacit, in addition to her explicit, ones. Consider, for example, an individual moral sentence of the form 'x is M'. Suppose i.H)-iv.H) are true and I hold only two direct moral beliefs, both of them explicitly: say, that ∎ all N_1's are M and that ∎ all N_2's are M. I could believe that x is M either because I believe it is N_1 or because I believe it is N_2 (or because, though I believe neither of these two things, I believe it has one or the other of those two properties): in specifying the content of 'x is M', as uttered by me, *via* reference to the non-moral properties picked out by the direct moral beliefs I hold, you had thus better take into account the non-moral properties picked out by both of my direct moral beliefs (which is what Hare's account of categorical moral sentences does), lest your proposal yield implausible results. Would it make any difference if I held one of those two direct moral beliefs explicitly and the other only tacitly? No: everything I have just said would still hold true. In putting forward her proposals about the content of moral sentences, my opponent is thus committed to treating the distinction between a speaker's explicit and tacit direct moral beliefs as irrelevant: in assigning content to moral sentences, as uttered by a given speaker, she is committed to taking into account also that speaker's tacit direct moral beliefs, in addition to her explicit ones.

One way of doing this is the following: ignore speakers who hold tacit direct moral beliefs and formulate the proposals in question considering only speakers whose direct moral beliefs are all explicit (once you have formulated the proposals this way, it is easy to extend them to speakers who hold tacit direct moral beliefs: just say about a speaker who holds certain direct moral beliefs, some of them tacitly, the same thing the proposals say about one who holds the

same direct moral beliefs, all of them explicitly). This is how I proceeded in section 2 when formulating my opponent's proposals: I did not distinguish between explicit and tacit direct moral beliefs; I introduced the notion of a user of moral language who has a fully worked out theory about what things are M, for a given moral predicate 'M', and in particular has made up her mind about what direct moral beliefs to accept; and I formulated my opponent's proposals as claims about the content of moral sentences as uttered by such a speaker, i.e. (among other things) a speaker whose direct moral beliefs are all explicit. The dialectical situation is therefore the following: my opponent is committed to treating the distinction between explicit and tacit direct moral beliefs as irrelevant; in order to do this, we are ignoring speakers who hold tacit direct moral beliefs and considering only speakers whose direct moral beliefs are all explicit; I introduce Mary to raise an objection against my opponent; she replies to it by suggesting that Mary holds a tacit direct moral belief. Clearly, this move does not address my objection: the reply we are considering fails because it goes against the supposition, made when introducing Mary, that she has a fully worked out theory about what things are morally obligatory, and in particular has made up her mind about what direct moral beliefs to accept.

The one I have just described is not the only way in which one can formulate my opponent's proposals so as to fulfill her commitment to treating the distinction between explicit and tacit direct moral beliefs as irrelevant, and it is important to see that nothing in my response hinges on how one formulates the proposals in question. A different way in which one can frame them is the following: formulate the proposals considering also speakers who hold tacit direct moral beliefs, in addition to speakers whose direct moral beliefs are all explicit, and frame the proposals so that they say, about a speaker of the former kind, the same thing they say about one of the latter, provided the direct moral beliefs held by the two speakers are the same. One could

actually do this even while keeping my formulations in terms of a speaker who has made up her mind about what direct moral beliefs to accept, provided one stipulated that the phrase 'a speaker who has made up her mind about what direct moral beliefs to accept and holds n such beliefs' should not be taken to exclude a speaker who holds tacit direct moral beliefs and should be used in such a way that a speaker's tacit direct moral beliefs are also taken into account, in addition to her explicit ones, in determining what number n is. But note: even if my opponent's proposals were formulated this way, the reply we are considering would still fail. The dialectical situation would be the following: my opponent is committed to treating the distinction between explicit and tacit direct moral beliefs as irrelevant; in order to do this, we are stipulating that when we say of a speaker that she 'has made up her mind about what direct moral beliefs to accept and holds n such beliefs' we are already counting her tacit direct moral beliefs, if she holds any, in addition to her explicit ones; I introduce Mary to raise an objection against my opponent; she replies to it by suggesting that Mary holds an additional direct moral belief, a tacit one. Once again, this move does not address my objection: it goes against the supposition, made when introducing Mary, that she has made up her mind about what direct moral beliefs to accept and holds two such beliefs.

So how one formulates my opponent's proposals really makes no difference with respect to the reply under consideration. In section 2 I chose the first of the two ways I have just described, and I will stick to it throughout the present work, because it makes exposition simpler. But no matter how one chooses to formulate my opponent's proposals, it remains true that the reply we are considering fails. My opponent is committed to treating the distinction between a

speaker's explicit and tacit direct moral beliefs as irrelevant: no surprise that she cannot, in response to my objection, avail herself of a reply that has that distinction do substantial work[27].

Having defended my argument against various possible replies, let me now state it in a more general and abstract way. Consider a user S of moral language who has a fully worked out theory about what things are M, for a given moral predicate 'M': in particular, S has made up her mind about what direct moral beliefs to accept. Suppose there are only two sorts of things S thinks are M: in the terminology of theses i.H)-iv.H), S has two direct moral beliefs, say that ■ all N_1's are M and that ■ all N_2's are M. Suppose further that S holds an individual moral belief about an object x: she thinks x is M. Finally, suppose S's non-moral beliefs about x are such that her moral belief about it fits her moral theory: maybe S believes that x is N_1, maybe she believes it is N_2, or maybe she believes neither of these two things but thinks x has one of these two non-moral properties. Imagine now S finds out that her non-moral beliefs about x are wrong: x, she realizes, is neither N_1 nor N_2. Having learnt this, S needs to revise some of her moral beliefs: what are her options? If i.H)-iv.H) are true, then for the reasons given above there is only one way S can go: stick to her moral theory and revise her belief that x is M. But this consequence of i.H)-iv.H) is false: giving this belief up is not the only option available to S. There is also another way she can go: stick to her belief about x and revise her moral theory. *Contra* i.H)-iv.H), even after learning that x is neither N_1 nor N_2, S can still end up sticking to her belief that x is M while developing her moral views in an epistemologically respectable way[28].

---

[27] I thank Karen Bennett and Nick Sturgeon for very helpful discussions of the argument presented in this section and possible replies to it.

[28] For simplicity's sake, in the text I state my argument in terms of an individual moral belief held by S, but it is easy enough to see that it can be stated, *mutatis mutandis*, in terms of any categorical moral belief held by S that counts, on i.H)-iv.H), as an inferred belief, be it individual or universal.

In raising my objection to i.H)-iv.H) I have simplified the description of my examples, so as to make the important details evident. The source of the problem can be described as follows. According to i.H)-iv.H), evidential support within one's moral belief system ultimately flows only one way, from the more to the less general: specifically, from one's direct moral beliefs – *via* appropriate non-moral beliefs- to the other categorical moral beliefs one subsumes under them. That this is false is shown by those cases in which non-moral evidence forces one to abandon the non-moral beliefs in question and thus prompts a revision in one's moral beliefs. The fact that one's options include that of sticking to one's less general moral beliefs and revising one's moral theory shows that, *contra* i.H)-iv.H), the former are (at least to a degree) evidentially independent of the latter. If evidential support did flow as described by i.H)-iv.H), then once one's inferred beliefs have been cut off from one's direct beliefs, due to one's rejection of the relevant non-moral beliefs, one's only option would be to stick to one's theory and revise one's less general beliefs.

As I say, in raising my objection I have simplified the description of my examples, but it should be clear that they point to a very common procedure in actual moral inquiry. This becomes even clearer once one notes that there is a feature of them that is not essential to my argument. My examples involve a user of moral language who has both a fully worked out moral theory and an individual moral belief about a certain object, and who is forced to revise some of her moral beliefs by a change in her non-moral beliefs. Go back, however, to our speaker S whose theory is made up of only two direct moral beliefs, i.e. that ■ all N_1's are M and that ■ all N_2's are M, and suppose this time that she has not thought about object x, so in particular she does not believe that x is M. Suppose you draw S's attention to x in order to challenge her moral theory: x is neither N_1 nor N_2, and you and S both acknowledge this, still (so goes your

challenge) isn't it M? Faced with your challenge, there is only one way S can go, if theses i.H)-iv.H) are true. She cannot accept the moral belief you are inviting her to accept, i.e. the individual moral belief that x is M: on i.H)-iv.H), this is a belief S can only hold by inferring it from other beliefs she has, and ultimately by inferring it *via* subsumption from her direct moral beliefs, plus appropriate non-moral beliefs, but S does not hold the non-moral beliefs she would need in order to infer it this way, since she agrees with you that x is neither N_1 nor N_2. So given i.H)-iv.H) S's only option is to stick to her moral theory and refuse to accept the belief that x is M. If S, who acknowledges that x is neither N_1 nor N_2, accepts the moral belief you are inviting her to accept and so revises her moral theory, she develops her moral views in an epistemologically objectionable way. Since this is exactly the way you are inviting her to go, this means that your challenge is really no legitimate challenge at all.

But it clearly is a legitimate challenge. Consider Mary again: imagine she buys no bottle of French wine and saves nobody's life. This time suppose you do: yesterday you were walking home holding your bag of groceries, when you saw a boy about to cross the street … , so you dropped your bag and ran toward the boy, pulling him off the street just in time to save his life. The boy was a perfect stranger to you, and your action of saving his life did not consist either in keeping one of your promises or in returning help to someone who had helped you. Today you are talking to Mary and, surprised by her moral views, you invite her to consider what you did yesterday: your action did not consist in either of those two things, you two agree, still wasn't it morally obligatory? If theses i.H)-iv.H) are true, then if Mary were to accept the moral belief that your action was obligatory and so revise her moral theory, she would be developing her moral views in an epistemologically objectionable way, but clearly she would not: challenging a moral theory the way you do in my example is entirely common in actual moral inquiry, and there is

99

nothing epistemologically objectionable about it. What you do in my example is to point out that, given the non-moral facts, the theory in question is unable to account for certain less general moral beliefs that, in your view, we should not give up: this is a perfectly legitimate way of challenging a moral theory. Theses i.H)-iv.H) entail it is not, so they must be false[29].

Since Blackburn's account of moral explanatory claims and Hare's account of categorical moral sentences are built on theses i.H)-iv.H), it is not surprising that the problem I have pointed out with these theses is mirrored by parallel problems with both accounts. As we saw, both accounts specify the content of the relevant class of sentences one way (i.e. in terms of simple substitution) for the case of a speaker S who has made up her mind about what direct moral beliefs to accept, for a given moral predicate 'M', and holds only one such belief; the accounts can then be extended in two different ways (relying on either multiple or disjunctive substitution) to the case in which S has made up her mind and holds more than one such belief. In pressing my

---

[29] Though the example in the text involves a user of moral language who has a fully worked out moral theory and someone else who draws her attention to a certain object in order to challenge that theory, it is worth pointing out that there is also an intra-personal version of the example: suppose you say nothing to Mary and she herself questions her moral theory, wondering 'Wasn't that morally obligatory?' after she sees you saving the boy's life. If i.H)-iv.H) are true, there is only one way Mary can answer her question: her only option is to stick to her moral theory and answer 'No'; if she answers 'Yes' and so revises her theory, she develops her moral views in an epistemologically objectionable way. Said otherwise, asking herself that question is not a legitimate way for Mary to question her moral theory. But it clearly is: questioning one's moral theory the way Mary does in this version of the example is entirely common in actual moral inquiry and there is nothing epistemologically objectionable about it. Once again, it might be suggested on behalf of i.H)-iv.H) that what happens, if Mary does decide to accept the belief that your action was obligatory and so revise her theory, is the following. After you draw her attention to your action of saving the boy's life (or she sees you saving the boy's life), Mary changes her moral theory, say by adding one more direct moral belief to the two she has held so far: for instance, the belief that ■ any action that consists in helping someone in need at little cost to oneself is obligatory. Having revised her theory this way, Mary now can, consistently with i.H)-iv.H), accept the belief that your action was obligatory, since she now has a direct moral belief, and an appropriate non-moral belief, from which she can ultimately infer that belief *via* subsumption. So Mary accepts the belief that your action was obligatory and revises her theory: nothing in i.H)-iv.H) prevents her from doing this. Ask, however, why Mary changes her theory: why does she add one more direct moral belief to the two she has held so far? If *this* is what happens, then it is not the belief that your action was obligatory that motivates Mary's revision of her theory: she does not revise her theory *by relying on* that belief. On i.H)-iv.H), that belief is one she cannot hold unless she infers it from direct moral beliefs different from those she has, so it cannot be what Mary relies on in changing the direct moral beliefs she has: though Mary can hold that belief once she has reached her new theory, she cannot rely on it in order to get there. Indeed, if this is what happens, Mary revises her theory for apparently no reason at all: on this account of the change in her moral beliefs, Mary is still developing her moral views in an epistemologically objectionable way.

objection against each account, I'll thus focus on a speaker who holds only one direct moral belief: showing there is a problem here is enough to show there is a problem with the account, no matter how it is extended to cover speakers who hold more than one such belief[30].

The problem with Hare's account is at this point easy to see. Consider a user S of moral language who has a fully worked out theory about what things are M, for a given moral predicate 'M': in particular, S has made up her mind about what direct moral beliefs to accept and holds only one such belief, say that ■ all N_1's are M. Suppose S also holds an individual moral belief about a certain object x: she thinks x is M. According to Hare's account of categorical moral sentences, S's belief that x is M amounts either to the belief that x is N_1 or to the belief that x is N_1 and ■ all N_1's are M (depending on whether one relies on non-enriched or enriched singular substitution). Imagine now S finds out that x is not N_1. Having learnt this, S needs to revise some of her moral beliefs: what are her options? Given Hare's account of categorical moral sentences, S cannot hold her belief that x is M any longer: on Hare's account, that belief amounts to the belief that x is N_1 (and ■ all N_1's are M), and S has just rejected the belief that x is N_1. So on Hare's account, after learning that x is not N_1 there is only one way S can go: stick to her moral theory and revise her belief that x is M. But as pointed out above this is not the only option available to S. There is also another way she can go: stick to her belief about x and revise her

---

[30] The same problem I am about to point out can also be raised for the case of a speaker who holds more than one direct moral belief, though for simplicity's sake I won't do it. In pressing my objection, I'll consider a speaker who has both a fully worked out moral theory and a moral belief of the relevant type, and who revises some of her moral beliefs because of new non-moral evidence; it is easy enough to see that the same objection can also be pressed by considering a speaker who has a fully worked out moral theory and someone else who challenges that theory by inviting her to accept a moral belief of the relevant type (or a speaker with a fully worked out moral theory who questions her own theory by wondering whether to accept a moral belief of the relevant type).

moral theory. Just like theses i.H)-iv.H), Hare's account of categorical moral sentences rules this other option out, so it must be false[31].

The problem I pointed out with theses i.H)-iv.H), which deal with categorical moral beliefs, is thus mirrored by a parallel problem with Hare's account, which is an account of categorical moral sentences. Blackburn's account, on the other hand, is an account of moral explanatory claims, and theses i.H)-iv.H) are silent about such claims. Interestingly, Blackburn's account too inherits the problem I pointed out with those theses, as shown by Sturgeon [1991], from which the two examples I am about to present are adapted. Consider again our speaker S who has a fully worked out theory about what things are M, for a given moral predicate 'M': in particular, S has made up her mind about what direct moral beliefs to accept and holds only one such belief, say that ■ all N_1's are M. Suppose S also accepts a moral explanatory claim: she thinks M-ness causes a certain type of effects E. According to Blackburn's account of moral explanatory claims, S's belief that M-ness causes E amounts to the belief that N_1-ness causes E. Imagine now S finds out that N_1-ness does not cause E. Given Blackburn's account, after learning this S cannot hold her belief that M-ness causes E any longer: on that account, after S finds out that N_1-ness does not cause E, she has no option but to give her moral explanatory belief up. But this seems wrong: this is not the only option available to S. Consider the following two examples.

One way our speaker S could find out that N_1-ness does not cause E is by observing that N_1 things are not followed by effects of type E[32]. Suppose Donald, a father of three young

---

[31] For simplicity's sake, in the text I state my objection to Hare's account in terms of an individual moral belief held by S, but it is easy enough to see that it can be stated, *mutatis mutandis*, in terms of any categorical moral belief held by S that counts, on 1.H)-4.H), as built, be it individual or universal.

[32] Plus a bunch of other things, such as (to mention just one) that no interfering or counterbalancing factors are at play that prevent the effects in question from occurring. For simplicity's sake, I'll omit such qualifications and take them as understood in what follows. Analogous considerations apply to the second of my two examples as well.

children, has a fully worked out theory about what it takes to treat people with decency and humanity. There is only one kind of education that Donald thinks will do that with one's children, one that combines various specific factors together: for brevity's sake, one that has the (quite complex) property of being N_1. Suppose Donald also has the following moral explanatory belief: a decent and humane upbringing makes children happy. So he sees to it that his three children receive an education that is N_1, hoping this will make them grow up happy; to his dismay, however, he sees them growing up all quite sad instead. On Blackburn's account of moral explanatory claims, there is only one option available to Donald after he sees this: stick to his moral theory and revise his moral explanatory belief. But this is not the only option available to him. There is also another way Donald can go: stick to his moral explanatory belief and revise his moral theory. There is only one kind of education Donald has so far thought to be decent and humane, one that is N_1, but his experience with his three children could end up changing this. It could lead him to conclude that, since a decent and humane education makes children happy, an education that is N_1 is not, after all, decent and humane. Which of these two options is more reasonable, whether to revise his theory or his moral explanatory belief, will depend on further details of Donald's evidential situation, but both of these options are available to him, even after learning that an education that is N_1 does not make children happy.

A different way our speaker S could find out that N_1-ness does not cause E is by observing that N_1 things are followed by effects of type E, but their being N_1 plays no role in bringing about those effects. Suppose Judy has a fully worked out theory of justice. It is only institutions with the property of being N_1 that Judy thinks are just. Suppose Judy also has the moral explanatory belief that justice makes institutions stable: justice, so thinks Judy, is a stabilizing feature, one that contributes to sustaining the institutions that display it. Studying the

history of her own country, Judy is pleased to find examples of institutions that were N_1 (or at least came close enough to being N_1); and she is not surprised to learn that such institutions, once established, lasted for long periods of time. Delving deeper into the history of these institutions, however, Judy learns that their being N_1 did not actually make any significant difference to how long they lasted: their stability was rather due to some other feature they also shared, say their being N_2. On Blackburn's account of moral explanatory claims, there is only one option available to Judy after she learns this: stick to her moral theory and revise her moral explanatory belief. Once again, however, this is not the only option available to her. There is also another way Judy can go: stick to her moral explanatory belief and revise her moral theory. It is only institutions with the property of being N_1 that Judy has so far thought to be just, but what she learnt could end up changing this. It could lead her to conclude that, since justice is a stabilizing feature, it is institutions with the property of being N_2 that are just (N_1 institutions only qualify as just if and to the extent that they are N_2). Which of these two options is more reasonable, whether to revise her theory or her moral explanatory belief, will again depend on further details of Judy's evidential situation (for example, on what property N_2-ness is: is there a plausible theory of justice, alternative to the one Judy has accepted so far, that accounts for justice in terms of this property?), but both of these options are available to her, even after learning that N_1-ness does not make institutions stable. Just like theses i.H)-iv.H) and Hare's account of categorical moral sentences, Blackburn's account of moral explanatory claims rules out as unavailable an option we do have available when we develop our moral views, so it must be false.

As I noted above, the objection I have just raised against Blackburn's account is due to Sturgeon; what I have done is to show that Sturgeon's objection can be extended to Hare's account of categorical moral sentences as well and locate the source of the problem with both

accounts in theses i.H)-iv.H). So far I have been framing the objection to Blackburn's account as follows: the account must be false, since it has a false consequence. I now want to show that the objection can be pressed even without assuming that the consequence in question is false[33]. In order to do this, let me start by observing that (just like my previous example about Mary) the two examples above point to a very common procedure in actual moral inquiry. This should be even clearer once one remembers that, though the two examples above involve a user of moral language who has both a fully worked out moral theory and a moral explanatory belief, and who revises some of her moral beliefs because of new evidence, I could (as I did with the example about Mary) have formulated them just as well by considering a speaker who has a fully worked out moral theory and someone else who challenges that theory by inviting her to accept a moral explanatory belief, or a speaker with a fully worked out moral theory who questions her own theory by wondering whether to accept a moral explanatory belief. What we do in cases like these is to check whether, given the non-moral facts, a moral theory is able to account for a certain moral explanatory belief and assume that, if the answer to this question is negative, we have both of the following options available: we can accept the theory and reject the moral explanatory belief, but we can also accept the moral explanatory belief and reject the theory. Said otherwise, it is entirely common in actual moral inquiry to assume that starting from a negative answer to that question is a legitimate way in which one can end up rejecting a moral theory. This is perhaps more evident when the moral feature at issue is a virtue or a vice, since associated with such features are various assumptions about what manifestations of them

---

[33] In making this case too I'll be building upon Sturgeon's work, using remarks made by him in different articles and for various purposes in order to argue that the objection to Blackburn's account can be pressed on an assumption weaker than what I have relied on so far. I'll make this case only about the objection to Blackburn's account, since the present work deals with proposals about moral explanatory claims that are put forward in order to accommodate our practice with regard to such claims. It is easy to see that the same considerations also apply to the objection against Hare's account of categorical moral sentences, *if* such account is not meant to have any radical revisionary consequences with regard to our actual practice (it is, I think, an interesting question whether this is how Hare thinks of his account, but I won't pursue this question here).

typically amount to, i.e. various moral explanatory claims of the form 'Virtue (vice) V causes E': it is thus common to object to a proposed account of V in terms of a certain property by relying on a moral explanatory claim of this form and pointing out that the property in question fails to cause E. And the same happens with discussions about other moral features as well, since it is not only about virtues and vices that we hold moral explanatory beliefs. For example, a standard way in which we are taught about many moral features, not just virtues and vices, is by being taught various platitudes about what a manifestation of them would amount to (as well as by being given examples of things that have the feature in question): it is thus common to object to a proposed account of a moral feature by relying on a moral explanatory claim about it and pointing out that the proposed theory is unable to account for it (just as it is common to object to a proposed account by relying on less general moral beliefs and pointing out that the proposed theory is unable to account for them).

A different way of noting that the two examples above point to a very common procedure in actual moral inquiry comes out if we focus on beliefs about people's reliability as moral judges[34]. One reason why such beliefs are relevant in the present context is that they provide one more example of moral explanatory beliefs against which we test moral theories[35]. For example, suppose we believe Marta is a reliable judge of right and wrong. One of the things we do, when considering an account of wrongness in terms of a certain property, is to check whether, given

---

[34] In what follows, I'll assume that reliability claims such as 'Marta is a reliable judge of right and wrong' are causal claims (roughly put, the claim just mentioned is a claim to the effect that rightness/wrongness causes Marta to react by thinking 'That is right/wrong') and so are a subset of the set of claims of the form 'M-ness causes E' (where 'M' is a moral predicate), i.e. of the set of claims that the present work deals with. This assumption is not uncontroversial. If the reader does not share it, my argument is not affected, since the considerations just offered in the main text are enough to show that the two examples I gave above point to a very common procedure in actual moral inquiry. On the other hand, reliability claims are one of the standard type of examples discussed in the literature I am dealing with (Harman [1977], which originated this literature, focuses on such examples) and the assumption in question is often shared by participants to this debate (for a possible exception, see Gibbard [2003]), so it is worth pointing out what follows from it (for the record, I am inclined to accept this assumption). I thank Harold Hodes for pressing me to think about the material in this footnote.

[35] An example, moreover, of moral explanatory beliefs that we hold about moral features of any kind.

the non-moral facts, the theory in question is able to account for our belief about Marta: when she sees, or thinks about, an action that has the relevant property, does Marta react by judging it wrong? Once again, in actual moral inquiry we assume that, if the answer to this question is negative, we have both of the following options available: we can go for the theory and reject the belief about Marta, but we can also go for the belief about Marta and reject the theory. Said otherwise, it is entirely common in actual moral inquiry to assume that starting from a negative answer to a question like this is a legitimate way in which one can end up rejecting a moral theory. Another reason why beliefs about people's reliability as moral judges are relevant in the present context is the following: the use of such beliefs that I have just illustrated provides the most natural explanation of something pretty much everyone who engages in moral inquiry does as a matter of course. What I have in mind is the appeal we make, when we discuss moral theories, to the moral judgments we do or would make under certain conditions, ones that we think of as (at least to a degree) favorable conditions (such as having a full and vivid appreciation of the facts of the case one knows, having no personal interest at stake, etc.). This appeal is widespread in our moral discussions. Why do we appeal to the judgments made under such conditions, when we discuss a moral theory? Why do we think of these conditions as favorable ones? The most natural explanation is that we believe we are reliable moral judges if we meet these conditions (or at least we are more reliable moral judges if we meet them than if we do not), so we check whether the moral theory at issue is able to account for this belief about our reliability: faced with an account of, say, moral goodness in terms of a certain property, we ask whether we would, under the relevant conditions, react to things that have the property in question by judging them morally good. And again, it is entirely common in actual moral inquiry to assume that starting from a negative answer to this question is a legitimate way in which one

can end up rejecting the theory at issue. This appeal to moral judgments made under favorable conditions is a staple of our moral discussions: if the one just sketched is the correct explanation of it, then the two examples I gave above point to a procedure that is not just very common, but indeed pervasive in actual moral inquiry[36].

Let us now go back to Blackburn's account. Like all accommodationist proposals, this account is meant to allow opponents of moral explanations to accommodate our practice with regard to moral explanatory claims: it is a tool, designed for a specific purpose, namely achieving accommodation. It is worth remembering the overall dialectic here. Opponents of moral explanations deny that moral properties are causally efficacious but are usually unwilling to call for a deep revision of our actual practice: even if moral properties do not produce any causal effects, they say, no serious mistake is involved in our practice of offering, discussing, and accepting moral explanatory claims. Accommodationist proposals are the tool they design with the aim of accommodating our practice with regard to such claims, and Blackburn's account is one such proposal.

Of course, Blackburn's account cannot allow opponents of moral explanations to achieve this aim unless it is true, and earlier in this section I argued that the account must be false, since it has a false consequence. Suppose, however, that the consequence in question is true, and grant that the account is true as well: the considerations I have just given show that it does not in any case deliver the good it is advertised for, namely accommodation.

---

[36] My earlier example about Mary pointed to the following procedure: check whether, given the non-moral facts, a moral theory is able to account for certain less general moral beliefs; my two examples about Donald and Judy point to the following procedure: check whether, given the non-moral facts, a moral theory is able to account for certain moral explanatory beliefs. If the explanation sketched in the text of the appeal, in our moral discussions, to moral judgments made under favorable conditions is correct, then in some cases the second procedure underpins the first: this is so when we take care to filter the less general moral beliefs against which we test a moral theory, dropping those held only under conditions we take to be distorting. Sturgeon remarks that we test moral theories against beliefs about people's reliability as moral judges in Sturgeon [1986a] and [1991]; he suggests that this explains the appeal, in our moral discussions, to moral judgments made under favorable conditions in Sturgeon [1992] and [1998].

If Blackburn's account is true and the content of our moral explanatory claims is what it says it is, then every time we ask whether a moral theory is able to account for a certain moral explanatory belief and take it that, in case of a negative answer to this question, we have two options available, we are making a mistake. If Blackburn's account is true, in such cases there is only one option available to us: reject the moral explanatory belief. We are thus mistaken when, in our actual moral practice, we assume that starting from a negative answer to that question is a legitimate way in which one can end up rejecting a moral theory. We are mistaken, it is worth pointing out, no matter which way we eventually go, whether in the end we do reject the theory or we end up rejecting the moral explanatory belief instead: the mistake consists not in taking one option rather than the other, but in assuming we have two options available.

As I have just shown, this is a very common assumption in actual moral inquiry, so the mistake that someone endorsing Blackburn's proposal is committed to detecting in our practice is no minor, local mistake, one that can be quarantined, so to speak, and corrected while saving our practice pretty much the way it is. The mistake affects large areas of our practice. If the natural explanation sketched above of the appeal, in our moral discussions, to moral judgments made under favorable conditions is correct, it indeed permeates it. A friend of Blackburn's proposal is thus committed to calling for a deep (possibly dramatic) revision of our actual practice.

So: I claim that the consequence of Blackburn's proposal I pointed out above is false, therefore that the proposal is false. But the problem with Blackburn's proposal runs even deeper than this and persists *even if* one grants that the consequence in question is true and that so is the proposal. The problem is that the proposal is useless, since it fails to accommodate our practice: it is a tool that fails to achieve what it is designed to achieve. You can have it if you want, it does

109

not matter: if you want to deny there are moral explanations, you had still better be willing to call for a deep revision of our actual practice.

I want to end this section by highlighting two points about the arguments I have presented in it. The first point concerns the range of phenomena that an opponent of moral explanations who wants to go accommodationist ought to pay attention to[37]. Suppose you deny that moral properties are causally efficacious but want to accommodate our practice with regard to moral explanatory claims. The most obvious mistake you need to deny we are making, when we offer, discuss, and accept such claims, is that of ascribing causal efficacy to moral properties. If a claim of the form 'M-ness causes E' (where 'M' is a moral predicate) ascribed causal efficacy to the property of being M, then of course a serious mistake would, in your view, be involved in our practice with regard to such claims. It would be a mistake in our metaphysics: we would be treating as causally efficacious properties that lack causal power. Accommodationist proposals take care of this possible source of mistakes in our practice by offering a sentence that ascribes no causal efficacy to the property of being M (or any other moral property) as capturing the content of 'M-ness causes E'. Though it is the most obvious, this is not, however, the only possible source of mistakes you need to attend to. Here is an example of another possible source you should watch out for. One of the things we do with our moral explanatory claims is to reason about them and use them when reasoning about our moral views more broadly. If you want to accommodate our practice with regard to such claims, you need to be able to accommodate this

---

[37] Since in my discussion of non-cognitivist accommodationist proposals in this and the next two chapters I am putting the payoff option aside, I'll formulate my first point considering only accommodationist proposals framed in terms of the content of moral explanatory claims. Modest, obvious amendments would extend my remarks to accommodationist proposals framed in terms of the payoff of such claims. In chapter 1, I pointed out that the position of opponents of moral explanations who want to go accommodationist is analogous to that of other writers who, in other areas of philosophy, also claim that their metaphysical views (for instance, there are no material objects) do not have any radical revisionary consequences with regard to our actual practice (for instance, our saying such things as 'My desk is to the left of my armchair'). A point analogous to the one I am about to make seems to me to hold about some of these other debates as well.

aspect of it as well. And it might be that the accommodationist proposal you adopt in order to avoid charging us with a mistake in our metaphysics commits you to detecting one in our epistemology, i.e. in our evidential practices with regard to moral explanatory claims: this is what I have argued happens with Blackburn's proposal. Opponents of moral explanations who want to go accommodationist often write as if the first kind of mistake, the possible mistake in our metaphysics, were the only one they need to take care of, so that a proposal on which a claim of the form 'M-ness causes E' does not ascribe causal efficacy to the property of being M (or any other moral property) is all they need in order to accommodate our practice. The arguments in this section show that this won't do. Though a proposal along these lines is a necessary first step, there is more to accommodating our practice with regard to moral explanatory claims than taking care of that possible mistake in our metaphysics. Any accommodationist account of our practice with regard to such claims ought to be sensitive to a whole range of phenomena (for example, to our evidential practices with regard to these claims): it ought to be able to avoid detecting serious mistakes in most (if not all) of the central aspects of it. This is a point I'll make use of in my discussion of accommodationist proposals throughout the present work.

As I pointed out at the beginning of this chapter, non-cognitivist accommodationist proposals are meant to mimic the accommodationist proposals available to a cognitivist that I described in chapter 2. It is thus instructive, in studying non-cognitivist proposals, to ask whether the objections they face also apply to the proposals available to a cognitivist.

As I noted above, the source of the problem I have raised for Blackburn's proposal -and for Hare's account of categorical moral sentences- lies in theses i.H)-iv.H), on which Blackburn's proposal –and Hare's account- are built. In section 3.1 I observed that the epistemological view formulated by theses i.H)-iv.H) has three noteworthy features: it is a

foundationalist view, it honors the doctrine of the autonomy of ethics, and it imposes a strict constraint on the kinds of beliefs that a categorical moral belief can ultimately be inferred from. The problem with i.H)-iv.H) that I have pointed out in this section is a problem with the last of these three features. According to i.H)-iv.H), evidential support within one's moral belief system ultimately flows only from the more to the less general: specifically, from one's direct moral beliefs to the other categorical moral beliefs one subsumes –by means of appropriate non-moral beliefs- under them; but, as I argued at the beginning of this section, this is false.

The three features just mentioned, I noted in 3.1, are ones that the view formulated by theses i.H)-iv.H) shares with the epistemological view I described in 2.4, which means that the latter view also faces the problem I have raised for theses i.H)-iv.H): the same argument I presented at the beginning of this section to show that i.H)-iv.H) are false can also be run against the view I described in 2.4. Theses i.H)-iv.H) articulate Hare's way of fleshing out the Assumption, the reading of the Assumption on which Blackburn's accommodationist proposal – and Hare's account of categorical moral sentences- are built. The cognitivist proposals presented in chapter 2, on the other hand, are *not* built on the epistemological view described in 2.4, so they do not inherit its problems.

The most natural way of fleshing out the Assumption, on which cognitivist proposals are built, is by means of the metaphysical picture articulated by theses i)-iv), and such a metaphysical picture does not commit its friends to the epistemological view described in 2.4: as I put it above, if you are a cognitivist your epistemology is not settled by your acceptance of the Assumption. On the other hand, if you are a non-cognitivist then theses i)-iv) are not available to you; you can adopt Hare's way of fleshing out the Assumption, but this way of reading the

Assumption does commit you to a certain epistemology: one, I have argued, with false consequences with regard to the way in which we develop our moral views.

The upshot -and this is the second point I want to highlight- is that it does make a difference whether you are a cognitivist or a non-cognitivist: *contra* what non-cognitivists typically claim, there are serious limits to the degree to which a non-cognitivist can mimic what a cognitivist can do. More accurately: the arguments in this section show that such limits exist, if the non-cognitivist tries to do the mimicking by relying on Hare's way of fleshing out the Assumption (I'll look at a different non-cognitivist reading of the Assumption starting from the next chapter). That way of fleshing out the Assumption is problematic, and its problems carry over to what one builds on it.

In particular, if you are an opponent of moral explanations who wants to go accommodationist, the prospects of your accommodationist project will be affected if you rely on Hare's way of fleshing out the Assumption: by relying on it, you will not be able to build accommodationist proposals able to mimic in all relevant respects the accommodationist proposals available to a cognitivist.

The lesson here is more general, moreover, and it is worth stressing, given the centrality in the non-cognitivist tradition of what in the previous section I called the 'Go by the speaker's standards' strategy for specifying the content of moral sentences and of the related doctrine of the descriptive meaning of moral expressions. As I pointed out in 3.2, Blackburn's accommodationist proposal can be seen as the application, to the particular case of moral explanatory claims, of that general strategy, which is the standard strategy non-cognitivists have appealed to in order to deal with features of our use of moral expressions that seem to call for a cognitivist treatment. That strategy is meant to work by mimicking the 'Go by the M-making

properties' strategy available to cognitivists that I described in chapter 2 and it is built on Hare's way of fleshing out the Assumption: if the arguments in this section are correct, they show that the 'Go by the speaker's standards' strategy cannot allow a non-cognitivist to carry the mimicking very far. In the next two sections, I'll give two more reasons for the same conclusion.

4. Second objection to the 'Go by the speaker's standards' strategy

One way to see that theses i.H)-iv.H) are false, I said at the beginning of the previous section, is by looking at their consequences with regard to the way in which we develop our moral views. In this section, I want to give a different reason for thinking that those theses are false, one that comes out if we look at what they require in order for us to hold moral views at all. The best way to introduce this further problem for i.H)-iv.H) is by going back to a feature of Hare's account of categorical moral sentences that I commented on in section 3.2, one that is shared by Blackburn's proposal about moral explanatory claims: in specifying the content of a moral sentence of the relevant type, as uttered by a given speaker, both accounts consider only a user S of moral language who has made up her mind about what direct moral beliefs to accept, for a given moral predicate 'M'. As I noted in 3.2, very few of us have made up our minds about this, and we are not told explicitly how exactly to move from assigning content to a moral sentence of the relevant type, as uttered by S, to assigning content to such a sentence, as uttered by speakers like most of us. After mentioning very quickly some alternatives available to a friend of either account to take at least some steps toward closing the gap between S and most of us, I said I would ignore this gap and argue that the accounts fail even on the assumption that we have all made up our minds about what direct moral beliefs to accept.

Let us continue then to consider only users of moral language who have made up their minds about this. In presenting the two accounts, I considered only the cases of a speaker who has made up her mind about what direct moral beliefs to accept, for a given moral predicate 'M', and holds either only one or more than one such belief: what about a user of moral language who has made up her mind about what direct moral beliefs to accept and holds none? Are there any such users of moral language? Yes, there are: they are called particularists. Moral particularism is the denial of all appropriately strong universal moral sentences using non-moral predicates: the moral particularist thus rejects all such sentences, so on i.H)-iv.H) she counts as someone who has made up her mind about what direct moral beliefs to accept and holds no such beliefs. The question I want to consider in this section is what theses i.H)-iv.H), and the two accounts built on them, entail about such a user of moral language.

Consider then Jonathan Dancy. According to i.H), any categorical moral belief Dancy holds is either an inferred belief or a direct belief. Given iv.H), any direct belief he has can be expressed by an appropriately strong universal moral sentence using a non-moral predicate denoting a non-moral property: as we have just seen, this means that, since Dancy is a particularist, none of his categorical moral beliefs are direct beliefs. Since on ii.H) any inferred belief he has is ultimately inferred from direct beliefs (plus appropriate non-moral beliefs) he has, it follows that Dancy has no inferred beliefs either, hence that he holds no categorical moral beliefs at all. More accurately: since i.H)-iv.H) are epistemological, not merely psychological, theses, it follows that Dancy holds no categorical moral beliefs at all *insofar as* he avails himself only of inferences of an epistemologically good sort. Said otherwise: if Dancy does hold any categorical moral beliefs, he is doing something epistemologically objectionable. But Dancy

does hold categorical moral beliefs, so if theses i.H)-iv.H) are true, he is doing something epistemologically objectionable in holding such beliefs.

Now, I suppose it can be granted that (just like any of us) Dancy is likely doing something epistemologically objectionable in holding *some* of the categorical moral beliefs he does hold: presumably, among his categorical moral beliefs are some that, were he to reason in an epistemologically fully respectable way, he would not hold. But what follows from i.H)-iv.H) is something much stronger than this. What I have just said we can grant is that Dancy is (like any of us) to some degree unreasonable in his categorical moral beliefs. What follows from i.H)-iv.H), on the other hand, is that he is unreasonable to an exceptionally high degree. If i.H)-iv.H) are true, Dancy is doing something epistemologically objectionable in holding *any* of the categorical moral beliefs he has: each categorical moral belief he has is one he is being unreasonable in holding. And this seems wrong: Dancy is not unreasonable to *this* degree. There is no such significant difference between Dancy and other users of moral language with respect to the degree to which they are unreasonable in their categorical moral beliefs. There are, of course, differences among users of moral language under this respect, and there may well be other users who fare better than Dancy does here; however, there is no such difference between Dancy and other speakers as theses i.H)-iv.H) entail there is. Here is, therefore, a further consequence of theses i.H)-iv.H), distinct from the one I pointed out in the previous section, that is false: a further reason for thinking that those theses are false.

But, one might ask, shouldn't it be obvious that i.H)-iv.H) have this consequence? If, as I put it at the beginning of this chapter, i.H)-iv.H) are one way of articulating the idea that use of moral language commits one to general moral principles, then of course they entail there is something wrong with a particularist who nonetheless holds categorical moral beliefs: what

exactly is the point of drawing attention to the fact that they do? Surely, one might think, anyone who accepts i.H)-iv.H) will not be moved by the problem I am trying to press here or s/he will have some way of dealing with it. And friends of i.H)-iv.H) do typically have a way of dealing with a more general version of the problem I am pressing here: the point of drawing attention to what those theses entail about moral particularists is that it shows the inadequacy of that typical way of dealing with the problem in question.

In order to see that there is a more general version of the problem I am pressing here, we need to consider what follows from i.H)-iv.H) if we temporarily drop the assumption that, as I reminded the reader at the beginning of this section, I am working under, namely that we have all made up our minds about what direct moral beliefs to accept (this assumption being a concession to my opponents).

Even if most of us have not made up our minds about what direct moral beliefs to accept and so are not (among other things) moral particularists, it looks as though we resemble the particularist under an important respect. At least on the face of it, what seems to be true of most of us is that if we do hold categorical moral beliefs that can be expressed by an appropriately strong universal moral sentence using a non-moral predicate denoting a non-moral property, we do not in any case hold enough such beliefs for us to ultimately infer *via* subsumption *all* other categorical moral beliefs we hold from them (plus appropriate non-moral beliefs). Indeed, on the face of it the share of categorical moral beliefs most of us hold that we do not have beliefs with the required features to subsume is not negligible. This is a familiar point about the difficulty of codifying our moral standards, at least if they are meant to meet strict, specific constraints; the difficulty the reader will encounter in formulating enough beliefs with the required features that s/he holds will confirm it.

This makes most of us similar to the particularist in the following respect: if i.H)-iv.H) are true, it looks as though most of us are unreasonable in holding many of the categorical moral beliefs we do hold (in the case of the particularist, this applies to all the categorical moral beliefs she holds). This too, I should think, is an obvious consequence of i.H)-iv.H) (obvious, that is, at least if one takes into account what our categorical moral beliefs actually look like). Interestingly, it is a consequence that friends of i.H)-iv.H) regard as problematic: far from being unmoved by it, what they do when faced with it is trying to rescue those theses from it. Their typical way of dealing with this problem consists in claiming that appearances are misleading here and that most of us do hold enough categorical moral beliefs with the required features: no doubt, they usually grant, it can at times be hard to formulate them, but this merely shows that such beliefs can at times be somewhat vague or possibly held only implicitly, not that we do not have them[38].

In order to assess this line of reply on behalf of i.H)-iv.H), one would need to determine whether most of us do or do not hold enough beliefs of the required sort at least in a vague or possibly only implicit form, a question I am not exactly sure how to address and one on which, given the appearances, the burden of proof falls on friends of i.H)-iv.H), I should think. Moral particularists are relevant here because, no matter what in the end one thinks of such a reply when applied to most of us, it is in any case unavailable when it comes to them. Drawing attention to moral particularists shows that this typical line of reply does not in any case solve the problem, so that friends of i.H)-iv.H) are still faced with the consequences of those theses: obvious consequences maybe, but still ones that are, and are regarded by them as, problematic[39].

---

[38] As we saw at the beginning of 3.1, this is what Hare claims when faced with the difficulty of formulating our moral standards.
[39] Is the line of reply described in the text really unavailable in the case of moral particularists? What if a friend of i.H)-iv.H) tries saying that particularists are wrong about their moral views and that they do in fact hold categorical

Like the problem with i.H)-iv.H) that I pointed out in the previous section, this further problem too is mirrored by parallel problems with the two accounts built on those theses. Both Hare's and Blackburn's account proceed by considering a user S of moral language who has made up her mind about what direct moral beliefs to accept, for a given moral predicate 'M', and specifying the content of a moral sentence of the relevant type, as uttered by S, *via* reference to the non-moral properties picked out by S's direct moral beliefs. On these accounts, a moral sentence of the relevant type, as uttered by S, thus *presupposes* (to use the same terminology borrowed from debates about the theory of definite descriptions I used in chapter 2) that there is at least one non-moral property that can be denoted by some non-moral predicate 'N' such that S believes that ∎ all N's are M[40]. If, in the case of a given speaker, there turn out to be no such properties, the upshot of the two accounts is therefore that a moral sentence of the relevant type, as uttered by that speaker, has no content. But there are speakers for whom there are no such properties: they are called particularists. Consider then Jonathan Dancy again. On Hare's account, a categorical moral sentence, as uttered by Dancy, has no content; so does a moral explanatory claim, as uttered by him, on Blackburn's account. When Dancy says things like 'That action is right' or 'Justice makes institutions stable', his remarks have no content, say the

---

moral beliefs that can be expressed by an appropriately strong universal moral sentence using a non-moral predicate denoting a non-moral property? Well, one can always claim that really, deep down, things are as one's theory predicts them, but this does not make it plausible that they are, and that particularists do in fact hold such beliefs is very implausible. *Maybe* friends of i.H)-iv.H) can manage to show that, appearances notwithstanding, most of us do hold enough such beliefs, at least in a vague or possibly only implicit form, but it is hard to believe that the same is true of particularists. In the text, I temporarily drop the assumption that we have all made up our minds about what direct moral beliefs to accept in order to show that there is a more general version of the problem I am pressing against i.H)-iv.H), one that applies to most of us (as we actually are) in addition to moral particularists. Since I am working under that assumption, however, my argument does not press this more general version of the problem against i.H)-iv.H): it focuses only on what i.H)-iv.H) entail about particularists, and not also on what they entail about most of us (as we actually are). As it turns out, this is no limitation on my argument, since the case of particularists is precisely the case for which the typical line of reply on behalf of i.H)-iv.H) described in the text is unavailable. I thank Karen Bennett and Nick Sturgeon for pressing me to think about the objection to my argument discussed in the text.

[40] I say 'presupposes', instead of 'presupposes or entails' (as I did in chapter 2 when discussing cognitivist proposals), because both accounts under discussion here rely on singular substitution.

two accounts. But this, I suppose it can be granted, is wrong: no matter what we think of what Dancy says, there is something he is saying. The two accounts thus have a false consequence, so they must be false[41].

It is easy enough to see that the objection I have just raised can also be framed, as I did with the one I presented in the previous section, as an objection not to the truth of Blackburn's account, but to its credentials as an accommodationist proposal[42]. In order to see this, let us suppose that the consequence of Blackburn's account I have just pointed out is true, and that the account is true as well. Obviously we do not, in our actual moral practice, treat moral explanatory claims uttered by particularists as having no content. If Dancy says something like 'Justice makes institutions stable', as a rule we do not dismiss his remark as empty: no matter what we think of what Dancy has said, we take him to have said something, and we usually engage with it, asking him to clarify it or support it with evidence, possibly raising doubts about or expressing agreement with it (or, if we ignore Dancy's remark, it is because we have no time for or interest in it, not because we take it to be empty). Under this respect, we treat Dancy no differently from other users of moral language who are not moral particularists. If Blackburn's account is true and the content of moral explanatory claims is what it says it is, we are mistaken in doing so: this is an aspect of our practice with regard to such claims that needs to be revised.

---

[41] As I pointed out in n. 14, there is, on Hare's account of categorical moral sentences, a limited class of uses of such sentences (what Hare calls the 'inverted commas' and the 'conventional' use) for which content is specified *via* reference to the non-moral properties picked out by the direct moral beliefs held by someone other than the speaker, so *very* strictly speaking categorical moral sentences can, on Hare's account, have a content even as uttered by Dancy, when used in one of these ways; and one might extend the notions of the inverted commas and the conventional use to moral explanatory claims as well. It should be obvious, however, that interpreting a moral sentence of the relevant type, as uttered by Dancy, as being used always either in the inverted commas or in the conventional way still yields terribly implausible results (which should of course come as no surprise, given that even on Hare's view these are meant to be peripheral, parasitic uses of a moral sentence).

[42] As I did in the previous section and for the same reasons given there (see n. 33), I'll bracket Hare's account and make this point only with regard to Blackburn's account.

Now, this mistake that a friend of Blackburn's account is committed to detecting in our actual practice is not so deep as the one that I argued in the previous section she is also committed to detecting. Surely I cannot claim (as I did about the mistake I pointed out in the previous section) that it affects large areas of, indeed possibly permeates, our practice: few users of moral language, after all, are moral particularists. Still, it is a mistake that is well worth noting: if you endorse Blackburn's account, you are committed to calling for ceasing to treat moral explanatory claims uttered by Dancy and other particularists as contentful claims.

Moreover, whether a revision of our practice is deep enough to count as a failure to accommodate the practice in question does not depend just on how large an area of the practice is affected by it: it can be a matter, so to speak, of quality as well as size. Two points are worth noting under this respect. One mistake we sometimes make, when we react to remarks made by other people, consists in misunderstanding what they have said: it consists in taking them to have said one thing, whereas they were saying a different thing. The first point is that on Blackburn's account the mistake we are making, in treating moral explanatory claims uttered by Dancy and other particularists the way we do, is more radical than this, since it consists in treating as contentful claims that have no content: it consists in taking particularists to have said something, whereas they were saying nothing. And -this is the second point- it is a mistake that should strike us even more as a serious mistake to detect in our actual moral practice, since the people whose claims we are to diagnose as devoid of content are among those who have reflected critically upon our practice and provided insightful observations about it. Though not widespread, the mistake in question is thus still pretty serious[43]. Finally, it is a mistake that adds to the one I

---

[43] The second of the two reasons I give in the text depends on the assumption that particularists have provided insightful observations about our actual moral practice. This assumption seems to me correct: whether or not we should in the end endorse moral particularism, it seems to me its defenders have made valuable contributions to our

pointed out in the previous section: if you endorse Blackburn's proposal, you are committed to detecting both mistakes in our actual practice, so surely you had better be willing to call for a deep revision of it.

So: I claim that the consequence of Blackburn's proposal I have pointed out in this section is false, therefore that the proposal is false. But, once again, the problem with Blackburn's proposal persists even if one grants that the consequence in question, and the proposal, are true. The problem is that the proposal is useless: it is unable to accommodate our practice. If you want to deny there are moral explanations, Blackburn's proposal cannot help you avoid being committed to calling for a deep revision of our actual practice.

In studying non-cognitivist accommodationist proposals it is instructive, I said in the previous section, to ask whether the objections they face also apply to the cognitivist accommodationist proposals described in chapter 2 that they are meant to mimic. The objection I have raised against Blackburn's proposal –and Hare's account of categorical moral sentences- in this section goes as follows: on Blackburn's proposal –and Hare's account-, a moral sentence of the relevant type, as uttered by a speaker S, presupposes that there is at least one non-moral property that can be denoted by some non-moral predicate 'N' such that S believes that ■ all N's are M; if, in the case of a given speaker, there turn out to be no such properties, the upshot is thus that a moral sentence of the relevant type, as uttered by that speaker, has no content; some speakers endorse moral particularism, i.e. the denial of all appropriately strong universal moral sentences using non-moral predicates; it follows that moral sentences of the relevant type, as uttered by moral particularists, have no content; but this consequence is false, and in any case it

---

understanding of our moral practice. I do grant, however, that any assessment of the merits of particularism is controversial, and so that the second reason I give in the text might not speak to everyone.

shows that Blackburn's proposal –and Hare's account- are unable to accommodate our actual practice.

Now, as I pointed out in the previous chapter[44], there is a possible argument against the cognitivist proposals presented in chapter 2 that mirrors the objection I have raised in this section. It goes as follows: on any of those proposals, a moral sentence of the relevant type (or the sentence capturing its payoff) presupposes or entails that there is at least one non-moral property that can be denoted by some non-moral predicate 'N' such that ■ all N's are M; if there turn out to be no such properties, the upshot is thus either that a moral sentence of the relevant type has no content (or payoff) or that it (or the sentence capturing its payoff) is false: in any case, there is something wrong with it; moral particularism, i.e. the denial of all appropriately strong universal moral sentences using non-moral predicates, is true; it follows that there is something wrong with moral sentences of the relevant type; this consequence shows that those proposals are unable to accommodate our actual practice.

As I say, this is a possible argument against the cognitivist proposals presented in chapter 2. Note, however, how hard it would be to press this argument against those proposals: it relies on the premise that moral particularism is true, and this is a premise that is quite difficult to establish, which is why I said I would not explore this argument in the present work. On the other hand, the argument against Blackburn's proposal –and Hare's account- relies on the weaker premise that there are moral particularists, and this premise is uncontroversial.

What is the source of this difference? The cognitivist proposals described in chapter 2 are built on theses i)-iv), which represent the most natural way of fleshing out the Assumption and put forward a straightforwardly metaphysical picture; as a result, those proposals specify the

_____

[44] See ch. 2, n. 33.

content (or payoff) of a moral sentence of the relevant type *via* reference to the non-moral properties picked out by *the brute moral truths* there are. What is needed, on those proposals, in order to draw the conclusion that there is something wrong with moral sentences of the relevant type is thus the assumption that there are no such truths, which follows from the premise that moral particularism is true, a premise that is quite hard to establish. If you are a non-cognitivist, theses i)-iv) are not available to you. You can adopt Hare's way of fleshing out the Assumption and build your proposals on theses i.H)-iv.H), which formulate an epistemological view; if you do, your proposals will specify the content of a moral sentence of the relevant type, as uttered by a speaker S, *via* reference to the non-moral properties picked out by *the direct moral beliefs* S holds. What is needed, on your proposals, in order to draw the conclusion that moral sentences of the relevant type, as uttered by certain speakers, have no content is thus the weaker assumption that there are speakers who hold no such beliefs, which follows from the weaker and uncontroversial premise that there are moral particularists.

The upshot is, once again, that it does make a difference whether you are a cognitivist or a non-cognitivist: there are serious limits to the degree to which a non-cognitivist can mimic what a cognitivist can do, at least if the non-cognitivist tries to do the mimicking by relying on Hare's way of fleshing out the Assumption. If you are a non-cognitivist opponent of moral explanations who wants to go accommodationist then, in the absence of a different non-cognitivist reading of the Assumption, you will not be able to build accommodationist proposals able to mimic in all relevant respects the accommodationist proposals available to a cognitivist. And, once again, the lesson here applies quite generally to the 'Go by the speaker's standards' strategy (of which Blackburn's accommodationist proposal can be seen as a particular application), and the related doctrine of the descriptive meaning of moral expressions, so central

to the non-cognitivist tradition. That general strategy is meant to work by mimicking the 'Go by the M-making properties' strategy available to cognitivists that I described in chapter 2 and is built on Hare's way of fleshing out the Assumption: the arguments in this section provide another reason, in addition to that given in the previous section, to believe that the 'Go by the speaker's standards' strategy cannot allow a non-cognitivist to carry the mimicking very far. A third reason will come out in the next section.

5. <u>Third objection to the 'Go by the speaker's standards' strategy</u>

In sections 3.3 and 3.4 I proceeded as follows: I started out by raising a problem for the epistemological view formulated by theses i.H)-iv.H) and I then showed that that problem is mirrored by parallel problems with Blackburn's proposal about moral explanatory claims and Hare's account of categorical moral sentences, which are built on those theses. The two objections I have so far raised against Blackburn's proposal and Hare's account thus have their source in the epistemological view articulated by i.H)-iv.H). The objection I want to raise in this section is different: there is a further problem with Blackburn's proposal, one that arises *even if* one grants that that epistemological view is true. The same, I'll argue, holds for Hare's account; indeed, the comparison I have throughout been drawing between proposals about moral explanatory claims and ones about categorical moral sentences will prove especially fruitful in our discussion of this problem.

   In order to press this problem, I'll start from a second objection raised against Blackburn's proposal by Sturgeon, in addition to the one I discussed in section 3.3; as I did in 3.3, I'll then build upon Sturgeon's objection in various ways. In particular, I'll argue that the objection to Blackburn's proposal can be pressed on weaker assumptions; I'll put forward, on

Blackburn's behalf, an amended version of that proposal designed to avoid the objection in question, and argue that it ultimately fails; and I'll argue that the same problem can (surprisingly, as we shall see) be pressed against Hare's account. In pressing this problem, I will (as I did in section 3.3 and for the same reason given there) focus on speakers who have made up their minds about what direct moral beliefs to accept, for a given moral predicate 'M', and hold only one such belief: consideration of this case is enough to raise a problem for Blackburn's proposal and Hare's account, no matter how they are extended to cover speakers who have made up their minds and hold more than one such belief[45].

Consider then a user of moral language S_1 who has made up her mind about what direct moral beliefs to accept, for a given moral predicate 'M', and holds only one such belief, say that ■ all N_1's are M; consider a different user of moral language S_2 who has also made up her mind about what direct moral beliefs to accept and who also holds only one such belief, but who accepts a different one, say that ■ all N_2's are M. The problem with Blackburn's proposal pointed out by Sturgeon is that it is unable to account for agreement and disagreement on moral explanatory claims among speakers who, like S_1 and S_2, have different moral standards.

Suppose S_1 says 'M-ness causes E' while S_2 says 'M-ness does not cause E'. On Blackburn's proposal, the content of the first remark is that N_1-ness causes E while the content of the second remark is that N_2-ness does not cause E, which means that S_1 and S_2 are not disagreeing with one another, rather they are talking past one another. But this consequence of Blackburn's proposal is false: intuitively, two speakers like S_1 and S_2 are as a rule disagreeing

---

[45] The same problem I'll discuss in this section can also be raised considering speakers who hold more than one direct moral belief, though for simplicity's sake I won't do it (at least, not explicitly: a careful reader will note that, in choosing examples to illustrate the problem, I'll feel free to choose ones that include speakers with more than one direct moral belief).

with one another[46]. Suppose (to adapt Sturgeon's example) I say 'Justice makes institutions stable' and you say 'Justice does not make institutions stable'; suppose I have a fully worked out theory of justice: say, I agree with Rawls; suppose you also have a fully worked out theory of justice, but a different one: say, you agree with Nozick. Does this mean that, in saying what we say, we are talking past one another? Hardly. *Contra* Blackburn's proposal, in saying what we say we are disagreeing with one another, despite the fact that we have different theories of justice. Indeed, that fact might well be one of the sources of the disagreement between us. Here is, therefore, still a further consequence of Blackburn's proposal, distinct from those I pointed out in the two previous sections, that is false: a further reason for thinking that the proposal is false.

Note, moreover, that the non-cognitivist herself has quite strong, independent reasons for thinking that this consequence of Blackburn's proposal must be false. This is so because the argument just presented turns against Blackburn's proposal the same objection that the Master Argument for non-cognitivism raises against the cognitivist who believes that moral predicates can be analyzed in terms of non-moral ones, and there seems to be no more reason for thinking the latter argument works against the cognitivist than there is for thinking the argument just presented is successful against Blackburn's proposal: the same intuitions of univocality marshaled by the Master Argument seem to be available here, and it is hard to see why such intuitions ought to be trusted in one case but not the other.

---

[46] I say 'as a rule' because, in order to raise a problem for Blackburn's proposal, one need not claim -what might not be very plausible- that it is *never* the case that the two speakers in question are talking past one another. The weaker claim in the text –which is undeniable- is enough.

Sturgeon raises this objection against Blackburn's proposal in his [1986b][47]. So far I have been framing it as an objection to the truth of that proposal: the proposal must be false, since it has a false consequence (one, moreover, that the non-cognitivist herself is committed to thinking must be false). Given my purposes here, it is interesting to note that it can be pressed even without assuming that the consequence in question is false: it too can, like those I presented in the two previous sections, be framed as an objection not to the truth of Blackburn's proposal, but to its accommodationist credentials.

In order to see this, suppose that the consequence of Blackburn's proposal I have just pointed out is true (bracket, moreover, the strong reasons the non-cognitivist has for thinking it must be false), and grant that the proposal is true as well. In our actual moral practice, we do not treat speakers like S_1 and S_2 above as talking past one another: quite to the contrary, as a rule we take them to be engaged in a genuine disagreement on a moral explanatory claim. Suppose again I say 'Justice makes institutions stable' and you say 'Justice does not make institutions stable'. Coming to realize that I agree with Rawls and you with Nozick does not lead us to end our discussion and conclude we are not disagreeing after all: rather, it leads us to conclude we might have located one of the sources of the disagreement we are engaged in (if it leads us to end our discussion, it is not because we take ourselves to be talking past one another, but because we now see our disagreement is so deep that it would take too long to settle it). Or, to borrow an example from a classic, when Thrasymachus maintains, in book I of Plato's *Republic*, that justice makes life unhappy and Socrates claims that, far from making it unhappy, justice makes life

---

[47] See also Sturgeon [1991] and, for an interesting application of this objection to the case of discussions of the problem of evil, [1995a]. Does Sturgeon's original formulation of the objection predate Blackburn's presentation of his proposal in Blackburn [1991a]? Here is the story: Blackburn [1991a] is a reply to Sturgeon [1986b]; in that article Sturgeon puts forward, on behalf of the non-cognitivist, a proposal very close to Blackburn's and then raises against it the objection I am now discussing; the same objection also applies to the proposal Blackburn puts forward in his [1991a]. In his reply, Blackburn embraces the consequence that two speakers like S_1 and S_2 in the main text are talking past one another, without addressing the problems that taking this route raises, especially for a non-cognitivist. For the reasons just given in the text and those I am about to give, I find this reply unsatisfactory.

happy, we take the two of them to be engaged in a genuine disagreement, even if we learn that Socrates does not share Thrasymachus' theory of justice.

If Blackburn's proposal is true and the content of moral explanatory claims is what it says it is, we are mistaken in treating speakers like S_1 and S_2 above the way we do: this is an aspect of our practice with regard to such claims that needs to be revised. And this is certainly a very serious mistake to detect in our practice of offering, discussing, and accepting moral explanatory claims. One of the central aspects of any such practice with regard to claims of any kind is the way in which we detect agreement and disagreement about claims of the relevant kind: if we treat as addressing one and the same issue -agreeing or disagreeing about it- speakers who are instead talking past one another, something defective lies at the very heart of our practice.

To see how deep a revision of our actual practice a friend of Blackburn's proposal is committed to calling for, imagine what the revised practice would look like. Consider a moral explanatory claim of the form 'M-ness causes E': our actual practice with regard to a claim like this, which is shaped like a discussion of one and the same issue that can be joined by parties who subscribe to many, different theories of the relevant moral feature, would sunder into many, parochial discussions of different issues, one for each different moral theory. If this does not already strike you as something very different from our actual practice, think about how often we currently compare different moral theories with respect to more specific claims, including moral explanatory claims, taking note of the fact that friends of two different theories can, despite their differences, agree on a certain moral explanatory claim, or observing that the difference between two theories shows up in the fact that their friends disagree on a certain moral explanatory claim; and think about how important we take such comparisons to be, given that we take some of the most interesting (or challenging) instances of agreement (or disagreement) on specific issues to

be precisely instances of agreement (or disagreement) among friends of different theories. All this would have to go.

So here we are again, for the third time: I claim that the consequence of Blackburn's proposal I pointed out at the beginning of this section is false, therefore that the proposal is false. But, once again, the problem with Blackburn's proposal runs even deeper than this: the proposal fails to deliver the good it is advertised for, namely accommodation. If you want to deny there are moral explanations, you are still committed to calling for a deep revision of our actual practice.

The objection I am discussing points out something worth noting about Blackburn's proposal, namely that it is in a respect a surprising proposal for a non-cognitivist to make[48]. The comparison with Hare's account of categorical moral sentences explains why. As we saw, in providing an account of the content of categorical moral sentences, Hare implements his strategy of 'going by the speaker's standards' in two different ways: for instances of a certain class of (peripheral) uses of categorical moral sentences, Hare's account relies on non-enriched singular substitution; for instances of a different (more typical) use of such sentences, the account relies on enriched singular substitution. The reason why we cannot give an account of the content of categorical moral sentences by relying only on non-enriched singular substitution is that, if we did, we would be unable to account for agreement and disagreement on such sentences among speakers with different moral standards: this is, in Hare's view, the mistake made by the cognitivist that the Master Argument warns us of.

---

[48] See Sturgeon [1991] and, especially, [1995a] for a similar assessment of Blackburn's proposal. In the remainder of this section I'll often need to refer to the operations of singular substitution *via* reference to the n_S's, and never to the operations of singular substitution *via* reference to the n's; for brevity's sake, I'll thus feel free to use 'singular substitution' as shorthand for 'singular substitution *via* reference to the n_S's'.

As pointed out above, Blackburn's proposal can be seen as the application to moral explanatory claims of the same strategy applied by Hare to categorical moral sentences. However, Blackburn's proposal implements that strategy in only one way: it accounts for the content of moral explanatory claims by relying only on non-enriched singular substitution. Blackburn's proposal therefore tries to do, with moral explanatory claims, precisely what the Master Argument is supposed to teach us will not work with categorical moral sentences. No wonder that it is vulnerable to the objection I am discussing.

A comparison with Hare's account of categorical moral sentences thus helps explain why Blackburn's proposal is a surprising proposal for a non-cognitivist to make, one bound, by the non-cognitivist's own standards, to suffer from the problem I am discussing. At the same time, however, it may also help find a way out of that problem. If the source of the problem lies in accounting for the content of moral explanatory claims by relying only on non-enriched singular substitution, why not put forward an accommodationist proposal that relies on enriched singular substitution? Blackburn does not do that, but we can try that on his behalf, putting forward the following amended version of Blackburn's proposal:

(B.H) The content of 'M-ness causes (does not cause) E' (as uttered by S) is that N_1-ness causes (does not cause) E and ■ all N_1's are M

for the case of a user S of moral language who has made up her mind about what direct moral beliefs to accept, for a given moral predicate 'M', and holds only one such belief (supposing that the predicate denoting the n_S is 'N_1'), and

(B.H') The content of 'M-ness causes E' (as uttered by S) is that N_1-ness causes E and

N_2-ness causes E and ■ all N_1's are M and ■ all N_2's are M; the content of 'M-ness does

not cause E' (as uttered by S) is that (N_1-ness does not cause E or N_2-ness does not cause

E) and ■ all N_1's are M and ■ all N_2's are M;

(B.H'') The content of 'M-ness causes (does not cause) E' (as uttered by S) is that the

disjunctive property of being $N_1 \vee N_2$ causes (does not cause) E and ■ all N_1's are M and

■ all N_2's are M

as two different ways of extending the amended proposal to the case of a user S of moral

language who has made up her mind about what direct moral beliefs to accept, for a given moral

predicate 'M', and holds two such beliefs (supposing that the predicates denoting the n_S's are

'N_1' and 'N_2') -it is easy to extrapolate from this case to cases in which S has made up her mind

and holds more than two such beliefs-[49].

Whereas on Blackburn's original proposal the content of a moral explanatory claim is

entirely non-moral, on the amended proposal a moral explanatory claim does have moral content,

captured by the additional clause ' ■ all N_1's are M' (or ' ■ all N_1's are M and ■ all N_2's are M'),

whose function is, on non-cognitivism, to express the relevant non-cognitive attitude toward N_1's

[49] As I did with Blackburn's original proposal, I keep the letter 'H' in my names for the theses formulating the amended version of that proposal because they too are built on Hare's way of fleshing out the Assumption. As usual, the clause ' ■ all N_1's are M' in (B.H) should read ' ■ all N_1's are M *and* not (■ all N_2's are M) and not (■ all N_3's are M) and …', but in order to make things simpler for the reader and easier for my opponent I'll frame it the way I have just done in the main text and drop the rest, taking it as understood (analogous remarks apply to the clause ' ■ all N_1's are M and ■ all N_2's are M', in all of its occurrences, in (B.H') and (B.H'')). In principle, the non-cognitivist can use the amended proposal in two different ways: she can claim it is all we need to account for the content of moral explanatory claims; or she can (as Hare does with categorical moral sentences) distinguish between two different uses of moral explanatory claims, a central, more typical use, to be accounted for in terms of the amended proposal, and a peripheral use, to be accounted for in terms of Blackburn's original proposal. In what follows, I won't distinguish between these two different ways of using the amended proposal that are available to the non-cognitivist.

(or N_1's and N_2's). And, one might think, the presence of this non-cognitive component will allow the non-cognitivist to avoid the problem I have raised against Blackburn's proposal in this section: it will allow her to account for agreement and disagreement on moral explanatory claims among speakers with different moral standards just as, according to the Master Argument, it allows her to account for agreement and disagreement on categorical moral sentences among such speakers. Consider again the two speakers S_1 and S_2 above, both of whom employ only one standard, though different ones, and who say, respectively, 'M-ness causes E' and 'M-ness does not cause E'. On the amended proposal, the content of S_1's remark is that N_1-ness causes E and ■ all N_1's are M, while the content of S_2's remark is that N_2-ness does not cause E and ■ all N_2's are M. Looking at the non-moral component of what they say does not help locate their disagreement; however, the moral component of what they say, which captures the only standard employed, respectively, by S_1 and by S_2, explains how they can be disagreeing with one another, showing that their disagreement is non-cognitive in nature[50].

Or, as I say, so one might think. In order to assess whether the amended proposal allows the non-cognitivist to avoid the objection I have raised in this section, we need to take note of the fact that that objection focuses on cases of agreement and disagreement among speakers with different moral standards, and so far we have only considered cases of disagreement. Consider then a third speaker S_3, who also employs only one standard, though a different one from that

---

[50] Remember that the clause expressing the moral component of what S_1 says is shorthand for ' ■ all N_1's are M *and* not (■ all N_2's are M) and not (■ all N_3's are M) and …', while the clause expressing the moral component of what S_2 says is shorthand for ' ■ all N_2's are M *and* not (■ all N_1's are M) and not (■ all N_3's are M) and …'. Just as in the case of Hare's use of the Master Argument, I am here bracketing a possible criticism of the suggestion on behalf of the non-cognitivist I am considering. According to this suggestion, what is meant to allow the non-cognitivist to vindicate the intuition that S_1 and S_2 are disagreeing with one another is that, on the amended proposal, one is rejecting a general principle the other holds, which means that the non-cognitivist is not in a position to claim she can vindicate the intuition that S_1 and S_2 are disagreeing with one another, unless she can account for what it is to reject a general principle. But, so goes the possible criticism, it is not clear that she can do that, given the worry that non-cognitivism may have trouble dealing with negation. Once again, though I find this worry compelling, in the present work I am granting that non-cognitivists can answer it, so I'll bracket this possible problem for the suggestion I am considering.

employed by either S$_1$ or S$_2$: say, that ∎ all N$_3$'s are M; and suppose S$_3$ says 'M-ness causes E'. Compare now the remarks made by S$_1$ and S$_3$. On the amended proposal, the content of S$_1$'s remark is that N$_1$-ness causes E and ∎ all N$_1$'s are M, while the content of S$_3$'s remark is that N$_3$-ness causes E and ∎ all N$_3$'s are M. Whether one looks at the non-moral or at the moral component of what they say, it follows from the amended proposal -just as it does from Blackburn's original proposal- that S$_1$ and S$_3$ are not agreeing with one another. Indeed, something else also follows from the amended proposal –which does not follow from Blackburn's original proposal-, as can be seen by noting that everything a friend of the revised proposal has to say about S$_1$ and S$_2$ in order to account for their *dis*agreement holds just as true of S$_1$ and S$_3$. If the moral component of what S$_1$ and S$_2$ say shows they are disagreeing with one another, then so does the moral component of what S$_1$ and S$_3$ say: the very considerations that are meant to allow a friend of the amended proposal to account for the disagreement between S$_1$ and S$_2$ commit her to the claim that S$_1$ and S$_3$ are disagreeing with one another as well.

But these consequences of the amended proposal are false: intuitively, two speakers like S$_1$ and S$_3$ are as a rule agreeing with one another. Suppose I say 'Justice makes institutions stable' and you say 'Justice makes institutions stable', and suppose again I agree with Rawls and you with Nozick. Does this mean that, in saying what we say, we are not agreeing with one another? Does it mean that, in saying what we say, we are actually disagreeing with one another? Hardly. *Contra* the amended proposal –and Blackburn's original proposal-, in saying what we say we are agreeing with one another, despite the fact that we have different theories of justice; indeed, the verdict yielded by the amended proposal that we are disagreeing with one another is, if anything, even less plausible that that yielded by Blackburn's original proposal, according to which we are talking past one another.

134

Even supposing that these consequences of the amended proposal are true and granting that the proposal is true as well, in our actual moral practice we do not in any case treat speakers like S_1 and S_3 above as disagreeing with one another: quite to the contrary, as a rule we take them to agree on the relevant moral explanatory claim. Suppose again I say 'Justice makes institutions stable' and you say 'Justice makes institutions stable'. Coming to realize that I agree with Rawls and you with Nozick does not lead us to conclude that we are not agreeing after all, and indeed that we are disagreeing with one another: rather, it leads us to conclude that, despite our differences, we at least agree on the moral explanatory claim in question.

If the amended proposal is true and the content of moral explanatory claims is what it says it is, we are mistaken in treating speakers like S_1 and S_3 above the way we do: this is an aspect of our practice with regard to such claims that needs to be revised. And, for reasons similar to those I pointed out above, this is a very serious mistake to detect in our practice of offering, discussing, and accepting moral explanatory claims. As I noted above, one of the central aspects of any such practice with regard to claims of any kind is the way in which we detect agreement and disagreement about claims of the relevant kind: if we treat as agreeing with one another speakers who are instead disagreeing, something defective lies at the very heart of our practice.

Once again, in order to appreciate how deep a revision of our actual practice a friend of the amended proposal is committed to calling for, imagine what the revised practice would look like. Consider a moral explanatory claim of the form 'M-ness causes E'. Surely I cannot claim (as I did when discussing Blackburn's original proposal) that our actual practice with regard to a claim like this would sunder into many, parochial discussions of different issues, one for each different moral theory, since on the amended proposal it would continue to be shaped like a

single discussion that can be joined by parties who subscribe to many, different theories of the relevant moral feature. Still, there would be only one possible upshot for the discussion among such parties: they would be bound to disagree with one another. If this does not already strike you as something very different from our actual practice, I invite you again to think about how often we currently compare different moral theories with respect to more specific claims, including moral explanatory claims, and take note of cases in which friends of two different theories can, despite their differences, agree on a certain moral explanatory claim; and about how important we take such comparisons to be, given that we take some of the most interesting instances of agreement on specific issues to be precisely instances of agreement among friends of different theories. All this would have to go. So the amended proposal cannot really help the non-cognitivist avoid the objection I have raised in this section, after all: it too is false, and in any case unable to accommodate our actual practice[51].

Remember how we came up with that proposal: the comparison with Hare's account of categorical moral sentences pointed us in its direction. We started by noting that, in Hare's view, the Master Argument shows that an account of categorical moral sentences that relies only on non-enriched singular substitution is unable to deal with agreement and disagreement among speakers with different moral standards, and that this is why his account relies (at least for instances of the central, more typical use of such sentences) on enriched singular substitution; we then noted that, somewhat surprisingly, Blackburn's proposal about moral explanatory claims relies only on non-enriched singular substitution, so it is an easy target for the objection I am discussing in this section; we finally came up with the suggestion that a proposal about moral explanatory claims that relies on enriched singular substitution might help the non-cognitivist

---

[51] It is easy to see that, just like Blackburn's original proposal, the amended proposal is also vulnerable to the two objections I raised in the two previous sections (this should come as no surprise, given that Hare's account of categorical moral sentences, after which the amended proposal is modeled, is also vulnerable to them).

avoid that objection. I have argued that it really does not: it too is unable to deal with agreement and disagreement among speakers with different moral standards. Shouldn't this mean that Hare's account too, despite its reliance on enriched singular substitution, is unable to deal with that? It should, and it does.

Consider again the three speakers S_1, S_2, and S_3 above, each of whom employs only one standard, though a different one from that employed by either of the other two, and suppose this time S_1 says 'x is M', S_2 says 'x is not M', and S_3 says 'x is M'. As we saw in section 3.2, Hare objects to the cognitivist who believes that moral predicates can be analyzed in terms of non-moral ones that she cannot account for the disagreement between S_1 and S_2, whereas he can. On (b.H), which relies on enriched singular substitution, the content of S_1's remark is that x is N_1 and ■ all N_1's are M, while the content of S_2's remark is that x is not N_2 and ■ all N_2's are M: the moral component of what S_1 and S_2 say shows, according to Hare, that they are disagreeing with one another.

Compare now S_1's with S_3's remark, whose content on (b.H) is that x is N_3 and ■ all N_3's are M. Whether one looks at the non-moral or at the moral component of what they say, it follows from (b.H) -just as, according to Hare, it does from the cognitivist's account- that S_1 and S_3 are not agreeing with one another. Indeed, something else also follows from (b.H) –which, as Hare would agree, does not follow from the cognitivist's account-. If the moral component of what S_1 and S_2 say shows they are disagreeing with one another, then so does the moral component of what S_1 and S_3 say: the very considerations that are meant to allow Hare to account for the disagreement between S_1 and S_2 commit him to the claim that S_1 and S_3 are disagreeing with one another as well.

But of course these consequences of (b.H) are false: intuitively, two speakers like S_1 and S_3 are as a rule agreeing with one another; indeed, the verdict yielded by (b.H) that they are disagreeing with one another is, if anything, even less plausible that that yielded, in Hare's view, by the cognitivist's account, according to which they are talking past one another. So here is something else we can learn from the comparison between Blackburn's proposal and Hare's account: Hare's account too is unable to deal with agreement and disagreement among speakers with different moral standards. And this is quite a surprising lesson, given Hare's reliance on the Master Argument: it means that Hare's account is vulnerable to the same objection Hare thinks is fatal to the cognitivist's account. Non-cognitivists have typically advertised their meta-ethical theory as the one best able to account for agreement and disagreement among speakers with different moral standards: surprisingly, it turns out that Hare's account of categorical moral sentences will not help them make good on their claim[52].

Non-cognitivist accommodationist proposals are meant to mimic the accommodationist proposals available to a cognitivist that I described in chapter 2. As pointed out in the two previous sections, it is therefore instructive, in studying non-cognitivist proposals, to ask whether the objections they face also apply to the proposals available to a cognitivist. The cognitivist

---

[52] In the text I argue that Hare's account of categorical moral sentences must be false, since it has false consequences. It is easy to see that considerations analogous to those I offered when discussing the amended version of Blackburn's proposal would show that, even granting that Hare's account is true, it fails to accommodate our actual practice; however, as I did with the objections raised in the two previous sections and for the same reasons given there (see n. 33), I won't argue for this. The objection raised in the text to Hare's use of the Master Argument is the one I promised in n. 17. It is instructive to compare it with a different criticism of Hare's argument, one that I mentioned in the same footnote and said I would bracket. This other criticism objects that Hare is unable to account for the disagreement between S_1 and S_2, given non-cognitivism's trouble dealing with negation. The criticism in the text grants that Hare can do that and argues that the very assumption that he can ends up generating problems for him elsewhere, i.e. in accounting for the agreement between S_1 and S_3. So we have here the materials for a dilemma for Hare: either he can or he cannot account for the disagreement between S_1 and S_2; if he cannot, then he is unable to account for agreement and disagreement among speakers with different moral standards; if he can, he ends up being unable to account for the agreement between S_1 and S_3, so he is still unable to account for agreement and disagreement among speakers with different moral standards; either way, Hare is unable to account for agreement and disagreement among speakers with different moral standards. Said otherwise, what I have shown is that this charge can be pressed against Hare *whether or not* non-cognitivists can in the end deal with negation.

proposals described in chapter 2 (be they about moral explanatory claims or categorical moral sentences) specify the content (or payoff) of a moral sentence of the relevant type *via* reference to the *n*'s, i.e. the non-moral properties that can be denoted by some non-moral predicate 'N' such that ∎ all N's are M. As a result, they assign the same content (or payoff) to a moral sentence, as uttered by any user of moral language, and so have an easy time accounting for agreement and disagreement, even among speakers who have different views about what properties are the *n*'s. Two speakers S_1 and S_2 may well have different views about what properties these are; to say the same thing in a more roundabout way: the non-moral properties that can be denoted by some non-moral predicate 'N' such that *S_1 believes that* ∎ all N's are M may well differ from the non-moral properties that can be denoted by some non-moral predicate 'N' such that *S_2 believes that* ∎ all N's are M. Still, even if these properties do differ, on those cognitivist proposals a moral sentence has the same content (or payoff), as uttered by either S_1 or S_2, since its content (or payoff) is specified *via* reference to the *n*'s, not to these properties.

On the other hand, Blackburn's proposal about moral explanatory claims (in either the original or the amended version) and Hare's account of categorical moral sentences specify the content of a moral sentence of the relevant type, as uttered by a speaker S, *via* reference to the n_S's, i.e. the non-moral properties that can be denoted by some non-moral predicate 'N' such that S believes that ∎ all N's are M. These properties will be different for two users of moral language S_1 and S_2 who have different moral standards, so on these non-cognitivist proposals a moral sentence will not have the same content, as uttered by S_1 or by S_2. As we saw, the idea behind Hare's use of the Master Argument is that this is not a problem when it comes to dealing with agreement and disagreement between S_1 and S_2, provided one's account of categorical moral sentences relies on enriched singular substitution: agreement and disagreement on such

sentences are non-cognitive in nature, and the non-cognitive component that such an account assigns to the content of categorical moral sentences will explain them; and we have extended this idea to moral explanatory claims by means of the amended version of Blackburn's proposal. But this idea, I have argued, can at most seem plausible when considering cases of disagreement: if we try to apply it to cases of agreement, it turns out to be quite implausible. The non-cognitivist proposals we have looked at in this chapter thus face a problem here, one not faced by the cognitivist proposals described in chapter 2[53].

What is the source of this difference? The cognitivist proposals described in chapter 2 are built on theses i)-iv), which put forward a straightforwardly metaphysical picture. As a result, those proposals specify the content (or payoff) of a moral sentence of the relevant type *via* reference to the non-moral properties picked out by *the brute moral truths* there are: these truths are the same for all users of moral language, and so is the content (or payoff) of a moral sentence on those proposals. If you are a non-cognitivist, theses i)-iv) are not available to you. You can adopt Hare's way of fleshing out the Assumption and build your proposals on theses i.H)-iv.H), which formulate an epistemological view. If you do, your proposals will specify the content of a

---

[53] In the text I claim that the cognitivist proposals described in chapter 2 are able to account for agreement and disagreement. But haven't we met the cognitivist in this chapter and said she was unable to do that? Yes: one of the claims made by the Master Argument is that the cognitivist who believes that moral predicates can be analyzed in terms of non-moral ones is unable to account for agreement and disagreement, and I have not disputed this claim. Note, however, that (*a*) cognitivists are not committed to the claim that moral predicates can be analyzed in terms of non-moral ones. Note, moreover, that (*b*) the account of categorical moral sentences the Master Argument ascribes to the cognitivist who does accept this claim differs from the cognitivist proposals of chapter 2 precisely in the respect pointed out in the text: according to the Master Argument, if a moral predicate can be analyzed in terms of a non-moral one, the non-moral predicate in question is to be determined by the standards a speaker uses in applying the relevant moral predicate, so that the cognitivist under consideration is committed to an account that specifies the content of a categorical moral sentence, as uttered by a speaker S, *via* reference to the non-moral properties that can be denoted by some non-moral predicate 'N' such that *S believes that* ■ all N's are M. When I presented the Master Argument, I noted (see n. 17) that one standard criticism raised against it is precisely that the cognitivist the argument criticizes is not committed to such an account, but I said that (though I find it compelling) I would bracket this criticism and grant that the cognitivist in question is so committed. *Contra* Hare's use of the Master Argument, I have argued that, even granting this much, Hare does not actually do a better job of accounting for agreement and disagreement than the cognitivist who believes that moral predicates can be analyzed in terms of non-moral ones, and that the cognitivist can indeed do better than Hare on this score, if she gives up such a claim and avails herself of the proposals described in chapter 2.

moral sentence of the relevant type, as uttered by a speaker S, *via* reference to the non-moral properties picked out by *the direct moral beliefs* S holds: these beliefs are not the same for all users of moral language, and neither is the content of a moral sentence on your proposals. Since the idea of accounting for agreement and disagreement by relying on enriched singular substitution ultimately fails, you end up facing a problem not faced by the cognitivist.

Here is, therefore, a third reason, in addition to the two I gave in the two previous sections, to think that it does make a difference whether you are a cognitivist or a non-cognitivist, an additional reason to believe that there are serious limits to the degree to which a non-cognitivist can mimic what a cognitivist can do, at least if the non-cognitivist tries to do the mimicking by relying on Hare's way of fleshing out the Assumption. If you are a non-cognitivist opponent of moral explanations who wants to go accommodationist then, in the absence of a different non-cognitivist reading of the Assumption, you will not be able to build accommodationist proposals able to mimic in all relevant respects the accommodationist proposals available to a cognitivist. And, once again, the lesson here applies quite generally to the 'Go by the speaker's standards' strategy (of which Blackburn's accommodationist proposal, in either the original or the amended version, can be seen as a particular application), and the related doctrine of the descriptive meaning of moral expressions, so central to the non-cognitivist tradition. That general strategy is meant to work by mimicking the 'Go by the M-making properties' strategy available to cognitivists that I described in chapter 2 and is built on Hare's way of fleshing out the Assumption: the arguments in this section provide a third reason to believe that the 'Go by the speaker's standards' strategy cannot allow a non-cognitivist to carry the mimicking very far.

It is time now to recap the main thread of the present chapter. In chapter 2 I presented three types of proposals about moral explanatory claims that a cognitivist opponent of moral explanations can avail herself of in order to accommodate our practice with regard to such claims. These accommodationist proposals are all built on theses i)-iv), which represent the most natural way of fleshing out the Assumption; and they can be seen as the application to the particular case of moral explanatory claims of an entirely general strategy that can be used to specify the content (or payoff) of all moral sentences of the relevant type, what I called the 'Go by the M-making properties' strategy. Theses i)-iv) put forward a straightforwardly metaphysical picture and are not available to a non-cognitivist. However, non-cognitivist opponents of moral explanations claim they can provide a different way of fleshing out the Assumption and rely on it in order to build accommodationist proposals that mimic the proposals available to a cognitivist. After presenting the first of the two non-cognitivist readings of the Assumption put forward in the literature, which is due to Hare (section 1), I introduced the accommodationist proposal built on it, which is due to Blackburn, pointing out that Blackburn's proposal can be seen as the application to the particular case of moral explanatory claims of an entirely general strategy, central to the non-cognitivist tradition, that non-cognitivists have used to specify the content of all moral sentences of the relevant type, one meant to allow them to mimic the cognitivist 'Go by the M-making properties' strategy: I called this the 'Go by the speaker's standards' strategy (section 2). I then offered three different reasons to reject Blackburn's proposal, three different arguments for the conclusion that the proposal is false, and does not in any case deliver the good it is advertised for: it fails to accommodate our actual practice (sections 3, 4, and 5). If you are a non-cognitivist opponent of moral explanations who wants to go accommodationist, you need a better accommodationist proposal than Blackburn's.

Actually, my arguments show that what is at stake for you is much more than the prospects of your accommodationist project. The most important class of sentences that a meta-ethical theory needs an account of is the class of categorical moral sentences, and the account that results from the application of the 'Go by the speaker's standards' strategy to these sentences is, as we saw, Hare's account. All the objections I raised against Blackburn's proposal apply to Hare's account of categorical moral sentences as well: they show that that account is false too. Said otherwise, the problems I pointed out are not specific to Blackburn's proposal about moral explanatory claims; they are general problems with the 'Go by the speaker's standards' strategy, which show that that strategy does not actually work: it is not a good recipe to assign content to moral sentences. Given the centrality in the non-cognitivist tradition of this strategy, and of the related doctrine of the descriptive meaning of moral expressions, this is a point worth stressing. My arguments show that the doctrine of the descriptive meaning of moral expressions can play no role in the articulation of a plausible version of non-cognitivism, one able to account for the content of moral sentences, starting from the central class of them made up by categorical moral sentences.

There is also another reason why there is more at stake for you here than the prospects of your accommodationist project. Non-cognitivists typically claim they can mimic many of the things a cognitivist can do. Among the things they claim they can mimic are, as we saw, the cognitivist proposals described in the previous chapter: the 'Go by the speaker's standards' strategy is meant to allow them to mimic the cognitivist 'Go by the M-making properties' strategy, and more specifically the proposals of type A and type B available to a cognitivist, that I presented in chapter 2. I argued that my objections do not apply to the cognitivist proposals of chapter 2, and so that the 'Go by the speaker's standards' strategy cannot really help the non-

cognitivist mimic those cognitivist proposals. This is so whether one is considering proposals about moral explanatory claims or categorical moral sentences: the problem is, once again, with the general strategy itself. If my arguments are correct, then the 'Go by the speaker's standards' strategy, and the related doctrine of the descriptive meaning of moral expressions, can play no role in the articulation of a version of non-cognitivism able to mimic, as non-cognitivists typically claim they can do, many of the things available to a cognitivist.

Finally, the objection I have raised in this section points out a third reason why there is much more at stake for you here than the prospects of your accommodationist project. What I have argued in this section is that the non-cognitivist proposals that result from the application (either to moral explanatory claims or to categorical moral sentences) of the 'Go by the speaker's standards' strategy are unable to account for the phenomenon of agreement and disagreement among speakers with different moral standards. This phenomenon has traditionally been claimed by non-cognitivists as their home turf, so to speak: the standard non-cognitivist position has been that non-cognitivism is able to account for it, indeed it is the meta-ethical theory best able to account for it, so much so that what deserves to be called the 'Master Argument' for non-cognitivism appeals precisely to this phenomenon. If my arguments in this section are correct, it turns out that the 'Go by the speaker's standards' strategy for specifying the content of moral sentences fails the non-cognitivist where she has traditionally claimed to be at a decisive advantage, leaving her without the support of what is meant to be one of the crucial reasons in favor of her meta-ethical position. That strategy, and the related doctrine of the descriptive meaning of moral expressions, can play no role in the articulation of a defensible version of non-cognitivism, one able to make good on those traditional non-cognitivist claims and retain the support of anything like the Master Argument.

144

As we saw, both Blackburn's proposal and, more generally, the 'Go by the speaker's standards' strategy are built on Hare's way of fleshing out the Assumption: for the sake of both your accommodationist project and your meta-ethical theory, you really need an alternative reading of the Assumption, on which to build a better accommodationist proposal for moral explanatory claims and, more generally, a better strategy for assigning content to moral sentences. As I mentioned, Hare's way of fleshing out the Assumption is one of two different readings of it that non-cognitivists have put forward: starting from the next chapter, we'll look at the second one and see whether it can help you solve the problems I have raised in this chapter.

Before turning to it, however, I want to end this chapter by going back to one of the arguments I have offered. One of the claims I have argued for in this section is that Hare's account of categorical moral sentences is unable to deal with agreement and disagreement among speakers with different moral standards. Given Hare's reliance on the Master Argument, my argument for this conclusion deserves further scrutiny: more specifically, I want to consider two possible suggestions one could make on Hare's behalf to help him avoid the problem I have raised. As it turns out, my discussion will also serve as an introduction to the alternative non-cognitivist reading of the Assumption we'll be looking at starting from the next chapter.

6. Hare on agreement and disagreement: further thoughts

Both the suggestions I want to consider propose that Hare should modify his account of categorical moral sentences in order to avoid the objection I raised in the previous section; they differ on the modifications they recommend[54]. The first suggestion proposes to implement the

---

[54] As we saw, Hare's account distinguishes between a class of peripheral uses of categorical moral sentences and a central, more typical use of such sentences: in my discussion in the present section, I'll consider only how his account handles instances of the central use, since this is the part of the account that, Hare claims, allows him to deal with agreement and disagreement among speakers with different moral standards and, I argued, fails to do this. For

strategy of 'going by the speaker's standards' in a slightly different way from that described in 3.2; the easiest way to introduce it is by focusing on individual moral sentences. On Hare's account, the content of a sentence of the form 'x is M' has two components, a non-moral and a moral one. One noteworthy feature of the account is that the moral component is specified in general terms: it is captured by a universal moral sentence, say ' ■ all N_1's are M' for a speaker, like S_1 above, whose only direct moral belief is that ■ all N_1's are M. This moral component is non-cognitive in nature; to put the point in terms of Hare's version of non-cognitivism, the sentence capturing it can be thought of as an imperative prescribing to do (avoid, etc., depending on what 'M' is) whatever is N_1. One natural response to the objection I raised in the previous section is to suggest that Hare should specify the moral component of the content of 'x is M' in particular rather than general terms: he should say it is captured by an imperative prescribing to do (avoid, etc., depending on what 'M' is) x. The suggestion would thus be that the content of 'x is M', as uttered by S_1, is captured by 'x is N_1 and do (avoid, etc.) x!'.

To see how this might help with the objection raised in the previous section, consider a speaker, like S_3 above, who also holds only one direct moral belief, though a different one from that held by S_1, say that ■ all N_3's are M, and suppose that both S_1 and S_3 say 'x is M'. Intuitively, two speakers like S_1 and S_3 are as a rule agreeing with one another, and the suggestion under consideration would allow Hare to account for this. On this suggestion, the content of 'x is M', as uttered by S_3, is captured by 'x is N_3 and do (avoid, etc.) x!'. Looking at the non-moral component of what S_1 and S_3 say does not help locate their agreement; however, the moral component of what they say explains how they can be agreeing with one another,

showing that their agreement is non-cognitive in nature: they are both prescribing to do (avoid, etc.) x[55].

The reason why this suggestion cannot help Hare avoid the objection I raised in the previous section is that it is in deep tension with Hare's claim that there is an important difference between an individual moral sentence of the form 'x is M' and an imperative sentence of the form 'Do (avoid, etc.) x!'. As we saw at the beginning of this chapter, this claim is central to Hare's meta-ethical theory. According to Hare, the function of both these kinds of sentences is to guide choices and actions: both can be used to frame answers to questions of the form 'What shall I do?'. However, the answers they can be used to frame differ in an important respect.

Someone who says 'Do (avoid, etc.) x!' is offering guidance only with respect to the particular situation under consideration. Since 'x is M' entails, on Hare's view, 'Do (avoid, etc.) x!', a speaker who says 'x is M' is also offering guidance with respect to the particular situation under consideration, but she is doing more than that: she is, according to Hare, offering guidance with respect to a whole *kind* of cases (of which the situation under discussion is an example), committing herself not just to 'Do (avoid, etc.) x!' addressed to the particular hearer in the particular situation at issue, but also to a whole range of other such imperative sentences. Within the linguistic tools we have available to guide choices and actions, what is specific to moral (and, more generally, evaluative) language, on Hare's view, is precisely that it is the tool that enables us to offer guidance with respect to *kinds* of choices and actions.

On this view of moral language, (*a*) a speaker who says 'x is M' is saying more than just 'Do (avoid, etc.) x!': in particular, she is committing herself to other imperative sentences like this, in addition to this one. Moreover, (*b*) since agreeing with someone who says 'x is M' is

---

[55] For simplicity's sake, I won't consider how the suggestion in question deals with the negation of an individual moral sentence and I'll grant that, if it can help Hare account for cases of agreement, it can help him with cases of disagreement as well.

taking a moral position, it takes more to agree with such a speaker than merely saying 'Do (avoid, etc.) x!': in particular, it takes committing oneself to other imperative sentences like this, in addition to this one. But neither of these things is true on the suggestion we are considering.

On this suggestion, (*a*) a speaker like S_1 who says 'x is M' is not committing herself to other imperative sentences in addition to 'Do (avoid, etc.) x!'. She is no doubt saying something else, in addition to 'Do (avoid, etc.) x!': she is saying 'x is N_1'. But this is not an imperative sentence. To put the point in a different way, 'x is N_1' captures the non-moral component of what S_1 says; so far as the moral component of what she says, i.e. the guidance she is offering, is concerned, she is not prescribing anything else but to do (avoid, etc.) x: this is, after all, the whole point of the suggestion under consideration. On this suggestion, (*b*) it is possible to agree with S_1 without committing oneself to other imperative sentences in addition to 'Do (avoid, etc.) x!': this is precisely what, on the suggestion in question, happens with a speaker like S_3. S_3 is no doubt saying something else, in addition to 'Do (avoid, etc.) x!': she is saying 'x is N_3'. But this is not an imperative sentence. And, after all, on the suggestion we are considering it is not meant to play any role in accounting for the agreement between her and S_1: 'x is N_3' captures the non-moral component of what S_3 says and, as we saw, the non-moral component of what S_1 and S_3 say is not what is meant to account for their agreement. The suggestion under consideration is thus at odds with Hare's view of moral language, so it cannot help him avoid the objection I raised in the previous section.

What about a non-cognitivist who is willing to give up Hare's view and endorse a picture of moral language on which the gap between an individual moral sentence of the form 'x is M' and an imperative sentence of the form 'Do (avoid, etc.) x!' is narrower than Hare would like? Can't such a non-cognitivist adopt the suggestion in question and avoid the objection that she

cannot account for agreement and disagreement on categorical moral sentences among speakers with different moral standards? In principle, she can; however, there are good reasons to think she should not.

To see why, let us begin by taking note of the fact that the suggestion we are considering is not available in the case of moral explanatory claims: it just cannot be applied to claims of the form 'M-ness causes E'. This means that at most the non-cognitivist we are imagining can avoid the objection I raised in the previous section against Hare's account of categorical moral sentences, but she cannot avoid the parallel objection I raised against the amended version of Blackburn's proposal: such a non-cognitivist still lacks an accommodationist proposal able to account for agreement and disagreement on moral explanatory claims among speakers with different moral standards. It also means that the non-cognitivist we are imagining cannot provide a uniform semantic account of the content of moral sentences: she is committed to telling different semantic stories for different kinds of moral sentences. This already sounds like a cost; it sounds even more like one if we look back at some of the points I made when presenting Hare's strategy of 'going by the speaker's standards' in 3.2.

As we saw, Hare adopts this strategy in order to account for a certain feature of the way in which we use categorical moral sentences, one that presents him with a problem because it seems to call for a cognitivist treatment. This type of problem, I pointed out, is not specific to categorical moral sentences: certain features of the way in which we use other moral sentences, such as moral explanatory claims, also seem to call for a cognitivist treatment, and so present non-cognitivists with a problem. The particular problem Hare faces, moreover, is not specific to categorical moral sentences either. As we saw, Hare's problem is to explain our use of categorical moral sentences to convey non-moral information, and moral explanatory claims can

be used the same way: to adapt the example from Hare I used at the beginning of 3.2, a parson can tell another 'Being a good person makes you feel at peace with yourself' in order to convey the information that going to church and obeying a certain set of Biblical injunctions make one feel at peace with oneself. So there is a type of problem for non-cognitivists that arises for different kinds of moral sentences, and Hare's specific problem also arises for different kinds of moral sentences.

One noteworthy feature of Hare's strategy of 'going by the speaker's standards' is that it can be (and, as we saw, it has been) applied to other sentences, in addition to categorical moral sentences: it provides the non-cognitivist with a uniform way of dealing with what is, by all appearances, one and the same (type of) problem non-cognitivists have, one that has nothing specific to do with categorical moral sentences. In the previous section, I raised the objection that the proposals that result from the application of this strategy (be they about categorical moral sentences or moral explanatory claims) are unable to account for agreement and disagreement among speakers with different moral standards. What the non-cognitivist we are imagining proposes to do is to implement the strategy of 'going by the speaker's standards' in a different way than Hare does, one that would allow her to avoid my objection in the case of categorical moral sentences, though it cannot be applied to moral explanatory claims.

But this, I should think, is bad theorizing: if a certain way of dealing with one and the same (type of) problem that arises for different kinds of sentences does not work, it is no good solution to come up with a different way of dealing with it that is only available for sentences of one kind. If the problem really is (as by all appearances it is) one and the same for different kinds of sentences, the adoption of a solution that is only available for one kind of sentences is theoretically unmotivated: such a solution is objectionably *ad hoc*.

So a non-cognitivist willing to give up Hare's picture of moral language can in principle adopt the suggestion we are looking at and put forward an account of categorical moral sentences that avoids the problem I raised in the previous section. If she does, however, she cannot provide a uniform semantic account of the content of moral sentences, and this is problematic both because there are general reasons for preferring a uniform semantic account and because there are specific reasons for thinking that the solution such a non-cognitivist adopts to avoid the problem raised in the previous section is objectionably *ad hoc*. If you are a non-cognitivist, for the sake of your meta-ethical theory you should therefore not adopt the suggestion in question. Moreover, if you are a non-cognitivist who wants to accommodate our practice with regard to moral explanatory claims, that suggestion is of no help for your accommodationist project. This first suggestion thus does not change much with respect to the conclusions I reached in the previous section[56].

The easiest way to introduce the second suggestion I want to consider on Hare's behalf is by means of the cognitivist proposals about categorical moral sentences I described in chapter 2. Those proposals, I argued, are not vulnerable to the objection I raised in the previous section and, as we saw, Hare's account of categorical moral sentences can be seen as an attempt to mimic the cognitivist who relies on singular substitution. One natural response to the objection I raised in the previous section is to grant it shows that such an attempt is unsuccessful and suggest that Hare should instead mimic the cognitivist who relies on general substitution. The point here is actually quite general: the cognitivist proposals, I argued, are not vulnerable to *any* of the objections I have raised in this chapter, so a natural response to each of those objections is to grant it shows that Hare's attempt to mimic the cognitivist who relies on singular substitution is

---

[56] It is at least arguable that the second pattern of analysis put forward by Stevenson [1944] amounts to the suggestion just discussed in the text rather than to the account of categorical moral sentences presented in 3.2: this is the third reason I promised in n. 24 for ascribing that account to Hare rather than Stevenson.

unsuccessful and suggest that he should mimic the cognitivist who relies on general substitution instead. Focusing again, as I did with the first suggestion, on individual moral sentences, the second suggestion I want to consider is thus that Hare should modify his account in order to mimic the proposal that the content of 'x is M' is that x has the n[57].

One question to ask about this suggestion is: *how* exactly should Hare do that? What account, different from the one we have been looking at in this chapter, can he put forward, using only materials available to a non-cognitivist, that would allow him to mimic that cognitivist proposal? Before addressing this question, however, I want to argue that, even granting it is available to him, this suggestion cannot in any case help Hare avoid the objection I raised in the previous section (or any of those I have raised in this chapter). The reason is that it too is at odds with a claim that is central to Hare's meta-ethical theory, namely the doctrine of the descriptive meaning of moral expressions.

As we saw, Hare emphasizes that there are similarities, in addition to differences, between moral and non-moral sentences, stressing that we can use a sentence of the form 'x is M' to convey non-moral information about x; and his account is driven by the idea that the way to deal with this aspect of moral language is by saying that (*1*) the content of 'x is M', as uttered by a speaker S, has a non-moral component, one that can be captured replacing the moral predicate with a non-moral one, determined by S's standards. That is, a parson can use 'Marta is a good girl' to convey the non-moral information that Marta goes to church and obeys a certain

---

[57] Some writers in meta-ethics read Hare as putting forward an account along these lines, instead of the account I ascribed to him in 3.2: for an example, see Harman [1977], ch. 7. What follows can be taken (together with the considerations offered in 3.2) as providing my reasons in favor of my reading of Hare. Since for simplicity's sake I am (see n. 54) focusing on users of moral language who have made up their minds about what direct moral beliefs to accept, for a given moral predicate 'M', and hold only one such belief, the cognitivist proposal I mention in the text is the one, among those relying on general substitution, for the case in which there is only one brute moral truth ascribing M-ness; when considering speakers who have made up their minds and hold more than one direct moral belief, the cognitivist proposals to look at are the ones, among those relying on general substitution, for the case in which there is more than one brute moral truth.

set of Biblical injunctions because (*2*) that Marta goes to church … is part of the content of 'Marta is a good girl', as uttered by him: it is, as Hare puts it, the descriptive meaning of that sentence, as uttered by the parson. The doctrine of the descriptive meaning of moral expressions is what, according to Hare, accounts for the similarities between moral and non-moral sentences he emphasizes; as we saw, it is the tool that the non-cognitivist tradition has borrowed from him and applied more broadly, to deal with other features of our use of moral expressions that seem to call for a cognitivist treatment. But an account mimicking the proposal that the content of 'x is M' is that x has the *n* has to give up such a doctrine.

This comes out if we remind ourselves of what such a proposal exactly says. Remember the two conventions I adopted in chapter 2 in framing the proposal in question as I did: I said that (*a*) I would use 'the *n*' as shorthand for the definite description 'the non-moral property that can be denoted by some non-moral predicate 'N' such that ■ all N's are M' and (*b*) in using it to frame that proposal, I would assume this definite description is treated *à la* Russell (so as to avoid having to write down the Russellian *analysans* every time I needed to frame the proposal), which means that 'x has the *n*' is an existential claim. On an account along the lines of this proposal, (*a*) the sentence capturing the content of 'x is M' replaces the moral predicate with an expression, 'the *n*', that still contains moral terminology and is independent of the speaker's standards, *contra* (*1*). On such an account, (*b*) it is no part of the content of 'Marta is a good girl', even as uttered by a parson, that Marta goes to church …, *contra* (*2*); indeed, for any non-moral property that can be denoted by some non-moral predicate 'N', it is no part of the content of 'x is

M', as uttered by anyone, that x is N. Adopting such an account would thus require giving up Hare's doctrine of the descriptive meaning of moral expressions[58].

Though this shows that the suggestion we are looking at cannot help Hare avoid the objection I raised in the previous section (or any of those I have raised in this chapter), we should ask again: what about a non-cognitivist who is willing to give up Hare's doctrine of the descriptive meaning of moral expressions? Can't such a non-cognitivist adopt the suggestion in question and put forward an account of categorical moral sentences that avoids the problem raised in the previous section (indeed, all the problems raised in this chapter)? After all, one of the lessons I drew from my objections is that the doctrine of the descriptive meaning of moral expressions can play no role in the articulation of a viable version of non-cognitivism, hence that, despite its central place in Hare's meta-ethical theory (and the non-cognitivist tradition), the non-cognitivist had better be willing to give up such a doctrine. This brings us back to the

---

[58] This is the reason Hare gives to deny that an account along the lines we are considering provides an adequate characterization of the content of categorical moral sentences. Focusing on 'ought'-sentences, in ch. 12 of his [1952] Hare introduces an artificial word, '*ought*', and proceeds to define (what I am here calling) individual moral sentences framed in terms of it along the lines we are considering; he then asks whether such '*ought*'-sentences are a good substitute for individual moral sentences framed in terms of our word 'ought'. He answers they are not, since they cannot perform the descriptive job performed by our 'ought'-sentences. Hare does go on to say that one might try to narrow the gap between '*ought*'- and 'ought'-sentences with the help of some further tool, but even so he is tentative about the full adequacy of the resulting account ("the latter [*sc.* the descriptive jobs of 'ought'] are not [adequately provided for by '*ought*'] without further definition, and then *not so handily*" –p. 192-; "by this further definition, we have succeeded in giving to '*ought*' *some* of the flexibility between evaluative and descriptive uses which 'ought' has in ordinary language" –p. 194-; "*One might* claim that, if we really were suddenly deprived of the use of the ordinary value-words, we might in time, by using my substitute value-words, come to use them with the same subtlety as we used the old ones" –p. 195-; all emphasis mine, except for occurrences of '*ought*'). Moreover, the difficulties here seem to be even greater than Hare suggests. The tool by which he says one might try to enable '*ought*'-sentences to perform the descriptive job performed by 'ought'-sentences is an appeal to the inverted commas use of moral sentences I mentioned in n. 14 (p. 194). Bracketing the details of how this is meant to go, there is a difficulty of principle here: both the inverted commas and the conventional use (though Hare mentions only the first, what he writes suggests he might have both in mind) make reference to the standards established within a group of people (either a specific, usually well understood, group or people generally), but elsewhere Hare is explicit that a sentence of the form 'x is M', as used by a given speaker, can have descriptive meaning even if there are no established standards, indeed no standards at all, for the class of objects x belongs to, and the speaker is setting up a standard for such objects for the first time. Here is for example what he writes about 'good': "Sometimes we learn to use 'good' in a new descriptive meaning through being taught it by an expert in a particular field […]. Sometimes, on the other hand, we make up a new descriptive meaning for ourselves. This happens when we start having a standard for a class of objects, certain members of which we have started needing to place in order of merit, but for which there has hitherto been no standard" (p. 118). So Hare is well advised to be tentative about the possibility of enabling '*ought*'-sentences to perform the descriptive job performed by 'ought'-sentences.

question I raised above: *how* exactly can a non-cognitivist adopt this suggestion and mimic the cognitivist proposal that the content of 'x is M' is that x has the *n*? What account, different from Hare's, can she put forward, using only materials available to her, that would allow her to mimic that proposal?

As we saw in the previous chapter, all cognitivist proposals specify the content of 'x is M' *via* reference to the non-moral property picked out by the brute moral truth ascribing M-ness, but they do so in different ways. The cognitivist proposals that rely on singular substitution make reference to the *n* in the following way: they specify the content of 'x is M' by replacing the moral predicate with the non-moral predicate denoting the *n* (and adding, in the case of enriched singular substitution, a clause expressing the moral relevance of the property in question). Hare's account can be seen as an attempt to mimic this sort of cognitivist proposals. Since non-cognitivists deny there is moral truth, Hare's account specifies the content of 'x is M', as uttered by a speaker S, *via* reference to the non-moral property picked out by the direct moral belief S holds, in the following way: it replaces the moral predicate with the non-moral predicate denoting the n_S (and adds a clause expressing the moral relevance of the property in question). Though Hare's attempt to mimic the cognitivist is, I argued, ultimately unsuccessful, it is at least clear how he can proceed in order to carry it out.

The suggestion we are considering is that the non-cognitivist should mimic the cognitivist proposal that relies on general substitution. This proposal makes reference to the *n* in the following way: it specifies the content of 'x is M' by replacing the moral predicate with the definite description 'the *n*'[59]. How can the non-cognitivist proceed in her attempt to mimic this sort of cognitivist proposal? Replacing the moral predicate with the definite description 'the n_S' is a non-starter: since 'x has the n_S' lacks any moral content, the upshot would be that the content

---

[59] And making the appropriate change to the verb in the sentence.

of 'x is M', as uttered by a speaker S, is entirely non-moral, which is not the result the non-cognitivist wants (and does not mimic the cognitivist proposal in question, on which 'x is M' does have moral content)[60]. But, if not by replacing the moral predicate with the definite description 'the n_S', how else can the non-cognitivist proceed in her attempt to mimic the proposal that the content of 'x is M' is that x has the n[61]?

Here is a different way to point out the problem the non-cognitivist needs to address. Go back to the core non-cognitivist idea: in uttering a sentence of the form ' ■ All N_1's are M', a speaker S is expressing the relevant non-cognitive attitude (commendation, condemnation, etc., depending on what 'M' is) toward N_1's. To illustrate the problem, let us suppose that 'M' is 'wrong' and take this idea in the rough-and-ready way it is often taken in discussions of non-cognitivism in the literature: in uttering a sentence of the form ' ■ All N_1's are M', a speaker S is saying 'Boo to N_1's!' (this is crude, but it will do for our purposes). If this is what S is doing when she says ' ■ All N_1's are M', what can she be doing when she says 'x has the n', i.e. 'x has the non-moral property that can be denoted by some non-moral predicate 'N' such that ■ all N's are M'? Maybe saying 'x has the non-moral property that can be denoted by some non-moral predicate 'N' such that boo to N's!'? This does not even make much sense. But then how is the

---

[60] It is worth pointing out that, on the view that the content of 'x is M', as uttered by a speaker S, is that x has the n_S, the content of 'x is M' depends on the speaker (though not, as on Hare's account, on the speaker's standards); this view is thus close to what Ayer [1936], ch. 6 calls "the orthodox subjectivist theory" of ethics, according to which sentences of the form 'x is M' "express propositions about the speaker's feelings" (p. 109). This is the very first theory Ayer mentions after introducing his non-cognitivist theory, in order to *distinguish* it from his own.

[61] Strictly speaking, a problem similar to the one I describe in the text also arises, as we saw in 3.2, for Hare's attempt at mimicking the cognitivist proposal that relies on enriched singular substitution. On this proposal, the content of 'x is M' has a moral component captured by the clause 'N_1-ness is the n' (supposing the predicate denoting the n is 'N_1'); Hare's account, however, cannot have the clause capturing the moral component of the content of 'x is M', as uttered by a speaker S, be 'N_1-ness is the n_S' (even supposing the predicate denoting the n_S is 'N_1'), since 'N_1-ness is the n_S' has no moral content. As we saw, the idea here is that Hare should have the clause in question be ' ■ all N_1's are M (remember that such a clause is shorthand for ' ■ all N_1's are M *and* not (■ all N_2's are M) and not (■ all N_3's are M) and …'). Interestingly, even assuming Hare can solve his problem this way, it remains unclear how the non-cognitivist I am considering in the text can solve her problem.

non-cognitivist going to mimic the proposal that the content of 'x is M' is that x has the *n*? It is worth pointing out that the problem here is an instance of a fully general problem for non-cognitivism. Suppose I say ' ■ All N₁'s are M'; I might infer from this 'There is a non-moral property that can be denoted by some non-moral predicate 'N' such that ■ all N's are M'. If in making the first remark I am saying 'Boo to N₁'s!', what can I be doing in making the second remark? Maybe saying 'There is a non-moral property that can be denoted by some non-moral predicate 'N' such that boo to N's!'? Again, this does not even make much sense. But then how is the non-cognitivist going to account for my second remark? The problem for the non-cognitivist is that it is not clear she can avail herself of even so basic an operation as existential generalization, hence of an existential claim like 'x has the *n*'.

My aim in raising these questions is not to suggest that the non-cognitivist cannot answer them, it is rather to point out that she has work to do here. On the other hand, this work might well be worth the non-cognitivist's effort: if she can find a way to proceed in her attempt to mimic the cognitivist who relies on general substitution, and if her attempt is ultimately successful, she will be able to put forward an account of categorical moral sentences that avoids the objection I raised in the previous section (indeed, all the objections I have raised in this chapter). And she might well be able to treat moral explanatory claims the same way, putting forward an accommodationist proposal about such claims that mimics the cognitivist accommodationist proposal relying on general substitution and avoids the objections I have raised against Blackburn's proposal. Though it cannot help Hare, this second suggestion might therefore prove quite helpful for a non-cognitivist willing to give up Hare's doctrine of the descriptive meaning of moral expressions. As a matter of fact, the second non-cognitivist reading of the Assumption we'll be examining starting from the next chapter is designed precisely to

enable the non-cognitivist to take advantage of this suggestion and mimic the cognitivist who relies on general substitution. It is time to look at it[62].

Chapter 4 – The second non-cognitivist strategy: good news

1. Gibbard's reading of the Assumption

In chapter 2 I described three different proposals about moral explanatory claims that an opponent of moral explanations can put forward in an attempt to accommodate our practice with regard to such claims. The first two of these accommodationist proposals rely on singular substitution, the third on general substitution. All of them are built on the reading of the Assumption articulated by theses i)-iv); since these theses freely assume that moral sentences can be true or false, they, and the accommodationist proposals described in chapter 2, are available only to a cognitivist. Non-cognitivist opponents of moral explanations who want to go accommodationist claim they can provide a different way of fleshing out the Assumption, one that does not assume that moral sentences can be true or false (or anything else that is not available to a non-cognitivist), and build on it accommodationist proposals that mimic the proposals described in chapter 2. In chapter 3 we looked at Hare's way of fleshing out the Assumption and at the non-cognitivist accommodationist proposals built on it, which attempt to mimic the cognitivist proposals that rely on singular substitution. I argued that these non-cognitivist proposals fail. Starting from this chapter, we'll look at a different non-cognitivist reading of the Assumption, alternative to Hare's, and at the accommodationist proposal built on it, which attempts to mimic the cognitivist proposal that relies on general substitution. I'll eventually argue that this non-cognitivist proposal fails too, and that the non-cognitivist's attempt to mimic the cognitivist who relies on general substitution is unsuccessful. The argument for this conclusion will take more than one chapter. In this chapter, I'll look at the good news for

159

the non-cognitivist, showing that the accommodationist proposal in question avoids all the problems I raised in chapter 3. I'll argue that there are other problems with it in the next chapter.

The alternative non-cognitivist reading of the Assumption we'll look at in this chapter is due to Gibbard [2003]; as I did in the previous chapter when dealing with Hare's reading, I'll start by giving a brief sketch of Gibbard's alternative reading and then move on to a closer examination of it[1]. As we saw in chapter 3, Hare fleshes out the Assumption that actions, people, institutions, etc. owe their moral features to their non-moral properties in terms of the idea that use of moral language commits one to general moral principles. Here is how I sketched out his way of doing this at the beginning of 3.1: if someone says 'x is M' (where 'M' is a moral predicate), then -unless s/he is abusing moral language- s/he is appealing to a general moral principle s/he takes to apply to x and yield this verdict about it. That is, if you say 'x is M', you are, according to Hare, committed to holding a general moral principle and to believing that that principle applies to x: you must infer your belief that x is M from a general moral principle and an appropriate belief about x.

A different way of articulating the idea that use of moral language commits one to general moral principles makes the commitment a matter not, so to speak, of where your belief must come from, but of where you can go with it. Suppose you say 'x is M'. Maybe you do not infer your belief that x is M from a general moral principle and an appropriate belief about x. Maybe your moral views are rather scattered, unsystematic and incomplete, and though you have a moral opinion about certain issues, you are undecided about others and do not have much to offer when it comes to general principles able to account for the views you do hold; so maybe

---

[1] Gibbard is explicit in framing his claims as claims about both language and thought, so (just as I did in chapter 3) in what follows I'll move back and forth between claims about moral sentences and claims about moral beliefs and I'll continue to speak of someone holding a moral belief as a user of moral language and moral predicates or as a speaker.

you do not even hold a general principle you take to apply to x. Even so, use of moral language commits you to general moral principles, one might claim, in the following sense: if you were to develop your views and become fully decided, then –unless you are abusing moral language in saying 'x is M'- you *would* hold a general moral principle you would take to apply to x and yield this verdict about it (provided, of course, that you did not change your mind about x and still held it is M). Use of moral language commits you to general moral principles in that, even though you may not, due to the incomplete state of your views, hold a general principle you take to apply to x, your belief that x is M calls for one: you are committed to holding one, if you develop your views so as to become fully decided.

On both ways of articulating the idea that use of moral language commits one to general moral principles, by saying 'x is M' you incur a commitment with regard to your broader set of beliefs: your belief set is constrained in a certain way. On Hare's way of articulating that idea, the constraint concerns the set you infer the belief that x is M *from*: you are committed to holding a general moral principle you take to apply to x and yield this verdict about it. On the alternative way we are looking at, the constraint concerns the set you end up incorporating the belief that x is M *into*, supposing you develop your views so as to become fully decided: you are committed to reaching a belief set that includes a general moral principle you take to apply to x and yield this verdict about it, if you become fully decided.

Think of the contrast in terms of the rules of a game. On Hare's view, the logic of moral predicates (i.e. the rules according to which competent users of moral language use its predicates) is such that one can only make the move of saying 'x is M' if one is coming from a certain place: if one infers that x is M from a belief set including a general moral principle and an appropriate belief about x. On the alternative I am sketching out, the logic of moral predicates is

such that, if one makes the move of saying 'x is M', one is constrained with regard to the place one can go next, or rather with regard to where one can end up further down the road: one can only end up with a belief set including a general moral principle and an appropriate belief about x, supposing one becomes fully decided.

The sketch I have just given provides one way of looking at (one aspect of) Gibbard's non-cognitivist theory: Gibbard [2003] can be read as putting forward an alternative non-cognitivist reading of the Assumption in terms of the alternative way, sketched out above, of articulating the idea that use of moral language commits one to general moral principles. Before taking a closer look at Gibbard's reading of the Assumption, a few remarks about my discussion of Gibbard's views are in order. Gibbard [2003] develops a complex and ambitious meta-ethical theory, which aims at addressing some of the most persistent worries about non-cognitivism (most notably, the Frege-Geach problem) and showing that, when properly articulated and defended, non-cognitivism ends up mimicking many of the claims available to a cognitivist (more specifically, to a realist *à la* Moore, who provides, according to Gibbard, the template that an adequate meta-ethical theory ought to match). In what follows, I'll deal with only one of the many, interesting aspects of such a theory; even with respect to the aspect I'll focus on, I'll abstract away from some of the details of Gibbard's theory and concentrate on the features directly relevant to my discussion in the present work. One noteworthy feature of Gibbard's theory (already present in his earlier [1990]) is its full generality: Gibbard claims his non-cognitivist theory applies to *all* language that, to use Wilfrid Sellars' phrase, seems to be "fraught with ought", or to all normative language, as it is standard to call it nowadays, of which moral language is a prime example[2]. His theory is therefore framed as one not directly about

---

[2] Hare's theory is also meant to account for the use of such words as 'good', 'right', and 'ought' outside the moral domain, in addition to their moral use, but Gibbard goes far beyond that: for instance, he is explicit (as Hare is not)

moral predicates, but about such predicates as 'is ok to do' or 'is the thing to do': predicates that, according to Gibbard, can be used to express decisions that are part of conditional plans we form for various (actual and hypothetical) contingencies, and in whose terms the predicates of normative language, including moral predicates, can be understood. When it comes specifically to morality, Gibbard (in his [1990]) endorses the view that moral language is, within the broader domain of normative language, the tool we have to raise and discuss questions on how to feel about the actions that people, ourselves included, choose: in particular, questions about whether to feel such emotions as guilt and remorse for things we choose, and blame, resentment, outrage, or indignation for things others choose. Though the full generality of Gibbard's theory matters for certain purposes, it does not for ours: a non-cognitivist who applied Gibbard's version of non-cognitivism directly to moral language would still be able to endorse Gibbard's reading of the Assumption and use it to develop an accommodationist proposal for moral explanatory claims. In dealing with Gibbard's views, I'll therefore bracket this feature of his theory and feel free to recast the claims he makes explicitly about predicates like 'is ok to do' as claims about moral predicates. This is how Gibbard's claims are often taken in discussions in the literature, and Gibbard himself explicitly points to parallels between the two sorts of claims throughout his book, including when he puts forward his proposal to accommodate explanatory claims, so with this much warning at the outset, I expect no confusion or misunderstanding to arise[3].

Gibbard frames his reading of the Assumption as a thesis about what he calls "hyperstates", i.e. maximally decided states of mind; in order to facilitate comparison with Hare's reading, which I formulated as a set of theses about moral beliefs, I'll frame Gibbard's

---

that talk of reasons for belief, i.e. the whole domain of epistemic rationality, also falls within the purview of his non-cognitivist theory (though see Hare [1952], p. 60; I am indebted to Sturgeon [1995b] for this reference).

[3] In order to keep track of the fact that Gibbard's theory belongs to the same meta-ethical family as Hare's theory, discussed in the previous chapter, I'll continue to use the traditional label 'non-cognitivism' to refer to it. Gibbard himself prefers the label 'expressivism'.

reading too in terms of beliefs. Let a <u>complete</u> belief set be a belief set with the following feature: for any categorical moral belief, the set contains either that belief or its negation, and for any non-moral belief, the set contains either that belief or its negation. Among complete belief sets, consider those that are also consistent. One way to think of a complete belief set that is consistent is as capturing the views of a user of moral language who has become fully decided: a (*highly* idealized) speaker who has acquired a determinate opinion about any categorical moral issue and about any non-moral issue, and who has done so while keeping her views free from inconsistencies. As Gibbard points out, a complete belief set that is consistent corresponds (at least roughly) to what he calls a "fact-plan world": such a belief set has a non-moral component, which one can think of as mirroring a possible world, and a moral component, which one can (on Gibbard's version of non-cognitivism) think of as capturing a "hyperplan", i.e. a plan for any situation, one that allows or rules out the alternatives available in that situation as permissible or impermissible, and does so in a consistent way[4].

Which complete belief sets are consistent and which are inconsistent is a matter of logic. Some of them are inconsistent because they violate entirely general logical rules: for example, they contain both a categorical moral belief and its negation. Others are inconsistent because they violate the specific rules governing the use of moral predicates: for example, a requirement of consistency for a set of beliefs is, according to Gibbard, that it not rule out as impermissible

---

[4] The hedge in the text is mine: Gibbard claims (see Gibbard [2003], pp. 57-58) that talk of hyperstates or, in my terminology, complete belief sets that are consistent and talk of fact-plan worlds are interchangeable. Without trying to settle the issue, here is a reason for thinking that this claim is at least not obvious. The notion of consistency Gibbard works with is a logical notion, whereas the notion of a possible world is a metaphysical one. It is thus possible to maintain, for example, that a complete belief set can be consistent while including a non-moral belief that is not just false, but necessarily false, hence that the non-moral component of a complete belief set that is consistent can fail to mirror any possible world, say by having the belief that water is not H_2O as one of its members. Since (as he says over and over in his book) Gibbard is interested in what we, as users of moral language, are committed to in virtue of sheer logic, in what follows I'll bracket talk of fact-plan worlds and speak in terms of complete belief sets that are consistent (as Gibbard himself, in the passage referred to above, recommends one should do, if one has worries about talk of fact-plan worlds). For Gibbard's take on how to deal with the philosophical issues raised by such identity statements as 'Water is H_2O', see Gibbard [2003], ch. 6.

all the alternatives available in a given situation, since a set of beliefs doing this would fail to offer any guidance with respect to the situation in question. A complete belief set that did this would thus be an inconsistent one, given the logic of moral predicates. Gibbard's reading of the Assumption formulates a further consistency requirement for a complete belief set[5].

As I pointed out above, one of Gibbard's main aims is to show that his version of non-cognitivism ends up mimicking many of the claims available to a cognitivist; central to his strategy aimed at showing this is a claim he calls the "Claim of Factual Constitution". His argument in support of this claim revolves around the thesis that all complete belief sets that are consistent must have a certain form: any such set "must amount to permitting all and only acts with some factual property *P*", so that for any such set "there is a prosaically factual property *P* such that [the set] can be viewed as taking the following form: for any situation […], [it] permits all acts with property *P*, and rules out all other acts". This thesis about the form of a complete belief set that is consistent captures Gibbard's reading of the Assumption: let us take a close look at it[6].

---

[5] In assuming that moral beliefs are subject to such general logical rules as the law of non-contradiction, I am granting (as I said I would: see ch. 3, n. 8) that such beliefs can stand in logical relations to one another even on non-cognitivism. For the requirement not to rule out all the alternatives available in a given situation, see Gibbard [2003], p. 56. The notion of consistency, and correspondingly of logic, Gibbard works with seems to come out as follows: moral language has a specific function, which is reflected in the logic of moral predicates, i.e. it gives rise to specific rules for competent use of moral predicates -in addition to those derived from entirely general logical rules-, which exclude certain sets of beliefs as inconsistent (as we saw at the beginning of chapter 3, an analogous idea is present in Hare).

[6] For the relevant discussion, see Gibbard [2003], pp. 94-98, and more generally ch. 5; the quotes in the main text come from p. 96. Gibbard formulates the thesis I quote not, as I do, as one about a complete belief set that is consistent (a hyperstate, in his terminology), but as one about the moral component of such a belief set (a hyperplan, in his terminology). However, if a speaker's plan that, say, permits / rules out individual actions (for instance, one that permits action x) is to satisfy the thesis in question, appropriate non-moral beliefs the speaker holds about the relevant actions (for instance, the belief that action x has property *P*) will have to be taken into account: the thesis in question cannot be true, strictly speaking, of a hyperplan, but only of a hyperplan plus a set of non-moral beliefs. True, framing it as a thesis about a complete belief set is going to include non-moral beliefs that are, strictly speaking, superfluous (say, because they are not about any permitted / ruled out action); at the same time, it does little harm and keeps things neat and simple, which is why I choose it.

If consistent, a complete belief set amounts, says Gibbard, to permitting all and only actions with property *P*, for some factual property *P*. In order to unpack this claim, let us for the time being bracket two aspects of it: I'll get back to them in due course. For any complete belief set that is consistent, says Gibbard, there is a property of which two things are true: the set amounts to (1) permitting all actions with that property and (2) permitting only actions with that property. The first aspect of Gibbard's claim I want to bracket is the second of these two things: for the time being, it is easier to focus only on the claim that the set amounts to permitting all actions with the property in question. Moreover, says Gibbard, a complete belief set that is consistent amounts to permitting all actions with one single property. This is the second aspect of Gibbard's claim I want to bracket: for the time being, it is easier to drop it and take the claim to be that, for any complete belief set that is consistent, there is a non-empty set of (one or more) properties such that the belief set amounts to permitting all actions with one of the properties in question.

These properties are, says Gibbard, factual, or prosaically factual, properties. What he means by this is that they are non-moral properties, which should of course come as no surprise, given that the claim we are looking at captures Gibbard's way of fleshing out the Assumption that actions, people, institutions, etc. owe their moral features to their non-moral properties[7]. Since my interest in this claim is as a tool that non-cognitivist opponents of moral explanations can use in order to develop an accommodationist proposal, moreover, the considerations in 2.2 apply here as well, so I'll take it to be part of the claim that these non-moral properties are ones for which we either already have, or can at least introduce into our language, a non-moral predicate: if a complete belief set is consistent, there is a non-empty set of non-moral properties

---

[7] See Gibbard [2003], p. 88.

that can be denoted by some non-moral predicate such that the belief set amounts to permitting all actions with one of the properties in question[8].

How should we take the latter part of this claim, i.e. that the belief set *amounts* to permitting all actions with one of the properties in question? Clearly, it cannot mean that the only categorical moral beliefs included in the belief set are universal moral beliefs to the effect that all actions with one of those properties are permissible: in general, the set will include, in addition to these, many other categorical moral beliefs as well. The idea here is rather that the belief set regards an action as permissible only if it counts it as instantiating one of the properties in question, so that any categorical moral belief the set includes, in addition to the universal moral beliefs just mentioned, is one that can be subsumed, *via* appropriate non-moral beliefs in the set, under one of those universal moral beliefs. Given the non-moral beliefs included in the set, those universal moral beliefs are enough to infer any other categorical moral belief in the set.

Consider now such universal moral beliefs: the beliefs to the effect that all actions with one of the properties in question are permissible. As I pointed out above, one can think of a complete belief set that is consistent as capturing the views of an idealized user of moral language who has become fully decided: thinking of it this way, the universal moral beliefs in question can be thought of as capturing the general principles or standards held by the speaker whose views are captured by the belief set. Gibbard takes such universal moral beliefs to cover not just actual cases but also possible counterfactual ones (compare with the considerations mentioned in 2.2): in my terminology, he takes them to be expressible by a strong universal moral sentence. What exactly is the degree of strength at play here? Gibbard's own view is that

---

[8] So far as I can see, Gibbard does not explicitly address the question whether the non-moral properties at issue are ones that can be denoted by some non-moral predicate, though at times what he writes seems to assume they are. However, he is explicit (see Gibbard [2003], p. 104) that they are properties we can have concepts for, concepts that are non-moral (or, as he usually calls them, "naturalistic" or "descriptive").

such universal moral beliefs are necessary. However, since a non-cognitivist can endorse Gibbard's reading of the Assumption and use it to develop an accommodationist proposal without taking a stand on this issue, I'll drop this view of Gibbard's and leave the claim we are looking at neutral with respect to this question, using (as I did in chapters 2 and 3) the symbol ' ■ ' for the relevant degree of strength, whatever exactly it turns out to be[9].

Let S be a complete belief set and S_{cm} be the set of categorical moral beliefs included in S. Gibbard's reading of the Assumption can then be framed in terms of the following thesis:

(G)   If S is consistent then, provided S_{cm} is not empty, there is a non-empty subset of S_{cm} –let it be the <u>axiomatic set</u>- such that

    i.G) all the beliefs in S_{cm} that are not in the subset –let their set be the <u>generated set</u>- follow from beliefs in the subset (plus appropriate non-moral beliefs in S) by subsumption;

    ii.G) all the beliefs in the subset can be expressed by an appropriately strong universal moral sentence using a non-moral predicate denoting a non-moral property.

According to thesis (G), if S is consistent, the categorical moral beliefs included in it can be arranged in a neat, orderly way, with some of them (making up the axiomatic set) taken as

---

[9] For Gibbard's own view, see Gibbard [2003], p. 95. As I pointed out in chapter 3 while presenting Hare's views (see ch. 3, n. 4), there might be questions to address for a non-cognitivist who wants to provide even a rough-and-ready characterization of the relation 'being stronger than' along the lines of the one I offered in chapter 2 (which appealed to the notion of truth), but I'll assume they do not raise any serious problem for the non-cognitivist. As I say in the main text, Gibbard takes the universal moral beliefs in question to be expressible by a strong universal moral sentence; he also claims that the non-moral concepts of the non-moral properties involved must meet certain constraints (see Gibbard [2003], pp. 102-108). This further claim too is one that a non-cognitivist can remain neutral about while still being able to rely on Gibbard's reading of the Assumption in order to develop an accommodationist proposal, so in presenting Gibbard's views I'll drop it.

starting points and all the others (making up the generated set) following from the former (plus non-moral beliefs in S) by subsumption.

What do I mean by 'following by subsumption'? Thesis (G) is formulated in terms of the logical relation 'following from': what sub-thesis i.G) says is that the beliefs in the axiomatic set are (together with non-moral beliefs in S) enough to infer all the other beliefs in S_{cm} by means of logic. Indeed, i.G) says something more specific than this: according to it, the beliefs in the generated set follow from beliefs in the axiomatic set (plus appropriate non-moral beliefs in S) *by subsumption*. Intuitively, what I mean by this is that the beliefs in the generated set can be subsumed under some universal moral belief (or beliefs) involving a non-moral property (or non-moral properties). More formally: I'll say that an individual / universal moral belief to the effect that x is M / all P's are M follows from beliefs in a given set <u>by subsumption</u> when the set contains a moral belief to the effect that all N's are M, plus a non-moral belief to the effect that x is N / all P's are N, for some non-moral property N-ness, or (in the case of a universal moral belief) when the set contains a series of moral beliefs to the effect that all N_1's are M, all N_2's are M, ..., where N_1-ness, N_2-ness, ... are different non-moral properties, plus a non-moral belief to the effect that each P is either N_1 or N_2 or ... [10]. Sub-thesis i.G) says not just that, for any belief in the generated set, the beliefs in the axiomatic set are (together with non-moral beliefs in S) enough to infer that belief by means of logic; it says more specifically that, for any belief in the generated set, there are in the axiomatic set (and among the non-moral beliefs in S) beliefs of a particular form -one that fits the definition just given of 'following by subsumption'-, which are enough to infer that belief by means of logic[11].

---

[10] Alternatively: plus a non-moral belief to the effect that all P's have the disjunctive property of being $N_1 \vee N_2 \vee$ ... .

[11] It is worth pointing out explicitly that i.G) does not rule out the possibility that a belief in the generated set follow by subsumption from different (series of) beliefs in the axiomatic set (plus appropriate non-moral beliefs in S). The definition of 'following by subsumption' matches in obvious ways the definitions of 'being ultimately made true in

For any belief in the generated set, says i.G), there are beliefs in the axiomatic set (and among the non-moral beliefs in S) from which the belief in question follows by subsumption. This entails that the relevant beliefs in the axiomatic set are universal moral beliefs involving non-moral properties. Sub-thesis ii.G) adds that such universal moral beliefs are strong (with the appropriate degree of strength) and involve non-moral properties that can be denoted by non-moral predicates. Indeed, ii.G) is a thesis about *all* the beliefs in the axiomatic set, and thus rules out, for example, the possibility (not ruled out by i.G)) that the axiomatic set include individual moral beliefs. As a result, on thesis (G) if S is consistent, it does not count anything as M (for a given moral predicate 'M') unless it counts it as instantiating one of the properties picked out by the universal moral beliefs in the axiomatic set. I noted above that one way to think of a complete belief set that is consistent is as capturing the views of a user of moral language who has become fully decided, with the beliefs in the axiomatic set capturing the general principles or

the relevant way' and 'being ultimately inferred *via* subsumption' given in the previous chapters. In comparing the latter two relations, I pointed out (see ch. 2, n. 45) that they differ in the following respect: the definition of 'being ultimately inferred *via* subsumption' is less restrictive since this is an epistemological relation, so the way it is defined takes uncertainty into account, allowing for some degree of non-moral uncertainty; the relation 'being ultimately made true in the relevant way', on the other hand, is a metaphysical relation, and when it comes to how things are made true in reality uncertainty is not an issue. Under this respect, the relation 'following by subsumption' is defined so as to resemble the metaphysical, not the epistemological, relation, since thesis (G) formulates a consistency requirement for a belief set that is complete, and when dealing with such a belief set uncertainty is, once again, not an issue. (Here is the sort of example I used to illustrate the point in chapter 2. Suppose one has two direct moral beliefs, i.e. that ■ all N_1's are M and that ■ all N_2's are M, and consider the following two cases: one comes to believe that x is N_1 and concludes it is M; one comes to believe that x has one or the other of the two non-moral properties in question and concludes it is M. On the definition of 'being ultimately inferred *via* subsumption', the individual moral belief that x is M counts as ultimately inferred *via* subsumption in both cases: in the first, from the moral belief that ■ all N_1's are M, plus the non-moral belief that x is N_1; in the second, from the two moral beliefs that ■ all N_1's are M and ■ all N_2's are M, plus the non-moral belief that x is N_1 or x is N_2. Suppose, on the other hand, that there are two brute moral truths, i.e. that ■ all N_1's are M and that ■ all N_2's are M, and that x is N_1 but not N_2. On the definition of 'being ultimately made true in the relevant way', the individual moral truth that x is M counts as ultimately made true in the relevant way by the moral truth that ■ all N_1's are M, plus the non-moral truth that x is N_1, but it does not also count as ultimately made true in the relevant way by the two moral truths that ■ all N_1's are M and ■ all N_2's are M, plus the non-moral truth that x is N_1 or x is N_2. Using now the same sort of example to illustrate the point with respect to the relation 'following by subsumption', consider the following two belief sets: the set made up just of the moral beliefs that ■ all N_1's are M and ■ all N_2's are M and the non-moral belief that x is N_1; and the set made up just of the moral beliefs that ■ all N_1's are M and ■ all N_2's are M and the non-moral belief that x is N_1 or x is N_2. On the definition of 'following by subsumption', the individual moral belief that x is M counts as following by subsumption from the beliefs in the first set, but not from those in the second.)

standards held by such a speaker. Thinking this way, we can say that on thesis (G) a user of moral language who has become fully decided does not judge anything to be M unless she holds a general principle she takes to apply to it and yield this verdict about it.

This allows us to see how (G) vindicates the way I sketched out at the outset of articulating the idea that use of moral language commits one to general moral principles. None of us is fully decided, of course, so the fact that, say, you think x is M but do not hold a general moral principle you take to apply to x and yield this verdict about it does not, on thesis (G), make you an inconsistent user of moral language. However, given thesis (G), any way in which you might develop your views and become fully decided in a consistent way is one in which you do end up holding a general moral principle you take to apply to x and yield this verdict about it (provided, of course, that you do not change your mind about x and preserve your belief that x is M). To put it as I did at the beginning of this section, your belief that x is M calls for such a general moral principle: you are committed to holding one, if you become fully decided.

It is instructive to compare Gibbard's reading of the Assumption, captured by thesis (G), with Hare's reading, formulated by theses i.H)-iv.H) from the previous chapter. Theses i.H)-iv.H) are framed in terms of the epistemological relation 'being inferred from': they formulate an epistemological view about what users of moral language infer their categorical moral beliefs from (insofar as they avail themselves only of inferences of an epistemologically good sort). Thesis (G), on the other hand, is framed in terms of the logical relation 'following from': it is a logical claim about how the categorical moral beliefs included in a complete belief set that is consistent are logically related to one another, one that is silent on the question which categorical moral beliefs are inferred from which others.

The contrast emerges with regard to actual users of moral language: since thesis (G) is about complete belief sets that are consistent, and since no such speaker's views are captured by a complete belief set that is consistent, thesis (G) does not apply to actual speakers, so in particular it makes no claim about what such speakers infer their categorical moral beliefs from. But, it is worth pointing out, the contrast also emerges for the (highly idealized) case of a user of moral language who has become fully decided. Thesis (G) does make a claim about the views of such a speaker. The claim, however, is the logical claim that the categorical moral beliefs of such a speaker make up two sets: the axiomatic set, whose members have the features mentioned by ii.G); and the generated set, whose members *follow* from beliefs in the axiomatic set (plus appropriate non-moral beliefs held by the speaker) by subsumption. Thesis (G) does not make the epistemological claim that the beliefs in the generated set are *inferred* by such a speaker from beliefs in the axiomatic set (plus appropriate non-moral beliefs she has).

So, for example, whereas (as we saw in chapter 3) theses i.H)-iv.H) formulate a foundationalist epistemological view, thesis (G) is compatible with the possibility that none of a fully decided speaker's categorical moral beliefs be held without inference. To give another example: whereas (as we also saw in chapter 3) on i.H)-iv.H) evidential support ultimately flows only from the more to the less general, (G) is compatible with the possibility that a fully decided speaker hold some beliefs in the generated set, say some individual moral beliefs, without inference and infer from them some of the universal moral beliefs in the axiomatic set. Quite generally, thesis (G) makes no claim about what even a fully decided user of moral language infers her categorical moral beliefs from.

True, if q logically follows from p, then one *can* infer q from p. And, given the way in which the two relations 'following by subsumption' (mentioned by thesis (G)) and 'being

ultimately inferred *via* subsumption' (mentioned by theses i.H)-iv.H)) are defined, the following is also true: if a categorical moral belief follows from other categorical moral beliefs (plus some non-moral beliefs) by subsumption, then one *can* ultimately infer it from those categorical moral beliefs (plus the non-moral beliefs in question) *via* subsumption, provided one holds the latter categorical moral beliefs without inference (and holds the relevant non-moral beliefs). So thesis (G) does entail this much about the categorical moral beliefs of a fully decided speaker: *if* she held the beliefs in the axiomatic set without inference, she *could* ultimately infer all the beliefs in the generated set from them (plus her non-moral beliefs) *via* subsumption. Thesis (G) thus guarantees that we can rationally reconstruct, so to speak, the categorical moral beliefs of a fully decided speaker *as if* they conformed with theses i.H)-iv.H). This claim, however, is different from the claim that the categorical moral beliefs of such a speaker conform with theses i.H)-iv.H): though thesis (G) entails the former claim, it stops short of committing its friends to any epistemological view about what even a fully decided user of moral language infers her categorical moral beliefs from.

Here then is one way of thinking of Gibbard's reading of the Assumption. The most natural way of fleshing out the Assumption, I pointed out in chapter 2, is by means of theses i)-iv), which put forward a straightforwardly metaphysical picture describing how categorical moral truths are arranged in terms of the metaphysical relation 'being made true by'. The non-cognitivist's strategy to provide her reading of the Assumption, we began to see in chapter 3, can be described as follows: preserve the shape of that metaphysical picture, replace the materials it is built out of. This is what Hare's reading does: it preserves the idea that the moral domain has a neat structure, but turns the metaphysical picture into an epistemological one, which describes how categorical moral beliefs are arranged in terms of the epistemological relation 'being

173

inferred from'. One way of looking at Gibbard's reading of the Assumption is as suggesting there is a better way of implementing the non-cognitivist strategy: the non-cognitivist should preserve the idea that the moral domain has a neat structure by invoking a different structure than the epistemological one invoked by Hare, one that operates at a deeper level.

Sure, the structure the non-cognitivist should appeal to involves categorical moral beliefs instead of categorical moral truths (since on non-cognitivism there can't be any such truths). It is not, however, a matter of how categorical moral beliefs are inferred from one another by users of moral language. Rather, it is a matter of how such beliefs are logically related to one another within a complete belief set that is consistent. Given this logical structure, it is no doubt possible to provide a rational reconstruction of the categorical moral beliefs of a fully decided speaker as if they exhibited the epistemological structure invoked by Hare. One point I have already made, though, is that this is not the same as saying these beliefs exhibit this structure. A second point, which I want to stress here, is that this epistemological structure is in any case merely a byproduct, resulting from the underlying logical relations that occur among the categorical moral beliefs of such a speaker. It is these logical relations that, on Gibbard's reading of the Assumption, give the moral domain its neat structure.

And, I should finally emphasize, all this is good news for the non-cognitivist, given the arguments I offered in the previous chapter to reject the epistemological view formulated by theses i.H)-iv.H). Thesis (G) provides the non-cognitivist with an alternative reading of the Assumption that does not commit her to that epistemological view.

I dwelt at some length on the difference between (G) and i.H)-iv.H) because they formulate the two readings of the Assumption put forward by non-cognitivists, so it is important to point out how they differ from one another. The difference between (G) and the reading of the

174

Assumption available to cognitivists, formulated by i)-iv), is more straightforward. Here I'll only remind the reader of a point I made when introducing i)-iv). As I remarked in section 2.2, the metaphysical relation 'being made true by' mentioned by i)-iv) is a different relation from the logical relation 'following from' mentioned by (G). More accurately (and, as I noted in 2.2, more interestingly, since of course q can follow from p even if p is not true or q is not true, whereas q can only be made true by p if both p and q are true), q's being made true by p is not the same as p's being true and q's following from p. This was the point of my example about Dave, one of the students in my same year in the philosophy graduate program. Consider the three sentences 'All the people in my year are good people', 'Dave is in my year', and 'Dave is a good person'. The belief expressed by the last sentence follows from the beliefs expressed by the other two, but (even on the assumption that the three sentences are all true) the truth expressed by the last sentence is not made true by the truths expressed by the other two: whatever is good about Dave, it is not that he is in my year. The logical relation 'following from' mentioned by thesis (G) is, as it is sometimes put, lighter than the straightforwardly metaphysical relation 'being made true by' mentioned by theses i)-iv).

Two final comments about thesis (G). Let S be a complete belief set and S_{cm} be the set of categorical moral beliefs included in S. Thesis (G) formulates a consistency requirement for S, and the requirement is conditional: *if* S_{cm} is not empty, there is a non-empty subset of S_{cm} with the features mentioned by i.G) and ii.G). This means that a complete belief set including no categorical moral beliefs at all satisfies the requirement in question: thesis (G) does not rule out as inconsistent a complete belief set that, for any categorical moral belief, contains the negation of that belief[12]. (G) thus makes room for more than Gibbard's formulations strictly speaking do.

---

[12] I should remind the reader that the negation of a categorical moral sentence is not a categorical moral sentence (see the definition of 'categorical moral sentence' in 2.2). Even so, it might be objected, isn't what I assume in the

My reconstruction of Gibbard's reading of the Assumption started from his thesis that all complete belief sets that are consistent must have a certain form: for any such belief set, there must be a non-empty set of (one or more) properties such that the belief set amounts to permitting all actions with one of the properties in question. This thesis does rule out as inconsistent a complete belief set including no categorical moral beliefs, since in order to have that form, a belief set must include some categorical moral beliefs: at the very least, universal moral beliefs to the effect that all actions with one of the properties in question are permissible. Gibbard's formulations effectively add, to the conditional requirement formulated by (G), the further requirement that S include categorical moral beliefs. Why did I drop this further requirement in my reconstruction of Gibbard's reading of the Assumption?

---

text false? I assume that a complete belief set that, for any categorical moral belief, contains the negation of that belief is one that includes no categorical moral beliefs at all. But suppose 'M' is a moral predicate and consider a complete belief set containing the belief that x is not M: no doubt 'x is not M' is not a categorical moral sentence, but can't the belief that x is not M be expressed also by a sentence that fits my definition of 'categorical moral sentence'? Take a moral predicate, 'M*', that is synonymous with 'not M' (maybe we already have such a predicate; if not, introduce it into our language): the belief that x is not M can be expressed by the categorical moral sentence 'x is M*', so it is a categorical moral belief. Doesn't this show that what I assume in the text is false? No, at least not in any problematic way. Thesis (G) formulates one reading of the Assumption and, as I pointed out in chapter 2 when introducing the first reading of the Assumption I looked at (see ch. 2, n. 7), there is a restriction on the scope of the Assumption, in any of its readings, one that comes out if we consider the view, held by some writers in meta-ethics, that some moral properties (predicates) are reducible to others. If we consider such a view, the Assumption is meant to commit its cognitivist friends only with respect to the categorical moral truths that ascribe the non-reducible moral property (or properties) to which the other moral properties can be reduced; and it is meant to commit its non-cognitivist friends only with respect to the categorical moral beliefs expressible by categorical moral sentences that use the non-reducible moral predicate (or predicates) to which the other moral predicates can be reduced (N-R categorical moral beliefs, for short). In chapter 2 I said that in the present work I would bracket the possibility that some moral properties (predicates) be reducible to others, thereby dropping such a restriction and framing the Assumption as a claim about *all* categorical moral truths (beliefs). Focusing now on thesis (G), if we bracket the possibility that some moral predicates be reducible to others, what I assume in the text is true. If, on the other hand, we take that possibility into account (as we have to, if we are to be able to raise the objection I am considering), then we have to re-frame thesis (G) so as to restrict its scope. A complete belief set has to be re-defined as a belief set with the following feature: for any N-R categorical moral belief, the set contains either that belief or its negation, and for any non-moral belief, the set contains either that belief or its negation. And thesis (G) has to be re-framed as imposing the following consistency requirement on a belief set S that is complete in this sense: if the set S_{cm} of N-R categorical moral beliefs included in S is not empty, there is a non-empty subset of S_{cm} with the features mentioned by i.G) and ii.G). If the scope of thesis (G) is restricted this way, what I assume in the text remains true, once its scope is restricted in a parallel way: a complete belief set (in the sense just defined) that, for any N-R categorical moral belief, contains the negation of that belief is one that includes no N-R categorical moral beliefs at all, and so satisfies the requirement formulated by thesis (G) (as just framed). The objection I am considering shows that such a set may still contain categorical moral beliefs, but it fails to show it may contain N-R categorical moral beliefs. As I said, in the present work I am, for ease of exposition, bracketing the possibility that some moral predicates be reducible to others, so I'll continue to use the unrestricted formulations in the text.

Because this further requirement is not part of the intuitive idea in whose terms non-cognitivists propose to flesh out the Assumption, i.e. the idea that use of moral language commits one to general moral principles. As we saw above, this is the idea that, in saying 'x is M' (where 'M' is a moral predicate), one incurs a commitment of a certain sort: if you hold a categorical moral belief, your belief set is constrained in a certain way. Think again in terms of the rules of a game: the idea is that the logic of moral predicates is such that, if you make the move of saying 'x is M', you must follow certain directions. It is not part of that idea that you *must* make that move, or that you *must* play the game of holding categorical moral beliefs. The idea is that there are rules to follow if one holds categorical moral beliefs; in itself, it is neutral with regard to whether one does hold any such beliefs.

Indeed, a conditional claim is what is needed in order to mimic, as non-cognitivist readings of the Assumption are meant to, the reading of the Assumption available to cognitivists, formulated by theses i)-iv). As I pointed out in 2.2, i)-iv) are neutral with regard to whether there are any categorical moral truths: the claim they make is the conditional claim that, if there are any such truths, they are arranged in a certain way. Analogously, i.H)-iv.H), which formulate Hare's reading of the Assumption, are neutral with regard to whether we hold any categorical moral beliefs: the claim they make is the conditional claim that, if we hold any such beliefs, they are arranged in a certain way. If Gibbard's reading is to mimic i)-iv), it too has to remain neutral with regard to whether a complete belief set includes any categorical moral beliefs, making only the conditional claim that, if it does, such beliefs are arranged in a certain way. The Assumption, in other words, makes a point about the nature, not the existence, of morality: it describes what morality looks like –what categorical moral truths look like, or what categorical moral beliefs

look like-, whether or not there is such a thing. The best way to formulate Gibbard's reading of the Assumption is thus by means of thesis (G).

If I am right, and the requirement that a complete belief set include categorical moral beliefs should form no part of Gibbard's reading of the Assumption, where does that requirement come from? Why do Gibbard's formulations take it for granted? Gibbard does not directly address the question why a complete belief set including no categorical moral beliefs is inconsistent, but here is a conjecture about why he should assume it is. I said above that a set of beliefs can, according to Gibbard, be inconsistent either because it violates entirely general logical rules or because it violates the specific rules governing the use of moral predicates. Before examining Gibbard's reading of the Assumption, which formulates a consistency requirement of the latter sort, I mentioned an example of another requirement that Gibbard takes to derive from the logic of moral predicates. This is the requirement that a set of beliefs not rule out as impermissible all the alternatives available in a given situation, the idea being that one of the specific functions of moral language is to provide guidance with respect to choices and actions and a set of beliefs ruling out all alternatives would fail to offer any. Given such a requirement, a complete belief set including no categorical moral beliefs does count as inconsistent, since for any categorical moral belief to the effect that an alternative is permissible, such a belief set contains the negation of that belief, thereby ruling out all the alternatives available in any given situation. If, however, this is why a complete belief set including no categorical moral beliefs is inconsistent, the requirement that a complete belief set include categorical moral beliefs is really a separate requirement from that formulated by thesis (G), one with a different and independent motivation: on the one hand, the idea that use of moral language commits one to general moral principles; on the other, the idea that one of the specific functions

of moral language is to provide guidance with respect to choices and actions. In our discussion of Gibbard's reading of the Assumption, we are then well advised to bracket it, despite the fact that his formulations run the two requirements together.

After pointing out one aspect of Gibbard's formulations that we should bracket, my second comment concerns two further aspects of them, which I said earlier I would bracket only temporarily. As we saw above, Gibbard frames his thesis that all complete belief sets that are consistent must have a certain form as follows: for any such belief set, there must be *one single* property such that the belief set amounts to permitting all *and only* actions with that property. I said I would temporarily bracket the two aspects I have just highlighted in italics and get back to them in due course: let me deal with them now in reverse order. Framing Gibbard's reading of the Assumption in terms of thesis (G) already re-introduces the second aspect. Let S be a complete belief set and suppose it includes categorical moral beliefs. Given thesis (G), if S is consistent the categorical moral beliefs included in it make up two sets: the axiomatic and the generated set. Consider now the properties picked out by the universal moral beliefs in the axiomatic set. As I pointed out above, on thesis (G) if S is consistent, it does not count anything as permissible unless it counts it as instantiating one of those properties. Said otherwise: on thesis (G) if S is consistent, what is true of it is not just that it permits all actions with one of the properties in question, but also that it permits only actions with one of those properties, i.e. S amounts to permitting all *and only* actions with one of the properties picked out by the universal moral beliefs in the axiomatic set. As for the first aspect mentioned above, there is actually less to it than meets the eye. Gibbard argues that, for any complete belief that is consistent, there must be a single property with the relevant feature by saying that such a property can always be

constructed disjunctively[13]. Thesis (G) already guarantees the existence of such a property: on (G) if S is consistent, it amounts to permitting all and only actions with the *single* disjunctive property built out of the properties picked out by the universal moral beliefs in the axiomatic set. This aspect of Gibbard's formulations thus adds nothing new to thesis (G).

2. **The 'Go by the hyperstates' strategy**

Gibbard relies on thesis (G) in order to develop an accommodationist proposal about moral explanatory claims. Just as in the case of the accommodationist proposals we looked at in chapters 2 and 3, Gibbard's proposal can be seen as the application to moral explanatory claims of a more general recipe that can be applied to other moral sentences as well. It is helpful, in studying that proposal, to look at the general recipe.

Consider an individual moral sentence, i.e. a sentence of the form 'x is M', where 'M' is a moral predicate and 'x' is a singular term. Gibbard proposes to account for the content of such a sentence in terms of the complete belief sets that are consistent and include the corresponding belief. A sentence like 'x is M' divides all the complete belief sets that are consistent into two sets: those that include the belief that x is M and those that include its negation. Gibbard's suggestion is to account for the content of 'x is M' in terms of the former set. If we think of a complete belief set that is consistent as capturing the views of an idealized user of moral language who has become fully decided, Gibbard's suggestion is that the content of 'x is M' is a matter of the ways in which one can become fully decided and hold that x is M. Indeed, Gibbard's view is that moral sentences quite generally can be accounted for along these lines: the content of a moral sentence is a matter of the various ways in which one can become fully decided and accept the corresponding belief. Using his term for complete belief sets that are

---

[13] See for instance Gibbard [2003], p. 96 or p. 99.

consistent, Gibbard's strategy in accounting for the content of moral sentences can thus be summarized by the slogan 'Go by the hyperstates'[14].

Is there a canonical way of representing the complete belief sets that are consistent and include the belief that x is M, hence of representing the content of 'x is M'? Yes, given thesis (G). On (G) any such belief set will include, in the first place, a subset of categorical moral beliefs making up the axiomatic set. (G) is neutral as to the number of categorical moral beliefs the axiomatic set can contain; no matter what their number is, these beliefs can all be expressed by an appropriately strong sentence of the form 'All N's are M', for some non-moral predicate 'N' denoting a non-moral property. Among all the complete belief sets that are consistent and include the belief that x is M, consider then those whose axiomatic set contains only one categorical moral belief: say, the belief that ■ all N_1's are M. On thesis (G) any such belief set will include, in the second place, an appropriate non-moral belief such that the belief that x is M follows from it, together with the moral belief that ■ all N_1's are M, by subsumption: that is, it will include the non-moral belief that x is N_1. What about the complete belief sets, among those that are consistent and include the belief that x is M, whose axiomatic set contains two categorical moral beliefs, say the beliefs that ■ all N_1's are M and ■ all N_2's are M? For the reason just given, on thesis (G) any such belief set will include either the non-moral belief that x is N_1 or the non-moral belief that x is N_2: either way, since we are dealing with belief sets that are complete and consistent, any such belief set will include the disjunctive belief that x is N_1 or

---

[14] See Gibbard [2003], ch. 3. Gibbard's view is actually that this is the way to account for the content of *any* sentence, moral or non-moral, and that seeing this is the key to solving the Frege-Geach problem for non-cognitivism. The content of a sentence is a matter of the ways in which one can become fully decided and accept it: keeping in mind that, as I pointed out above, talk of what he calls "hyperstates" and talk of what he calls "fact-plan worlds" are, according to Gibbard, interchangeable, this idea is (as Gibbard himself points out: see Gibbard [2003], pp. 57-58) a version of the popular idea that the content of a sentence can be accounted for in terms of the set of possible worlds in which the sentence holds. With his strategy of 'going by the hyperstates', Gibbard aims at providing a version of this popular idea that makes it applicable also to a domain, such as the moral domain on non-cognitivism, in which there can be no truths.

x is N_2. Quite generally, on (G) any complete belief set that is consistent and includes the belief that x is M will include two things: (1) the categorical moral beliefs making up the axiomatic set and (2) a non-moral belief like those just mentioned, so it can be represented by the conjunction of these two elements. The set of all such belief sets, hence the content of 'x is M', we can then represent by means of the big disjunction of all these conjunctions. One way of implementing the strategy of 'going by the hyperstates' thus yields the following proposal:

(c.G') The content of 'x is M' is that

(x is N_1 and   ■ all N_1's are M) or

(x is N_2 and   ■ all N_2's are M) or … or

((x is N_1 or x is N_2)   and   ■ all N_1's are M and ■ all N_2's are M) or

((x is N_3 or x is N_4)   and   ■ all N_3's are M and ■ all N_4's are M) or … or

… .

Another way of implementing the strategy of 'going by the hyperstates' relies on a slightly different way of providing a canonical representation of the complete belief sets that are consistent and include the belief that x is M. Consider again, among such belief sets, those whose axiomatic set contains two categorical moral beliefs, say the beliefs that ■ all N_1's are M and ■ all N_2's are M. I said above that, given thesis (G), any such belief set can be represented by conjoining two elements: the beliefs in the axiomatic set and the disjunctive belief that x is N_1 or x is N_2. We can get the same result using, instead of a disjunctive belief, one that ascribes a

disjunctive property: the property of being $N_1 \lor N_2$[15]; analogously for all the other belief sets whose axiomatic set contains more than one categorical moral belief. On this way of implementing the strategy of 'going by the hyperstates', what we get is the following:

(c.G'') The content of 'x is M' is that

(x is N_1 and   ■ all N_1's are M) or

(x is N_2 and   ■ all N_2's are M) or … or

(x has the disjunctive property of being $N_1 \lor N_2$ and   ■ all N_1's are M and ■ all N_2's are M) or

(x has the disjunctive property of being $N_3 \lor N_4$ and   ■ all N_3's are M and ■ all N_4's are M) or … or

… [16].

---

[15] Once again, since we are dealing with belief sets that are complete and consistent, any of the belief sets in question will include the belief that x has the disjunctive property of being $N_1 \lor N_2$.

[16] I use the letter 'c' in my names for Gibbard's proposals because, as I pointed out at the beginning of this chapter, they aim at mimicking the cognitivist proposals of type C described in chapter 2, i.e. those that rely on general substitution. Later in this chapter I'll discuss how exactly Gibbard's proposals try to achieve this aim, and begin to examine whether they succeed. On both (c.G') and (c.G''), how to deal with the negation of an individual moral sentence is pretty straightforward: just take the negation of the big disjunction capturing the content of the individual moral sentence. Gibbard introduces his strategy of 'going by the hyperstates' in chapter 3 of his book, when dealing with moral sentences in general, including categorical moral sentences; he then articulates that strategy more clearly, making further details explicit, in chapter 10, where he deals with moral explanatory claims. Since the strategy is meant to be entirely general, in presenting Gibbard's views I am building the details he makes explicit only in his treatment of moral explanatory claims into his account of categorical moral sentences as well. Gibbard himself adopts the second of the two ways of implementing the 'Go by the hyperstates' strategy described in the text, since (as we saw at the end of the previous section) he favors the second way of representing complete belief sets that are consistent and whose axiomatic set contains more than one categorical moral belief. I mention also the first way of implementing that strategy because it is useful to distinguish and lay out all the options available to a non-cognitivist who wants to take inspiration from Gibbard; a comparison with the cognitivist proposals relying on general substitution, which Gibbard's account attempts to mimic, also suggests two ways of implementing his strategy. In what follows, I'll feel free to ascribe both (c.G') and (c.G'') to Gibbard. Finally, look at the first disjunct mentioned by (c.G') and (c.G''): I framed it as 'x is N_1 and ■ all N_1's are M'. As I pointed out in the text, this disjunct is meant to represent those, among the complete belief sets that are consistent and include the belief that x is M, whose axiomatic set contains only one categorical moral belief, i.e. the belief that ■ all N_1's are M. But if it is to do that, shouldn't the second conjunct be stronger than ' ■ all N_1's are M'? Shouldn't it reflect the fact that the axiomatic set includes *only* one categorical moral belief? Yes. And the same holds, of course, for all the other disjuncts making up

One noteworthy feature of (c.G') and (c.G'') is that the moral content of 'x is M' is captured entirely by means of sentences of the form 'All N's are M' (where 'N' is a non-moral predicate denoting a non-moral property). On both proposals, the content of 'x is M' is captured by a big disjunction, each of whose disjuncts conjoins two elements: of these two conjuncts, only the second has moral content, and it is made up of sentences of that form; the content of the first conjunct, on the other hand, is entirely non-moral. If we look only at the first conjunct of each conjunction, what (c.G') and (c.G'') do is to replace an individual moral sentence's apparent reference to a moral property with a reference to non-moral properties. This is a feature of (c.G') and (c.G'') that should strike you as appealing, if you are a non-cognitivist who wants to accommodate our practice with regard to moral explanatory claims, since you want to do the same with respect to such claims. On the face of it, a claim of the form 'M-ness causes E' (where 'M' is a moral predicate) ascribes causal efficacy to the property of being M: replacing such a claim's apparent reference to a moral property with a reference to non-moral properties is what

---

the big disjunctions mentioned by (c.G') and (c.G''). This is the same problem that, as I pointed out in chapter 3 (see ch. 3, n. 15 and n. 17), arises with the formulation of Hare's account of categorical moral sentences. As in chapter 3, the solution is to strengthen the conjunct in question by having it be ' ■ all N_1's are M *and* not (■ all N_2's are M) and not (■ all N_3's are M) and …'; analogously for all the other relevant conjuncts. Why, then, if such conjuncts need to be strengthened this way, do I frame them the way I do in the main text? Because, as in chapter 3, I want to make things simpler for the reader and easier for my opponent. Once we strengthen the relevant conjuncts, it is not clear that Gibbard can account for the content of 'x is M' by means of (c.G') or (c.G''), given the worry that non-cognitivism may have trouble dealing with negation. Whether non-cognitivists can answer this worry is controversial: Gibbard claims he can (see Gibbard [2003], ch. 4); as I said in chapter 3, I believe this is a worry non-cognitivists cannot answer. However, as I also said in chapter 3, in the present work I won't press this worry and I'll grant non-cognitivists that they can answer it, since the issues it raises are distinct from those I want to raise here. I'll thus grant Gibbard that he can strengthen the relevant conjuncts and still be able to account for the content of 'x is M' by means of (c.G') or (c.G''). Indeed, to keep things simple, I'll phrase the conjunct in question as ' ■ all N_1's are M' and drop the rest, taking it as understood (I'll remind the reader of it when necessary); analogously for all the other relevant conjuncts. In addition to the problem related to negation, there is another problem for Gibbard here, one that I *will* press in what follows; I won't be in a position to do so, however, until we get to chapter 5. Keeping in mind that I am here bracketing the worry about negation also answers the following question: what can a non-cognitivist mean by a disjunction such as those mentioned by (c.G') and (c.G'')? How can Gibbard account for such disjunctions in a way compatible with his non-cognitivism? Since I am granting Gibbard that he can account for negation, and since nobody doubts that non-cognitivists can account for conjunction, I am thereby granting him disjunction as well.

you need to do in order to accommodate our practice of offering, discussing, and accepting claims like this. It should thus strike you as a plausible strategy, in accounting for the content of such claims, to 'go by the hyperstates': start with the idea that the content of 'M-ness causes E' is a matter of the ways in which one can become fully decided and hold the corresponding belief; given thesis (G), it should strike you as plausible to represent such a content in one of the two ways described above, which will allow you to achieve the result you wanted. This is what Gibbard's proposal about moral explanatory claims does:

(C.G') The content of 'M-ness causes E' is that

(N_1-ness causes E   and   ■ all N_1's are M) or

(N_2-ness causes E   and   ■ all N_2's are M) or … or

(N_1-ness causes E and N_2-ness causes E   and   ■ all N_1's are M and ■ all N_2's are M) or

(N_3-ness causes E and N_4-ness causes E   and   ■ all N_3's are M and ■ all N_4's are M) or … or

…;

(C.G'') The content of 'M-ness causes E' is that

(N_1-ness causes E   and   ■ all N_1's are M) or

(N_2-ness causes E   and   ■ all N_2's are M) or … or

(the disjunctive property of being $N_1 \vee N_2$ causes E   and   ■ all N_1's are M and ■ all N_2's are M) or

(the disjunctive property of being $N_3 \vee N_4$ causes E   and   ■ all N_3's are M and

■ all N_4's are M) or … or

…[17].

As with the proposals about individual moral sentences, on (C.G') and (C.G'') the content of a moral explanatory claim is captured by a big disjunction, each of whose disjuncts conjoins two distinct elements: one with an explanatory, entirely non-moral content, another with a purely moral, non-explanatory content. As a result, on both proposals a moral explanatory claim has both moral and explanatory content, but the former does not mesh in an undesired way with the latter: on both proposals, no ascription of causal efficacy is made to any moral property.

Some interesting features of Gibbard's strategy of 'going by the hyperstates' come out if we compare it with Hare's strategy of 'going by the speaker's standards', which we looked at in chapter 3[18]. In order to assign content to moral sentences, both Hare and Gibbard start with an idealized user of moral language. As we saw in chapter 3, Hare's semantic account is built on

---

[17] See Gibbard [2003], ch. 10, especially pp. 203-207. In his treatment of moral explanatory claims, Gibbard focuses mostly on examples of particular claims of the form 'x's being M causes effect *e*' (e.g. 'The injustice of that regime causes that revolution), so strictly speaking his proposal is formulated as one about claims of this form; in presenting his views, I feel free to translate that proposal into one about general claims of the form 'M-ness causes effects of type E' (e.g. 'Injustice causes revolutions'), which are my topic in the present work. In his own formulation of his accommodationist proposal, Gibbard avails himself of the definite description 'the property that constitutes being the thing to do' or (since I am recasting his claims as claims about moral predicates) 'the property that constitutes being M': I'll get back to this aspect of his proposal later in this section. Once again, Gibbard himself adopts the second of the two ways of accounting for the content of 'M-ness causes E' described in the text. I mention also the first way because it is useful to distinguish and lay out all the options available to a non-cognitivist who wants to take inspiration from Gibbard; a comparison with the cognitivist who relies on general substitution, whom Gibbard's proposal attempts to mimic, also suggests two ways of accounting for the content of moral explanatory claims. In what follows, I'll feel free to ascribe both (C.G') and (C.G'') to Gibbard. (C.G') and (C.G'') both deal with negations in the same straightforward way as (c.G') and (c.G'') do. The same considerations mentioned in n. 16 apply here as well: the second conjunct of the first disjunct mentioned by (C.G') and (C.G''), which I framed as ' ■ all N_1's are M', should read ' ■ all N_1's are M *and* not (■ all N_2's are M) and not (■ all N_3's are M) and …', but in order to make things simpler for the reader and easier for my opponent I'll frame it the way I have just done in the main text and drop the rest, taking it as understood; analogously for all the other relevant conjuncts.

[18] As we saw in chapter 3, Hare applies his 'Go by the speaker's standards' strategy to categorical moral sentences, while its application to moral explanatory claims is due to Blackburn. Though in what follows I'll illustrate them with reference to Hare, my remarks concern the general strategy itself.

theses i.H)-iv.H), which make an epistemological claim about how our categorical moral beliefs look like, not as a matter of fact, but insofar as we avail ourselves only of epistemologically good inferences: whereas actual speakers meet this condition only to a limited extent, Hare starts with a user of moral language who, in the first place, does so to a full extent. Hare's speaker, in the second place, has a fully worked out theory about what things are M, for a given moral predicate 'M': in particular, she has made up her mind about what general moral principles (in Hare's terminology) or direct moral beliefs (in the terminology of i.H)-iv.H)) to accept. As I pointed out in section 3.2, this makes Hare's speaker different from most actual speakers at least under the following two respects: on the one hand, most of us have general moral principles that are often somewhat vague (I believe *something* like ' ■ All N_1's are M', though the non-moral property in question might not be quite N_1-ness, it might be some other property in its vicinity); on the other, even when our general principles are clear-cut, we are often open to the possibility that they not be exhaustive (I believe that ■ all N_1's are M and ■ all N_2's are M, but am uncertain whether that is all or there are other items on the list)[19].

Gibbard also starts with a user of moral language who is idealized both in the way she reasons and in how much she believes, though not quite in the same ways as Hare's speaker is. As we saw in the previous section, Gibbard's semantic account is built on thesis (G), which deals with complete belief sets that are consistent. As pointed out above, one way to think of a belief set with these features is as capturing the views of a user of moral language who, in the first

---

[19] A third respect of dissimilarity came out in section 3.3, where I noted that none of the direct moral beliefs held by Hare's speaker are tacit: they are, all of them, explicit. As I pointed out in that section, however, this is due to the way I chose to formulate Hare's semantic account, which can be formulated even starting with a speaker who lacks this feature. Hare's speaker has made up her mind about what direct moral beliefs to accept so, under the respects just mentioned, she differs, as I say, from most actual speakers: are there *any* actual speakers who resemble Hare's speaker under these respects? The best candidates are moral philosophers with a systematic account of the principles pertaining to some moral feature. Such speakers make up what is in any case a tiny group even among actual speakers who are moral philosophers, not to mention among all actual speakers.

place, is consistent and, in the second place, has acquired a determinate opinion about any categorical moral issue and about any non-moral issue -with the beliefs in the axiomatic set capturing the general principles held by such a speaker-. Under both these respects, such a user of moral language is of course very different from any actual speaker.

The way Hare and Gibbard use their respective idealized speakers to develop their semantic accounts also differs, this time quite a lot. Hare's strategy is to 'go by the speaker's standards': in order to assign content to a moral sentence, replace the moral expression occurring in it with a non-moral one, determined by the standards of the speaker uttering the sentence. On Hare's account, the content of a moral sentence, as uttered by a given idealized speaker, is thus specified by looking at *that* speaker's standards: it is determined by the specific standards of the speaker uttering the sentence. As we saw in the previous chapter, this is done by having the content of a moral sentence, as uttered by a given speaker, be given by a conjunction of two elements: a moral conjunct, listing the standards of that speaker, and a non-moral one, determined by those standards[20].

Gibbard's strategy, on the other hand, is to 'go by the hyperstates': the content of a moral sentence is a matter of all the various ways in which one can become fully decided and accept it. On Gibbard's account, the content of a moral sentence is thus specified by looking at *all* idealized speakers. As we saw above, this is done by having the content of a moral sentence be given by the big disjunction of all the conjunctions of the sort figuring in Hare's account. The upshot is that, on Gibbard's account, the content of a moral sentence is not tied to the specific standards of any idealized speaker, not even those of the speaker uttering the sentence. Take, for

---

[20] As we saw in chapter 3, Hare distinguishes between a class of peripheral uses of categorical moral sentences and a central, more typical use of such sentences; he relies on non-enriched singular substitution to account for the former and on enriched singular substitution to account for the latter. In the text I focus on Hare's account of the central, more typical use of categorical moral sentences. Analogous remarks apply, *mutatis mutandis*, to his account of the peripheral uses of such sentences.

instance, individual moral sentences. On Gibbard's account, the content of 'x is M' is given by a big disjunction. The standards of a given speaker determine which specific disjunct that speaker will accept, if she accepts that x is M: for example, in the case of an idealized speaker whose only standard is that ■ all N_1's are M, the relevant disjunct will be 'x is N_1 and ■ all N_1's are M'. The content of 'x is M', however, is not tied to the specific standards of any given speaker, since it is not, in contrast with Hare's account, captured by any one of the disjuncts: it is captured by the disjunction.

As a result, though both Hare and Gibbard start with an idealized user of moral language, only Gibbard ends up assigning content to moral sentences as uttered by actual, non-idealized speakers as well. On Hare's account, the content of a moral sentence is determined by looking at the specific standards of the speaker uttering the sentence. A specific set of standards is easily available in the case of an idealized speaker; however, when it is a speaker like most of us who utters the sentence, it is not clear which, among all the possible sets of standards, to look at, since most of us have not settled on any specific one: most of us have not made up our minds about what direct moral beliefs to accept. As I pointed out in 3.2, it is thus not clear how to assign content to a moral sentence as uttered by non-idealized speakers like most of us. Hare is not explicit about how exactly to deal with non-idealized speakers, and though, as I noted in the previous chapter, there are some suggestions he could try in order to take at least some steps toward closing the gap between idealized and non-idealized speakers, there is further work for him to do here[21].

On Gibbard's account, on the other hand, the content of a moral sentence is not determined by looking at the specific standards of any speaker. Even in the case of an idealized

---

[21] As I also noted in the previous chapter, none of the objections I there raised against the 'Go by the speaker's standards' strategy rely on the existence of such a gap.

189

speaker, the specific set of standards she holds plays no special role in determining the content of a given moral sentence, which is captured by a big disjunction. No further work is thus required in order to deal with a speaker who, like most of us, does not hold any specific set of standards: the same disjunction can be offered as capturing the content of the relevant moral sentence, even as uttered by such a speaker. An idealized speaker and I differ in that, in her case, there is a specific disjunct she will accept, if she accepts a given moral sentence, whereas in my case there is no such disjunct. The content of the sentence, however, is given for both of us by the disjunction.

Gibbard's strategy achieves another noteworthy result as well. As we saw in chapter 3, Hare's strategy is an attempt to mimic the cognitivist proposals described in chapter 2 that rely on singular substitution. Though, as I argued in the previous chapter, such an attempt fails, it is pretty clear how Hare proceeds in order to carry it out: whenever the cognitivist makes use of the non-moral predicates denoting the n's (or their disjunction), use the non-moral predicates denoting the n_S's (or their disjunction) instead. One natural response to my arguments, I pointed out, is for the non-cognitivist to grant that Hare's attempt is unsuccessful and try mimicking the cognitivist who relies on general substitution. To illustrate with individual moral sentences: the idea is that the non-cognitivist should mimic the proposal that the content of 'x is M' is that x has one of the n's (or that x has the disjunctive property built out of the n's). The problem with this suggestion is that it is not at all clear *how* exactly the non-cognitivist can do that: what account, different from Hare's, can she put forward, using only materials available to her, that would allow her to mimic the cognitivist who relies on general substitution? Clearly, she cannot proceed by using 'the n_S's' whenever the cognitivist makes use of 'the n's': since 'x has one of the n_S's' (or 'x has the disjunctive property built out of the n_S's') lacks any moral content, the

upshot would be that the content of 'x is M' is entirely non-moral, which is not the result the non-cognitivist wants. As I noted in 3.6, the source of the problem for the non-cognitivist is that 'x has one of the *n*'s' (or 'x has the disjunctive property built out of the *n*'s') is an existential claim (given the stipulation that 'the *n*'s' be treated *à la* Russell, so that I can avoid having to write down the Russellian *analysans* every time I need to frame the cognitivist proposal), and it is not clear the non-cognitivist can avail herself of even so basic an operation as existential generalization, hence of an existential claim like those the cognitivist makes use of.

Gibbard's strategy shows the non-cognitivist a way out of this problem: it shows her how to proceed in her attempt to mimic the cognitivist who relies on general substitution. One way to think of an existential claim is as a (possibly infinite) disjunction. According to the cognitivist who relies on general substitution, the content of 'x is M' is captured by the existential claim 'x has one of the *n*'s' (or 'x has the disjunctive property built out of the *n*'s'). The non-cognitivist cannot avail herself of such a claim. She can, however, try to mimic it by means of the big disjunctive claim that (c.G') (or (c.G'')) offers as capturing the content of 'x is M'. The way to proceed for the non-cognitivist, in her attempt to mimic the cognitivist who relies on general substitution, is to 'go by the hyperstates' and adopt Gibbard's account of moral sentences. Whether Gibbard's account ultimately succeeds in mimicking the cognitivist proposals that rely on general substitution will depend on whether it is immune to objections that do not apply to those proposals, such as the objections I raised in the previous chapter against Hare's account: we'll look into this starting from the next section. But at the very least it enables the non-cognitivist to carry out her mimicking attempt: argument is now needed to show that the attempt fails[22].

---

[22] Remember (n. 16) that in the present work I am bracketing the worry that non-cognitivism may have trouble dealing with negation, since the issues it raises are distinct from those I want to raise here, so what I say in the text is

Indeed, armed with Gibbard's account, the non-cognitivist can, and Gibbard does, take a further step. In chapter 2 I stipulated that the expression 'the *n*'s' be treated *à la* Russell. This is a stipulation because 'the *n*'s' is shorthand for 'the non-moral properties that can be denoted by some non-moral predicate 'N' such that ■ all N's are M', and how such an expression is to be treated is controversial. As I said in chapter 2, my stipulation is not meant to play any substantive argumentative role: I make it only in order to avoid having to write down the Russellian *analysans* every time I need to frame the cognitivist proposals that rely on general substitution. According to such proposals, the content of a moral sentence is that there are some properties for which certain conditions hold true, where the conditions in question are those mentioned by a treatment *à la* Russell of the expression 'the *n*'s'; in order to have a quick way of framing such proposals, I thus appropriated the expression 'the *n*'s', stipulating that it be treated *à la* Russell.

Given this stipulation, the non-cognitivist cannot avail herself of the claims that the cognitivist who relies on general substitution puts forward as capturing the content of moral sentences. Using again individual moral sentences to illustrate the point, the content of 'x is M' is, according to the cognitivist, that x has one of the *n*'s (or that x has the disjunctive property built out of the *n*'s), i.e. that there are some properties of which the following things are true: each of them is a non-moral property that can be denoted by some non-moral predicate 'N' such that ■ all N's are M, …; but on non-cognitivism there are no such truths as that ■ all N's are M. What the non-cognitivist can do is to try and mimic the claims that the cognitivist who relies on general substitution makes use of, and Gibbard's account shows her how to do this.

---

subject to a qualification: Gibbard's strategy enables the non-cognitivist to carry out her mimicking attempt, *on the assumption* that the worry about negation can be answered, an assumption which is controversial (and which, as I said, I believe is false). Even so, the progress for the non-cognitivist is real, since I was bracketing the worry about negation in the previous chapter too, when raising the problem that it is not at all clear how the non-cognitivist can proceed in her mimicking attempt. What solves this problem is the insight that the non-cognitivist can use disjunctions in her attempt to mimic existential claims, and the full articulation and systematic defense of this insight are due to Gibbard.

However, once she has gotten here, the non-cognitivist can, and Gibbard does, go further and claim to have earned the right to use the expression 'the n's' itself. She can remind us that my stipulation is just for ease of exposition and is not meant to play any substantive argumentative role; if we ask the substantive question how 'the n's' is to be treated, the answer, she can claim, is that it is not to be treated *à la* Russell, but in terms of the big disjunctions employed by Gibbard's account: for instance, she can claim that 'x has one of the n's' is to be accounted for in terms of the big disjunction mentioned by (c.G'). If she takes this further step, the non-cognitivist thus can, as Gibbard does, frame her proposals about the content of moral sentences using the same claims that the cognitivist who relies on general substitution makes use of: for instance, she too can claim that the content of 'x is M' is that x has one of the n's. The disagreement with the cognitivist will be about how to account for such claims as 'x has one of the n's'[23].

As I said, my stipulation is not meant to play any substantive argumentative role, so I regard this further step by the non-cognitivist as entirely legitimate. Given the non-cognitivist's overall position, though, this is not the crucial step. The non-cognitivist's aim is to mimic the cognitivist who relies on general substitution, and Gibbard provides her with an account of moral sentences that enables her to carry out her mimicking attempt. The crucial question for the non-

---

[23] This explains why (as I pointed out in n. 17) Gibbard feels free to frame his proposal about moral explanatory claims using the definite description 'the property that constitutes being M' (or rather, strictly speaking, 'the property that constitutes being the thing to do', since he frames his claims as claims about normative language). I should mention explicitly that my presentation of Gibbard's views here is a reconstruction of his position: he does not present things quite the way I do in the text. In particular, it is a confusing fact about his book that he feels free throughout to frame his claims using the definite description 'the property that constitutes being M', without explicitly addressing the question how he can make use of it. What I offer in the main text is my attempt to reconstruct an answer on his behalf. (One especially confusing passage is Gibbard's discussion of what he calls the "Claim of Factual Constitution", which I mentioned in the previous section: see Gibbard [2003], pp. 94-98. The claim he sets out to defend there is that there is such a property as the property that constitutes being M, all his argument shows is that, for any user of moral language who has become fully decided, there is such a property as the property that that speaker takes to constitute being M. In my terminology, Gibbard's discussion seems to make the mistake, which I noted above the non-cognitivist should avoid, of replacing 'the n_S' for 'the n'. On my reconstruction of Gibbard's position, his discussion can be seen to avoid this mistake, though I won't argue for this here.)

cognitivist is whether the attempt is ultimately successful: does Gibbard's account succeed in mimicking the cognitivist proposals that rely on general substitution? Whether it does depends on whether it is immune to objections that do not apply to those proposals. If it is, then the non-cognitivist has achieved her aim, *whatever* the fate of her further claim to have earned the right to use the expression 'the *n*'s'.

My point here is quite general. The nature of the mimicking project is sometimes described as follows. The cognitivist can, in putting forward her claims about morality, avail herself of certain notions / expressions that are, at least on the face of it, bound up with the notion of moral truth: for instance, she can avail herself of the notion of knowledge, or of reference, or –as in our case- of the expression 'the *n*'s'. In order to show that she can mimic those cognitivist claims, the non-cognitivist needs to show that she can, using only materials available to her, reconstruct those notions / expressions, so that she too can avail herself of them: she has to, as it is often put, "earn the right" to use those notions / expressions. This is how non-cognitivists themselves sometimes describe their mimicking project. But it is not, it seems to me, the best way of characterizing that project. It is of course a very interesting result if the non-cognitivist can show she can avail herself of the notions / expressions in question, but she need not get this far in order to show she can mimic the relevant cognitivist claims. In order to do this, what she needs to show is that she can, using only materials available to her, put forward a claim that does for her the same philosophical work (for example, accommodating our practice with regard to moral explanatory claims) done for the cognitivist by the cognitivist claim she is trying to mimic. *This* is what winning the mimicking game amounts to. If the non-cognitivist succeeds in doing this, her success will no doubt lend plausibility to the view that the notions / expressions employed by the cognitivist are within the non-cognitivist's reach. It is not obvious, however,

194

that it will settle the matter. Whether the notion of, say, knowledge, or of reference, or –as in our case- the expression 'the n's' is available to the non-cognitivist will depend, one might think, also on general considerations drawn from epistemology, or the philosophy of language. Assessment of the view that the notions / expressions in question are available to the non-cognitivist quickly becomes an interesting but also quite complex matter. My point here -and it is a friendly point to the non-cognitivist- is that, whether or not it is in the end correct, that view goes in any case beyond what is required for the mimicking project to succeed: the fate of such a project does not hinge on it.

In the discussion to follow, I'll therefore bracket Gibbard's further contention that he has earned the right to use the expression 'the n's'. Indeed, since I'll continue to need a quick way of framing the cognitivist proposals that rely on general substitution, I'll stick to the stipulation that 'the n's' be treated *à la* Russell. As I noted above, given this stipulation the non-cognitivist cannot avail herself of the claims that the cognitivist who relies on general substitution puts forward as capturing the content of moral sentences; what she can do is to try and mimic those claims. Gibbard's crucial contention is that his account of moral sentences succeeds in doing precisely this: I'll look into it starting from the next section.

Just as in the case of the proposals we looked at in chapters 2 and 3, it is useful to have a name for the substitution operation that Gibbard's proposals rely on. Since, as we have just seen, Gibbard's proposals are close to the cognitivist proposals that rely on general substitution, let us use 'multiple / disjunctive general substitution' as a name also for the substitution operation that (c.G') and (C.G') / (c.G'') and (C.G'') rely on. As we saw in chapter 2, the operation of general substitution relied on by the cognitivist proposals is defined in terms of the n's, i.e. the non-moral properties that can be denoted by some non-moral predicate 'N' such that ∎ all N's are M:

that operation takes in a moral sentence and returns a sentence that replaces the original sentence's apparent reference to a moral property with a reference to the n's. The operation of general substitution relied on by Gibbard's proposals, on the other hand, takes in a moral sentence and returns a disjunction: each disjunct is a conjunction of two elements, the first of which replaces the original sentence's apparent reference to a moral property with a reference to the properties picked out by the universal moral sentences that make up the second conjunct (or with a reference to the disjunction of such properties). As we saw above, such universal moral sentences can –for each disjunct- be thought of as capturing the standards of a user of moral language who has become fully decided and accepts the original moral sentence -with one disjunct for each speaker with a different set of standards-, so we can say that the operation of general substitution relied on by Gibbard's proposals is defined in terms of the non-moral properties, denotable by some non-moral predicate, picked out by S's standards, for each possible user S of moral language who has become fully decided: for short, in terms of the n_H's - keeping in mind that Gibbard's term for the views of such speakers is "hyperstates"-. Since I am using the name 'general substitution' for the operation relied on by both the cognitivist's and Gibbard's proposals, I'll add the phrase '*via* reference to the n's (the n_H's)' when I need to refer only to the operation of general substitution relied on by the cognitivist's (Gibbard's) proposals.

With this terminology in place, we can now articulate Gibbard's 'Go by the hyperstates' strategy in a fully general way. So far I have been dealing only with individual moral sentences and moral explanatory claims, but the strategy can be applied more broadly, to yield an account of all moral sentences (of the relevant type). Here is how one can apply that strategy to specify the content of all moral sentences, starting from a subset of them whose content is assumed as given:

1.G)   All moral sentences (of the relevant type) are either <u>built sentences</u>, i.e. their content is specified by means of other moral sentences (of the relevant type) or <u>basic sentences</u>, i.e. their content is not specified by means of other moral sentences (of the relevant type).

2.G)   The content of all built moral sentences is specified by means of basic moral sentences.

3.G)   The set of basic moral sentences is the set of appropriately strong universal moral sentences that can be formulated using a non-moral predicate denoting a non-moral property.

4.G)   The content of all built moral sentences is specified in terms of (multiple / disjunctive) general substitution, *via* reference to the non-moral properties, denotable by some non-moral predicate, picked out by S's standards, for each possible user S of moral language who has become fully decided[24].

As I said above, 1.G)-4.G) provide a recipe to specify the content of all moral sentences, starting from a subset of them whose content is assumed as given. The subset in question is that defined by thesis 3.G): the set of appropriately strong universal moral sentences that can be formulated

---

[24] It is easy enough, starting from the case of individual moral sentences and moral explanatory claims, to define the operation of general substitution *via* reference to the n_H's for the case of universal moral sentences, so I'll spare the reader the technical details. Theses 1.G)-4.G) match theses 1)-4) presented in chapter 2, which articulate the cognitivist 'Go by the M-making properties' strategy, and theses 1.H)-4.H) presented in chapter 3, which articulate Hare's 'Go by the speaker's standards' strategy. Indeed, the only difference among the three sets of theses lies in the fourth thesis, i.e. in the substitution operations in terms of which the content (or, in the case of 4), payoff) of built moral sentences is specified.

using a non-moral predicate denoting a non-moral property[25]. Since Gibbard's account is a non-cognitivist account, the content of these sentences is specified (as in the case of Hare's account) in terms of the core non-cognitivist idea that, in uttering a sentence of the form ' ■ All N's are M', a speaker is expressing the relevant non-cognitive attitude (commendation, condemnation, etc., depending on what 'M' is) toward N's[26].

Theses 1.G)-4.G) articulate in some detail the strategy summarized by the slogan 'Go by the hyperstates'. As I noted above, given thesis (G) this is a strategy that should strike you as a plausible way to go, if your aim is to replace a moral sentence's apparent reference to a moral property with a reference to non-moral properties. The application of this general recipe to the particular case of moral explanatory claims yields Gibbard's accommodationist proposal, which is the second non-cognitivist accommodationist proposal we need to discuss.

Gibbard's 'Go by the hyperstates' strategy is designed to be the strategy that, as I argued in the previous chapter, the non-cognitivist needs. At the end of my discussion of Blackburn's accommodationist proposal in section 3.5, I noted there are several reasons why the non-cognitivist needs a better strategy for assigning content to moral sentences than the 'Go by the speaker's standards' strategy I presented in chapter 3, which is central to the non-cognitivist tradition and of which Blackburn's proposal is an application. One reason has to do with our main topic here, i.e. her accommodationist project: if the non-cognitivist wants to accommodate our practice with regard to moral explanatory claims, she needs a better accommodationist proposal than Blackburn's, since Blackburn's proposal fails to accommodate our actual practice.

---

[25] I should remind the reader that the modal expression 'sentences that *can* be formulated' covers both sentences that require, for their formulation, a non-moral predicate we already have and sentences that require a non-moral predicate we do not already have but can come up with.

[26] For Gibbard's own take on this idea, which he fleshes out in terms of the acceptance of norms, see Gibbard [1990].

The other reasons have to do with non-cognitivism itself: the 'Go by the speaker's standards' strategy fails to provide an adequate account of the content of moral sentences, starting from the central class of them made up by categorical moral sentences; it does not allow the non-cognitivist to mimic, as she typically claims she can do, the cognitivist 'Go by the M-making properties' strategy I presented in chapter 2; and it fails the non-cognitivist where she has traditionally claimed to be at a decisive advantage, i.e. accounting for agreement and disagreement among speakers with different moral standards, thereby leaving her without the support of the Master Argument for non-cognitivism.

In section 3.6, I pointed out that, since the 'Go by the speaker's standards' strategy is an attempt to mimic the cognitivist who relies on singular substitution, a natural thought for the non-cognitivist in search of a better strategy is to try and mimic the cognitivist who relies on general substitution instead. Gibbard's 'Go by the hyperstates' strategy promises to be just the strategy the non-cognitivist needs. Before I begin to discuss whether it lives up to this promise, looking into whether it allows the non-cognitivist to solve the problems I raised in the previous chapter for the 'Go by the speaker's standards' strategy, I want to end this section by pointing out one crucial aspect of the older non-cognitivist strategy that the newer one leaves behind.

This is the doctrine of the descriptive meaning of moral expressions. Consider again the idealized users of moral language that the two strategies start with. On the 'Go by the speaker's standards' strategy, the content of a moral sentence, as uttered by a given such speaker, has a non-moral component, one that can be captured replacing the moral expression in the sentence with a non-moral one, determined by that speaker's standards. So, for example, it is part of the content of 'x is M', as uttered by an idealized speaker whose only standard is that ∎ all N_1's are M, that x is N_1: this is the descriptive meaning of 'x is M', as uttered by that speaker. On the 'Go

by the hyperstates' strategy, on the other hand, the content of a moral sentence is given by a big disjunction, and the standards of any given idealized speaker play no special role in determining that content. So, for example, even in the case of an idealized speaker whose only standard is that ■ all N_1's are M, the content of 'x is M' is given by a big disjunction, in which the disjunct 'x is N_1 and ■ all N_1's are M' is just one disjunct among others[27]. On that strategy, it is therefore no part of the content of 'x is M', even as uttered by such a speaker, that x is N_1: that sentence does not, even as uttered by such a speaker, ascribe the property of being N_1 to x. Indeed, on the 'Go by the hyperstates' strategy it is no part of the content of 'x is M', as uttered by anyone, that x is N, for any non-moral property that can be denoted by some non-moral predicate 'N'. Gibbard's account of moral sentences thus gives up the non-cognitivist doctrine of the descriptive meaning of moral expressions. Given the centrality of this doctrine in the non-cognitivist tradition, this is a point worth stressing; it is also, I should add, good news for the non-cognitivist, given the arguments I offered in the previous chapter to reject this doctrine[28].

In what follows, I'll offer two reasons for thinking that Gibbard's accommodationist proposal fails, and more generally for rejecting his account of moral sentences. I'll proceed as

---

[27] Though, as I noted above, this is the disjunct such a speaker will accept, if she accepts that x is M, *what* she accepts, in accepting that x is M, is given by the disjunction.

[28] As we saw in chapter 3, Hare's doctrine of the descriptive meaning of moral expressions was introduced to address a problem for non-cognitivists, i.e. explaining our use of categorical moral sentences to convey non-moral information. If the non-cognitivist gives up that doctrine, she needs a different solution to this problem. The obvious option is to turn to the solution that (as we saw in 3.2: see ch. 3, n. 15) Hare himself considered but eventually rejected, i.e. saying that the non-moral information is something the hearer infers from the speaker's utterance of the categorical moral sentence, together with her knowledge of the speaker's standards, rather than part of the meaning of the sentence. It is in this connection interesting to note that, although Gibbard's meta-ethical theory is in many respects close to Hare's, one does not find in Gibbard's work the same emphasis one finds in Hare's on the fact that moral sentences can be used to perform some of the same jobs performed by non-moral sentences, such as conveying non-moral information. In addition to the one mentioned in the text, a further difference worth noting between Gibbard and the non-cognitivist tradition (though one not directly related to the topic of the present work) is that the sort of agreement and disagreement Gibbard takes to be key to understanding moral thought and language is agreement and disagreement in plan, which (as he himself points out: see Gibbard [2003], pp. 68-71) differs in important respects from the agreement and disagreement in attitude, central to the non-cognitivist tradition, that Stevenson focused on, so much so that people can agree (disagree) in plan while disagreeing (agreeing) in attitude. I thank Nick Sturgeon for pointing out to me the importance of this further difference.

follows. I'll start by discussing, in the remainder of this chapter, whether Gibbard's account solves the three problems I raised in the previous chapter for the 'Go by the speaker's standards' strategy. I'll show that it does: new arguments are needed against the newer non-cognitivist strategy, different from those that work against the older one. I'll then offer two such arguments in the next chapter.

3. <u>Good news for the 'Go by the hyperstates' strategy</u>

In the previous chapter, I raised three problems for the 'Go by the speaker's standards' strategy. Each of them, I argued, can be raised both for categorical moral sentences and for moral explanatory claims. In this section, I'll show that Gibbard's account of moral sentences is immune to all of them. For simplicity's sake, I'll focus on Gibbard's account of categorical moral sentences; analogous considerations apply to his account of moral explanatory claims as well.

The first objection I raised against the 'Go by the speaker's standards' strategy is that it rules out as unavailable an option we do have available when we develop our moral views. Consider a user S of moral language who has a fully worked out theory about what things are M, for a given moral predicate 'M'. Suppose there is only one sort of things S thinks are M: those with the property of being N_1. Suppose further that S holds an individual moral belief about an object x: she thinks x is M. Imagine now S finds out that x is not N_1. Having learnt this, S needs to revise some of her moral beliefs: what are her options? Intuitively, there are at least two ways S can go: she could stick to her moral theory and revise her belief that x is M, or she could stick to her belief about x and revise her moral theory. But, so goes my objection, Hare's account of

categorical moral sentences rules the second option out. On that account, S can only go one way: stick to her moral theory and revise her belief that x is M.

Turn now to Gibbard's account of categorical moral sentences. Gibbard's account is built on thesis (G), which is about complete belief sets that are consistent, and (as we saw in 4.1) one way to think of a belief set with these features is as capturing the views of a user of moral language who has become fully decided, with the beliefs in the axiomatic set capturing the general principles held by such a speaker. Our speaker S need not, of course, be a fully decided user of moral language; she is, however, decided in one respect, we are supposing: she has settled on a specific set of standards, holding only one general principle, namely that ∎ all N_1's are M. In the terminology of thesis (G), S's views are thus modeled, not by a single complete belief set that is consistent, but by a set of such belief sets: those whose axiomatic set contains only one categorical moral belief, i.e. the belief that ∎ all N_1's are M; or, more accurately, those whose axiomatic set contains only the belief that ∎ all N_1's are M and that include the other belief we are supposing S holds, i.e. that x is M. On Gibbard's account of categorical moral sentences, S's belief that x is M amounts to the belief in the disjunction mentioned by (c.G') (or (c.G'')). Given this belief, and given that she has settled on a specific set of standards and holds only one general principle, i.e. that ∎ all N_1's are M, there is a specific disjunct of that disjunction S is committed to accepting, namely 'x is N_1 and ∎ all N_1's are M'. And, of course, after she finds out that x is not N_1, S cannot accept this disjunct any longer. This, however, does not mean she cannot hold her belief that x is M any longer, since this belief amounts to the belief, not in that disjunct, but in the disjunction. So on Gibbard's account, S does have the option of sticking to her belief that x is M, even after learning that x is not N_1; if S goes this way, she cannot hold on to her set of standards any longer: she is committed to revising her moral

theory. A different option S also has is to stick to her moral theory; if S goes this way, she cannot accept the disjunction mentioned by (c.G') (or (c.G'')) any longer: she is committed to revising her belief that x is M. Gibbard's account thus makes room for both of the options that intuitively S has available in developing her moral views[29].

---

[29] Gibbard does not address the three problems I raised in chapter 3 directly, and what I present in this section are the solutions to them that his account of moral sentences puts him in a position to offer. Indeed, when he briefly touches on the question of how we reason about our moral views, what he does say is at odds with the solution to the first of those problems presented in the text. In order to raise that problem, I consider a user of moral language who has a fully worked out moral theory and, as I noted above, such a speaker need not be fully decided. She need not, but she can: I could raise the problem just as well focusing on a user of moral language who has become fully decided and is forced to revise some of her moral beliefs by new non-moral evidence. Talking about fully decided speakers in his discussion of moral explanatory claims, Gibbard contrasts the kind of evidence in support of these claims that such speakers have with that available to ordinary speakers, who lack (among other things) a fully worked out moral theory: whereas an ordinary speaker can accept a moral explanatory claim on such grounds as authority or induction, the evidence of a fully decided speaker, claims Gibbard, "can be factored into a purely causal part and a purely [moral] part" and is expressed by one of the disjuncts (a different one for each fully decided speaker with a different set of standards) making up the disjunction that, on Gibbard's account, captures the content of the moral explanatory claim (Gibbard [2003], p. 206; the same claim is made on p. 204 as well). So, for instance, if a fully decided speaker whose only standard is that ■ all N_1's are M accepts that M-ness causes E, her evidence in support of this claim is expressed, claims Gibbard, by 'N_1-ness causes E and ■ all N_1's are M'. Note two things about Gibbard's claim. One: on Gibbard's account of moral explanatory claims, there is no reason to expect this claim to be true. On that account, the content of a moral explanatory claim is given by a disjunction, and the evidence in support of a disjunction need not in general amount to any of its disjuncts, as Gibbard would of course agree, given what he says about the evidence available to ordinary speakers. Why expect things to be different when it comes to fully decided speakers? Two: if Gibbard's claim is true, then the solution presented in the text fails, once the problem it is meant to solve is raised focusing on a fully decided speaker. Consider again a fully decided speaker whose only standard is that ■ all N_1's are M and who accepts that M-ness causes E. If such a speaker finds out that N_1-ness does not cause E, she cannot of course accept the disjunct 'N_1-ness causes E and ■ all N_1's are M' any longer. As I say in the text, this in itself does not mean she cannot accept that M-ness causes E any longer, since on Gibbard's account the content of this claim is given, not by that disjunct, but by the whole disjunction mentioned by (C.G') (or (C.G'')). However, if we now plug in Gibbard's claim that our speaker's evidence for the moral explanatory claim is captured by that disjunct -a disjunct she cannot accept any longer-, we are back to the result that our speaker has only one way to go: stick to her moral theory and revise her moral explanatory belief. And that Gibbard's claim about the kind of evidence available to a fully decided speaker has this consequence should come as no surprise, given that it comes very close to making, about the moral explanatory beliefs held by a fully decided speaker, the same epistemological claim that theses i.H)-iv.H) make about our inferred categorical moral beliefs (whereas, as I pointed out in section 4.1, thesis (G) makes no claim about what even a fully decided speaker infers her categorical moral beliefs from). As I said, the claim of Gibbard's we have been looking at in this footnote occurs in a brief discussion on how we reason about our moral views, and it plays no role in Gibbard's overall meta-ethical theory, so in my presentation of Gibbard's views I dropped it. I take its presence in Gibbard's discussion as a sign both of the persistent influence of Hare's reading of the Assumption and of how tempting it is to blur the line I drew in 4.1 between saying (as Gibbard can say, and is guaranteed by thesis (G)) that we can reconstruct the beliefs of a fully decided speaker *as if* they conformed with i.H)-iv.H) and saying (as Gibbard cannot say) that i.H)-iv.H) are true of the beliefs of such a speaker. (It might be objected that the claim of Gibbard's we have been looking at in this footnote is not really problematic, since it is about fully decided speakers, whereas the first of my three problems concerns how speakers develop their moral views, so it cannot be raised with regard to a speaker who has become fully decided. But I take it that a fully decided speaker can change her mind, moving from a fully decided state to a not fully decided, or a different fully decided, state; after all, it is not that we should think of a fully decided speaker

The second objection I raised against the 'Go by the speaker's standards' strategy is that, on that strategy, moral sentences uttered by moral particularists have no content. On Hare's account, a categorical moral sentence, as uttered by a speaker S, presupposes that there is at least one non-moral property that can be denoted by some non-moral predicate 'N' such that S believes that ■ all N's are M; if, in the case of a given speaker, there turn out to be no such properties, the upshot is thus that a categorical moral sentence, as uttered by that speaker, has no content; some speakers endorse moral particularism, i.e. the denial of all appropriately strong universal moral sentences using non-moral predicates; it follows that categorical moral sentences, as uttered by moral particularists, have no content; but, so goes my objection, when Dancy makes a remark like 'That action is right', there is something he is saying.

Gibbard's account avoids this problem for the same reason why, as we saw in 3.4, the cognitivist proposals described in chapter 2 do. That is: there is a possible argument against Gibbard's account that mirrors this objection to Hare's account, but it is a much harder argument to press, since it relies on a premise that is quite difficult to establish. It is easier to frame the argument if I introduce a little bit of terminology. Consider the disjunction that, on Gibbard's account, captures the content of a categorical moral sentence: each disjunct is a conjunction of two elements; of these two conjuncts, the second is made up of appropriately strong sentences of the form 'All N's are M' (where 'N' is a non-moral predicate denoting a non-moral property). Consider now the disjunction built out of such conjuncts: '(■ all N_1's are M) or (■ all N_2's are

---

as blinding herself to new evidence. Remember that what I need in order to raise the problem is a speaker who is forced to revise some of her moral beliefs by new non-moral evidence, so for the present purposes you can think of it this way: a fully decided speaker is defined as a speaker who, among other things, has a determinate opinion about any non-moral issue; it is no part of the definition of a fully decided speaker that she holds *true* non-moral opinions. The reason why, in raising the problem in the text, I choose not to focus on fully decided speakers is not that such speakers cannot change their mind, but that they involve so high a degree of idealization, and I prefer, when possible, not to rely on so highly idealized cases in my arguments.)

M) or … or ($\blacksquare$ all N$_1$'s are M and $\blacksquare$ all N$_2$'s are M) or ($\blacksquare$ all N$_3$'s are M and $\blacksquare$ all N$_4$'s are M) or … or …'. As shorthand for this disjunction, use 'Disj'[30].

Here then is a possible argument against Gibbard's account: on Gibbard's account, a categorical moral sentence entails that Disj[31]; if not (Disj), the upshot is thus that nothing is right or wrong, good or bad, etc.; assume moral particularism, i.e. the denial of all appropriately strong universal moral sentences using non-moral predicates; it follows that nothing is right or wrong, good or bad, etc.; this consequence shows that Gibbard's account is unable to accommodate our actual practice. Note, however, how much harder it would be to press this argument against Gibbard's account: whereas the objection to Hare's account relies on the uncontroversial premise that there are moral particularists, one of the premises of this argument is moral particularism itself, and this is a premise that is quite difficult to establish, which is why I won't explore this argument in the present work[32].

---

[30] I should remind the reader (see n. 16) that each disjunct of 'Disj' is shorthand for something more complex: the disjunct ' $\blacksquare$ all N$_1$'s are M' is shorthand for ' $\blacksquare$ all N$_1$'s are M *and* not ($\blacksquare$ all N$_2$'s are M) and not ($\blacksquare$ all N$_3$'s are M) and …'; the disjunct ' $\blacksquare$ all N$_1$'s are M and $\blacksquare$ all N$_2$'s are M' is shorthand for ' $\blacksquare$ all N$_1$'s are M and $\blacksquare$ all N$_2$'s are M *and* not ($\blacksquare$ all N$_3$'s are M) and not ($\blacksquare$ all N$_4$'s are M) and …'; analogously for all other disjuncts.

[31] Each disjunct of the disjunction that, on Gibbard's account, captures the content of a categorical moral sentence entails one disjunct of 'Disj', hence it entails that Disj; therefore, the categorical moral sentence entails that Disj.

[32] Here is a point worth noting about the relationship between Gibbard's meta-ethical theory and moral particularism. As I pointed out in 4.1, thesis (G) formulates a consistency requirement for complete belief sets, and the requirement is conditional: if a complete belief set includes categorical moral beliefs, there is a non-empty subset of those beliefs with the features mentioned by i.G) and ii.G). Given what ii.G) says, it follows from (G) that it is a consistency requirement for a complete belief set including categorical moral beliefs that it include beliefs that can be expressed by an appropriately strong universal moral sentence using a non-moral predicate denoting a non-moral property. As I also pointed out in 4.1, a further consistency requirement Gibbard takes to apply to complete belief sets is that any such belief set include categorical moral beliefs. Since this further requirement forms no part of Gibbard's reading of the Assumption, I said I would bracket it in the present work. Consider now what follows from the conjunction of the two requirements: if a complete belief set is consistent, it includes beliefs that can be expressed by an appropriately strong universal moral sentence using a non-moral predicate denoting a non-moral property. It follows from this that there is no complete belief set that is consistent and contains the negation of all beliefs that can be expressed by an appropriately strong universal moral sentence using a non-moral predicate, i.e. that moral particularism is inconsistent: it violates the logic of moral predicates. And, no matter what we should in the end think of moral particularism, this is quite a strong claim to make about it. Gibbard does not note that this claim follows from his meta-ethical theory. On the one hand, one would expect to hear more from him about why we should accept it. As I noted in 4.1, Gibbard does not say much in defense of the requirement that a complete belief set include categorical moral beliefs, nor does he devote more than a brief paragraph and a footnote to what I there conjectured is its source, namely the requirement that a set of beliefs not rule out as impermissible all the

The third objection I raised against the 'Go by the speaker's standards' strategy is that it is unable to account for agreement and disagreement among speakers with different moral standards. Consider three users of moral language, S_1, S_2, and S_3, each of whom has a fully worked out theory about what things are M, for a given moral predicate 'M'. Suppose each of them employs only one standard, though a different one from that employed by either of the other two: S_1's only standard is that ■ all N_1's are M, S_2's only standard is that ■ all N_2's are M, and S_3's only standard is that ■ all N_3's are M. Suppose further that S_1 says 'x is M', S_2 says 'x is not M', and S_3 says 'x is M'. Intuitively, two speakers like S_1 and S_2 are as a rule disagreeing with one another, while two like S_1 and S_3 are as a rule agreeing with one another. But, so goes my objection, Hare's account can at most deal with the disagreement between S_1 and S_2: when it comes to S_1 and S_3, it follows from the account that they are not agreeing with one another. Indeed, the very feature of the account that is meant to allow Hare to deal with the disagreement between S_1 and S_2 commits him to the claim that S_1 and S_3 are disagreeing with one another as well.

---

alternatives available in a given situation (see Gibbard [2003], p. 56). On the other, there are *prima facie* good reasons to reject the claim in question: it seems to build too much into the logic of moral predicates, stipulating away what on the face of it looks like a substantive issue. Moreover, non-cognitivists themselves seem to have strong, independent reasons for finding the claim suspicious. Consider the negative part of the Master Argument for non-cognitivism, i.e. the part designed to show that the cognitivist who believes that moral predicates can be analyzed in terms of non-moral ones is unable to account for agreement and disagreement among speakers with different moral standards. One way in which the insight behind it is sometimes put is by saying that the cognitivist ends up building a specific moral principle into the meaning of moral predicates, thereby building too much into it: those rejecting the principle in question raise a substantive issue that cannot just be stipulated away from moral discussion (Gibbard himself puts things along these lines in his [1990]: see pp. 11-12). If this insight is right, it seems to apply to the case of the particularist as well. If it is too much to build a specific moral principle into the logic of moral predicates, the same seems to hold of building the availability of some moral principle or other into it: the issue raised by the particularist seems to be as much one that cannot be settled by a mere appeal to logic as does the issue raised by someone rejecting the cognitivist's favored moral principle (for a non-cognitivist who appeals to this sort of considerations to reject Hare's claim that 'universalizability' is built into the logic of moral predicates, see Blackburn [1984], p. 220 and [1991a], pp. 8-9). I mention this point about the relationship between Gibbard's meta-ethical theory and moral particularism only because it is, as I say, worth noting: it will not play any role in my discussion. For another writer claiming that Gibbard's requirement that a set of beliefs not rule out all the alternatives implausibly stipulates away (this time all by itself) a view that cannot be dismissed this easily, see Sturgeon [2009], p. 86; the view in question is that there are moral dilemmas, i.e. situations in which none of the alternatives available to an agent is morally permissible.

Turn now to Gibbard's account. As I noted above, the views of a speaker like S_1 are, in the terminology of thesis (G), modeled by a set of complete belief sets that are consistent: those whose axiomatic set contains only the belief that ■ all N_1's are M and that include the other belief we are supposing S_1 holds, i.e. that x is M; analogously for S_2 and S_3. On Gibbard's account, the content of S_1's remark is given by the disjunction mentioned by (c.G') (or (c.G'')), and so is the content of S_3's remark, while the content of S_2's remark is given by the negation of that disjunction. The account can therefore vindicate the intuitions about agreement and disagreement in a straightforward way: in saying what they say, S_1 and S_3 are accepting the same disjunction, hence they are agreeing with one another; in saying what they say, S_1 and S_2 are, respectively, accepting and rejecting the same disjunction, hence they are disagreeing with one another.

What turns out to be problematic for Hare's account is the agreement between S_1 and S_3, so let us take a closer look at how Gibbard's account deals with it. On Gibbard's account, I have just said, the content of S_1's remark 'x is M' is given by the disjunction mentioned by (c.G') (or (c.G'')). Given that she accepts this disjunction, and given that she holds only one general principle, i.e. that ■ all N_1's are M, there is a specific disjunct of the disjunction S_1 is committed to accepting, namely 'x is N_1 and ■ all N_1's are M'; analogously for S_3 with respect to the disjunct 'x is N_3 and ■ all N_3's are M'. And these two disjuncts are at odds with one another[33]. This, however, does not mean that, in saying what they say, S_1 and S_3 are disagreeing with one another, since the content of their remarks is given, not by these disjuncts, but by the disjunction.

---

[33] I should again remind the reader (see n. 16) that the clause ' ■ all N_1's are M' in the first disjunct is shorthand for ' ■ all N_1's are M *and* not (■ all N_2's are M) and not (■ all N_3's are M) and …', while the clause ' ■ all N_3's are M' in the second disjunct is shorthand for ' ■ all N_3's are M *and* not (■ all N_1's are M) and not (■ all N_2's are M) and …'.

Gibbard's account can thus avoid the implausible verdict yielded by Hare's account and vindicate the intuition that S_1 and S_3 are agreeing with one another.

So far, so good for Gibbard's account. It is time now for a recap and some scorekeeping. Gibbard's account avoids all the problems I raised in the previous chapter for the 'Go by the speaker's standards' strategy. Those problems, I argued in chapter 3, show that Blackburn's accommodationist proposal about moral explanatory claims fails. Moreover, since they are general problems with the 'Go by the speaker's standards' strategy, which is central to the non-cognitivist tradition, they affect the prospects, not only of the non-cognitivist accommodationist project, but of non-cognitivism itself: specifically, they affect the non-cognitivist's ability to provide an adequate account of the content of moral sentences, starting from the central class of them made up by categorical moral sentences; her ability to mimic, as she typically claims she can do, the cognitivist 'Go by the M-making properties' strategy I presented in chapter 2; and her ability to account for agreement and disagreement among speakers with different moral standards.

The 'Go by the speaker's standards' strategy is built on Hare's reading of the Assumption, which is one of two non-cognitivist readings put forward in the literature. In this chapter, after presenting the second non-cognitivist reading of the Assumption, which is due to Gibbard (section 1), I introduced the accommodationist proposal Gibbard builds on it, pointing out that this proposal can be seen as the application to the particular case of moral explanatory claims of an entirely general strategy that Gibbard uses to specify the content of all moral sentences of the relevant type, one meant to mimic the cognitivist 'Go by the M-making properties' strategy, and specifically the cognitivist proposals that rely on general substitution: I called this the 'Go by the hyperstates' strategy (section 2). I then showed that the 'Go by the

hyperstates' strategy is immune to all of the problems raised in chapter 3 for the 'Go by the speaker's standards' strategy (section 3).

As a result, one conclusion we can already draw is that Gibbard succeeds in providing the non-cognitivist with the means to account for moral agreement and disagreement. Given how important it is for the non-cognitivist that she be able to account for this phenomenon, this is no small result[34]. On the other hand, the following questions are still open: in the first place, whether Gibbard's proposal about moral explanatory claims succeeds in accommodating our practice with regard to such claims; secondly, whether his 'Go by the hyperstates' strategy provides an adequate account of the content of moral sentences quite generally, and one that succeeds in mimicking the cognitivist who relies on general substitution. How to answer these questions depends on whether there are other problems with Gibbard's proposals, different from those I raised in chapter 3. In the next chapter, I'll argue that the answers are negative. There are two further problems with Gibbard's accommodationist proposal, which show that that proposal fails. These problems also show that Gibbard's general account of moral sentences is inadequate. Finally, neither of these two problems applies to the cognitivist who relies on general substitution, or to any of the cognitivist proposals presented in chapter 2: Gibbard's attempt to mimic the cognitivist is ultimately unsuccessful, and despite all that he achieves on the non-cognitivist's behalf, it still makes a difference whether you are a cognitivist or a non-cognitivist.

---

[34] While this result allows the non-cognitivist to continue appealing to the Master Argument in support of her meta-ethical theory, how strong the support is will depend on how well the cognitivist in turn fares when it comes to accounting for moral agreement and disagreement. As I noted in chapter 3 (see ch. 3, n. 17), there are problems with the Master Argument's claim that the cognitivist fails here, but in the present work I am bracketing them. Another thing I am here bracketing (see n. 16) is the worry that non-cognitivism may have trouble dealing with negation, so what I say in the text is subject to the usual qualification: Gibbard succeeds in providing the non-cognitivist with the means to account for moral agreement and disagreement, *on the assumption* that the worry about negation can be answered.

Chapter 5 - The second non-cognitivist strategy: why it fails too

1. <u>Two lessons</u>

In chapter 2 I described three different proposals about moral explanatory claims that a cognitivist opponent of moral explanations can put forward in an attempt to accommodate our practice with regard to such claims. All of them are built on theses i)-iv), which represent the most natural way of fleshing out the Assumption and are available only to a cognitivist. In chapters 3 and 4 we looked at two different non-cognitivist attempts to mimic the accommodationist proposals available to a cognitivist. In chapter 3, we examined Hare's way of fleshing out the Assumption and the non-cognitivist accommodationist proposals built on it. I argued that those proposals fail. In chapter 4, we looked at Gibbard's alternative non-cognitivist reading of the Assumption and at the accommodationist proposal he builds on it. Our assessment of Gibbard's proposal has so far been positive: that non-cognitivist proposal, I showed, avoids all the problems raised in chapter 3. But, I'll argue in this chapter, there are other problems with Gibbard's proposal: it fails too, and Gibbard's attempt to mimic the cognitivist is, like Hare's, ultimately unsuccessful. In the next two sections, I'll raise two different problems for Gibbard's proposal. In order to press them, however, we first need to go back, in this section, to the cognitivist views I presented in chapter 2. Why?

Because there is a complication: I have so far bracketed it, but we now need to attend to it. The way to formulate the reading of the Assumption available to cognitivists is not quite by means of i)-iv), but in terms of a slightly different set of theses. In chapter 2, where I framed it by means of i)-iv), I simplified things a little: we now need to add the piece I then left out.

Analogously for the cognitivist proposals presented in that chapter. The need for these revisions comes out if we consider a certain kind of cases, which I have so far bracketed. So I'll start by going back, in this section, to the cognitivist: consideration of cases of a certain kind shows, I'll argue, that the cognitivist will want to flesh out the Assumption by means of a revised version of i)-iv); analogously for the cognitivist proposals presented in chapter 2. I'll then turn, in the next two sections, to Gibbard and ask whether he has the resources to mimic this revised reading of the Assumption, and these revised proposals, and deal with the relevant sort of cases. I'll argue that he does not. Indeed, the kind of cases we'll be looking at in this chapter points to two different problems for Gibbard, neither of which, I'll argue, he can solve.

Why did I not deal with the needed revisions right away, instead of postponing consideration of them until this chapter? Because doing so would have required addressing the issues raised by the kind of cases we'll be looking at in this chapter, unnecessarily complicating my exposition throughout the present work. So for ease of exposition, I bracketed consideration of those cases in the previous chapters and said all I had to say that could be said without looking at them. Having done that, we can now turn to the issues raised by those cases and the revisions they require.

Back then to the cognitivist package I presented in chapter 2. Remember its core ideas. First, there is the Assumption: actions, people, institutions, etc. owe their moral features to their non-moral properties; or –fleshing this out just a little- all the M-ness instantiated in the world ultimately stems, for a given moral property M-ness, from a certain set of non-moral properties: the M-making properties. Providing a few more details: the categorical moral truths in any possible world form a neat, structured set, with some of them –the brute ones- that are not made true by other truths and are (together with non-moral truths) what ultimately makes all the others

–the derived ones- true. Second, there is the idea that if you want to replace a moral sentence's apparent reference to a moral property with a reference to non-moral properties, the Assumption points to a plausible way of achieving your aim: you should 'go by the M-making properties'. If there is a job, apparently performed by M-ness, that you are not willing to assign to it, it should strike you as plausible to assign it to the ultimate source of all the M-ness instantiated in the world. The application of this general idea to the case of moral explanatory claims yields the accommodationist proposals available to a cognitivist opponent of moral explanations.

I now want to point out two lessons to be learnt from consideration of a certain kind of cases and show that, in order to take those lessons into account, both the cognitivist reading of the Assumption presented in chapter 2 and the cognitivist proposals built on it need to be revised. The first lesson is negative, in the following sense: it points to results that a cognitivist friend of the Assumption will *not* want her reading of the Assumption to commit her to. The second lesson is positive: it points to results that a cognitivist friend of the Assumption *will* want her reading of the Assumption to yield. For both lessons, the results in question are verdicts about cases of a certain kind: I'll define the relevant kind as we proceed, after offering some examples.

It will be helpful in what follows to keep in mind a feature of the Assumption that will by now be obvious, but that is worth mentioning explicitly. Suppose you are a cognitivist friend of the Assumption. Then you will think of substantive moral inquiry as aiming at finding out what the brute moral truths are. Categorical moral truths form a hierarchy with brute truths at the top, so substantive moral inquiry aims at figuring out what the brute truths are. These, you will think, are the principles pertaining to M-ness, for a given moral property M-ness; and developing a theory of M-ness (of goodness, of rightness, etc.) amounts, you will think, to figuring out what the brute moral truths are that ascribe M-ness. The Assumption, in other words, provides a more

detailed articulation and systematization of our intuitive notions of a moral principle and of a moral theory. What are moral principles? If you are a cognitivist friend of the Assumption, you have an answer to offer: they are brute moral truths. And what is a moral theory? It is an account of the brute moral truths there are. And why is such an account important? Because if we know about brute moral truths, we know about what ultimately makes all the other categorical moral truths true: there is more to the moral domain than brute moral truths, but everything else comes from what we know about, plus non-moral truths. Given this feature of the Assumption, if you are a friend of it you will then want your reading of the Assumption to yield plausible verdicts about what does, or does not, count as a moral principle and about what does, or does not, count as a moral theory distinct from a given one[1].

Since I'll eventually argue they need to be revised, another thing that it will be helpful to keep in mind is what theses i)-iv) say. So by way of reminder, here are i)-iv):

i)   All categorical moral truths are either <u>derived truths</u>, i.e. they are made true by other categorical moral truths (plus appropriate non-moral truths) or <u>brute truths</u>, i.e. they are not made true by other truths.

ii)   All derived moral truths are ultimately made true by brute moral truths (plus appropriate non-moral truths).

---

[1] Though in the text I illustrate the point only with regard to the cognitivist reading of the Assumption, analogous remarks apply also to the non-cognitivist readings we looked at in chapters 3 and 4: the Assumption, on *any* of its readings, provides an articulation and systematization of our intuitive notions of a moral principle and of a moral theory, with different readings offering different and competing answers to the question 'What is a moral principle?' and, correspondingly, 'What is a moral theory?'. In a nutshell, moral principles are: universal moral truths that ultimately make all the other moral truths true; universal moral beliefs from which all the other moral beliefs are ultimately inferred; universal moral beliefs from which all the other moral beliefs logically follow. Thinking of the different readings of the Assumption as offering different answers to the question 'What is a moral principle?', one feature worth pointing out is the similarity between the answers offered to this question (at least, the first and the third answer) and some of the standard types of answers offered in the philosophy of science / metaphysics literature to the question 'What is a law of nature?' (think, for instance, of the debate between D.M. Armstrong and D. Lewis). This is, as I say, a feature worth pointing out, though I won't have time to discuss it here.

iii)   All derived moral truths are ultimately made true in only one way, i.e. the relevant way.

iv)   All brute moral truths can be expressed by an appropriately strong universal moral sentence using a non-moral predicate denoting a non-moral property.

And now let us turn to the **first lesson**. According to iv), being a truth expressible by an appropriately strong universal moral sentence using a non-moral predicate denoting a non-moral property –or, as I will from here on say, being a U truth- is a necessary condition for being a brute moral truth. Thesis iv) does not say that being a U truth is also sufficient for being a brute truth. Would it be plausible to say so? No. Suppose ' ■ All N_1's are M' expresses a brute moral truth. Consider ' ■ Everything with the conjunctive property of being N_1&F is M' –or, for short, ' ■ All N_1&F's are M'-, where 'F' is a non-moral predicate denoting a non-moral property. It is true that ■ all N_1&F's are M; this, moreover, is a U truth. Is it a brute truth? I take it that, for many, many replacements for 'F', saying yes is counterintuitive and implausible. Suppose ' ■ All murders are wrong' expresses a brute moral truth. Is ' ■ All murders committed on a Wednesday are wrong' also a brute truth? Surely not. If we say that being a U truth is sufficient for being a brute truth, then the number of brute truths explodes, with a lot of them mentioning quite irrelevant properties. It would be a serious count against a reading of the Assumption if it yielded this result, so it is good news for cognitivist friends of the Assumption that iv) does not commit them to it.

So far, so good for theses i)-iv): no revisions needed yet. Strictly speaking, however, the cognitivist proposals presented in chapter 2 (be they about moral explanatory claims or categorical moral sentences) do need to be revised in order to accommodate the lesson that being

214

a U truth is not sufficient for being a brute truth. This lesson has an immediate semantic import, given that the idea behind the cognitivist proposals is that we can, when specifying the content (or payoff) of moral sentences, 'go by the M-making properties', i.e. replace a moral sentence's apparent reference to the property of being M with a reference to the M-making properties. If being a U truth is not sufficient for being a brute truth, then the properties to replace for being M, when specifying the content (or payoff) of a moral sentence, cannot be read off the U truths: being a property picked out by a U truth is not sufficient for being a property to 'go by' when doing the semantics of moral sentences. Suppose again that ' ■ All N_1's are M' expresses a brute moral truth. ' ■ All N_1&F's are M' is (for many, many replacements for 'F'), though a U truth, not a brute truth. When specifying the content (or payoff) of a moral sentence, being N_1 is therefore one of the properties to replace for being M, whereas being N_1&F is not.

Strictly speaking, however, given the way I framed the cognitivist proposals in chapter 2, being N_1&F does end up among the properties to 'go by' when doing the semantics of moral sentences. This is due to the fact that I defined the substitution operations relied on by the proposals in terms of the *n*'s, i.e. the non-moral properties that can be denoted by some non-moral predicate 'N' such that ■ all N's are M, or –what is the same- the non-moral properties that can be denoted by some non-moral predicate 'N' such that ' ■ All N's are M' expresses a *truth*, and of course being N_1&F is one such property. The M-making properties, however, those we are meant to replace for being M when specifying the content (or payoff) of a moral sentence, are only the non-moral properties that can be denoted by some non-moral predicate 'N' such that ' ■ All N's are M' expresses a *brute truth*. My formulation of the proposals is too permissive.

At least, it is too permissive if 'the *n*'s' is read, as in chapter 2 I invited you to do, as shorthand for 'the non-moral properties that can be denoted by some non-moral predicate 'N'

such that ■ all N's are M'. Since in chapter 2 I was bracketing consideration of cases like those just looked at, this reading of 'the *n*'s' kept things simpler while doing no harm, so I adopted it. Now that we are taking these cases into account, my original formulation of the proposals needs to be revised by having 'the *n*'s' be shorthand for 'the non-moral properties that can be denoted by some non-moral predicate 'N' such that ' ■ All N's are M' expresses a brute truth'. If 'the *n*'s' is read this way, my formulation of the proposals is not too permissive, and everything I said in the previous chapters remains true. From now on then, this is what 'the *n*'s' will be shorthand for. In section 3, I'll show that an argument against Gibbard's semantic proposals (indeed, against all the non-cognitivist semantic proposals we looked at in chapters 3 and 4) can be constructed once we pay attention to the need for this revision to the cognitivist proposals.

It is worth mentioning here (because I'll make use of use it in section 3) that an analog of our first lesson holds for theses i.H)-iv.H), which formulate Hare's reading of the Assumption, and for the non-cognitivist proposals built on them. By way of reminder, here are i.H)-iv.H):

i.H)   All categorical moral beliefs we have are either <u>inferred beliefs</u>, i.e. they are inferred from other categorical moral beliefs (plus appropriate non-moral beliefs) we have or <u>direct beliefs</u>, i.e. they are not inferred at all.

ii.H)   All inferred moral beliefs are ultimately inferred from direct moral beliefs (plus appropriate non-moral beliefs).

iii.H) All inferred moral beliefs are ultimately inferred in only one way, i.e. *via* subsumption.

iv.H)   All direct moral beliefs can be expressed by an appropriately strong universal moral sentence using a non-moral predicate denoting a non-moral property.

Cases like those we have just looked at –think again of '■ All murders are wrong' and '

■ All murders committed on a Wednesday are wrong'- show it would be counterintuitive and

implausible to say that being a belief expressible by an appropriately strong universal moral

sentence using a non-moral predicate denoting a non-moral property –for short, being a U belief-

is sufficient for being a direct belief. No problem here for iv.H), which says only that being a U

belief is necessary for being a direct belief. On the other hand, the non-cognitivist proposals built

on Hare's reading of the Assumption do need to be revised by having 'the n_S's' be shorthand, not

(as I did in chapter 3) for 'the non-moral properties that can be denoted by some non-moral

predicate 'N' such that S believes that ■ all N's are M' –or, what is the same, 'the non-moral

properties that can be denoted by some non-moral predicate 'N' such that ' ■ All N's are M'

expresses a *belief* held by S'-, but for 'the non-moral properties that can be denoted by some

non-moral predicate 'N' such that ' ■ All N's are M' expresses a *direct belief* held by S'. From

now on then, this is what 'the n_S's' will be shorthand for (once again, if 'the n_S's' is read this

way, everything I said in the previous chapters remains true).

The first lesson, I said above, points to verdicts that a cognitivist friend of the

Assumption will *not* want her reading of the Assumption to commit her to. The **second lesson**,

on the other hand, points to verdicts that a cognitivist friend of the Assumption *will* want her

reading of the Assumption to yield. Suppose again that ' ■ All murders are wrong' expresses a

brute moral truth. Cognitivist friends of the Assumption, we have just seen, will not want to say

that ' ■ All murders committed on a Wednesday are wrong' is a brute truth too, and theses i)-iv)

do not commit them to saying so. What they will want to say is that the latter is a derived truth.

Do i)-iv) yield this verdict? They do. It is true, we are supposing, that ■ all murders committed on a Wednesday are wrong. What is wrong about them? Well, that they are murders. In the terminology of i)-iv), what makes it true that ■ all murders committed on a Wednesday are wrong is that ■ they are all murders and ■ all murders are wrong. '■ All murders committed on a Wednesday are wrong' is thus a truth that is made true by other categorical moral truths (plus appropriate non-moral truths), ultimately by the brute truth that ■ all murders are wrong. Thesis i) yields that it is a derived truth, which is the result cognitivist friends of the Assumption will want.

We have considered the case in which '■ All murders are wrong' is a brute moral truth. This is just one example: analogous considerations apply to analogous cases. If '■ All N_1's are M' is a brute truth, then cognitivist friends of the Assumption will want to say that '■ All N_1&F's are M' is a derived truth (for many, many replacements for 'F'), and this is the result i)-iv) yield, given the definitions under i).

Let us use 'case #1' for the case involving the truths that ■ all murders are wrong and ■ all murders committed on a Wednesday are wrong. The cases we have been looking at so far – case #1 and those like it- are examples of the following *kind* of cases:

- t is a U truth to the effect that ■ all N_2's are M;

- t is made true by other categorical moral truths (plus appropriate non-moral truths), ultimately by a brute truth to the effect that ■ all N_1's are M, for some property N_1-ness such that ■ all N_2's are N_1

(case #1 and those like it are the cases of this kind in which N_2-ness is the conjunctive property of being $N_1\&F$, for many, many replacements for 'F'). It is cases of this kind that we'll be considering in this chapter. When a truth meets the two conditions above, I'll say that it <u>fits the description</u>. For example, in case #1 the truth that ■ all murders committed on a Wednesday are wrong fits the description[2].

Given the definitions under i), i)-iv) yield the result that in *all* cases of our kind, the truth that fits the description is a derived truth: thesis i) defines derived truths as those that are made true by other categorical moral truths (plus appropriate non-moral truths), and any truth that fits the description meets this definition. Case #1 and those analogous to it provide one class of examples in which this is the right result, the result cognitivist friends of the Assumption will want. Here is another such case, which I'll call case #2.

Suppose you and I are discussing what deserves respect. Human beings are said by many to deserve respect. Non-human animals too are often mentioned. Some say works of art deserve respect. Others that the environment does. Of course, treating a human being with the respect s/he deserves amounts to something different than does treating a work of art with the respect it

---

[2] If a truth fits the description, then (given the first condition) it is a universal moral truth and (given the second condition) there is a brute truth that ultimately makes it true. It is not built into the second condition that there be *exactly one* such brute truth: as we saw in 2.2, theses i)-iv) allow for the possibility that there be, for one categorical moral truth, multiple brute truths each of which ultimately makes the truth in question true, so the second condition only requires that there be *a* brute truth that ultimately makes the truth fitting the description true. As we also saw in 2.2, i)-iv) allow for the possibility that a universal moral truth be ultimately made true, not by a *single* brute truth, but by a *series* of different brute truths, so the definition of the kind of cases I want to draw attention to in this chapter should, strictly speaking, be broadened to include the latter possibility as well. That is, the second condition should be: t is made true by other categorical moral truths (plus appropriate non-moral truths), ultimately by either (*a*) a brute truth to the effect that ■ all N_1's are M, for some property N_1-ness such that ■ all N_2's are N_1 or (*b*) a series of different brute truths to the effect that ■ all N_1's are M, ■ all N_3's are M, …, where N_1-ness, N_3-ness, … are properties such that ■ each N_2 is either N_1 or N_3 or … . In what follows, however, I'll rely on the definition in the text, since it is enough to make all the points I want to make about the kind of cases in question. Since the second condition cannot be met unless ■ all N_2's are N_1, I should remind the reader that in 2.2 I left theses i)-iv) neutral as to the degree of strength expressed by ' ■ ', assuming only (fairly uncontroversially) that it is not greater than that of logical truths. It follows that the condition that ■ all N_2's are N_1 is met, provided that 'All N_2's are N_1' expresses a logical truth (as it does in case #1 and those like it, in which being N_2 = being $N_1\&F$). And of course the same holds, if 'All N_2's are N_1' expresses a necessary truth. Since the cases we'll be looking at in what follows fall under one or the other category, the condition that ■ all N_2's are N_1 will be met in all of them.

deserves; still, it makes sense to say, of your little child, my dog and the Mona Lisa, that it deserves respect, and we are trying to figure out which of these things, if any, is true. Suppose I propose the following theory: there is only one sort of things that deserve respect, namely sentient beings. There is only one principle when it comes to what deserves respect: that sentient beings deserve respect. Human beings, I agree with what many say, do deserve respect. But they do because they are sentient: they deserve respect *qua* sentient beings. Nothing special about us, says my theory: what is deserving of respect about us is that we are sentient beings, i.e. something we share with other, non-human beings.

On my theory, cognitivist friends of the Assumption will want to say, there is only one brute truth, namely ' ■ All sentient beings deserve respect'. ' ■ All human beings deserve respect' is, though a truth, a derived truth, they will want to say. And, once again, this is the verdict i)-iv) yield. In the terminology of i)-iv), ' ■ All human beings deserve respect' is a U truth that is made true by other categorical moral truths (plus appropriate non-moral truths), ultimately by the brute truth that ■ all sentient beings deserve respect. That is, it is a truth that fits the description, so i)-iv) yield that it is a derived truth.

So far, so good for theses i)-iv). However, the general claim they yield that in *all* cases of the kind defined above, the truth that fits the description is a derived truth is, I now want to suggest, implausible: there are cases in which a U truth is made true by other categorical moral truths (plus appropriate non-moral truths), but cognitivist friends of the Assumption will nonetheless want to count it as a brute truth. Hence, the definitions under i) must be revised.

It is easy enough to see what my next example, case #3, will be. Suppose that, while discussing with me what deserves respect, you propose the following theory: there are two sorts of things that deserve respect, two principles when it comes to what deserves respect, i.e. that

sentient beings deserve respect and that so do human beings. Human beings, you agree with me and many others, deserve respect. But they do for *two* reasons: because they are sentient and because they are human. There *is* something special about us, says your theory: two things are deserving of respect about us, and while one is something we share with other, non-human beings, the other is the very fact that we are human beings.

On your theory, cognitivist friends of the Assumption will want to say, there are two brute truths: ' ■ All sentient beings deserve respect' and ' ■ All human beings deserve respect'. But note: given the definitions under i), i)-iv) do *not* allow them to say so. This is due to the fact that, no matter what else they might want to say about ' ■ All human beings deserve respect', one thing they cannot avoid saying is that it is made true by other categorical moral truths (plus appropriate non-moral truths). Hence, if they stick to the definitions under i), ' ■ All human beings deserve respect' cannot count as a brute truth: it must count as derived.

The point is worth stressing. Consider human beings: they deserve respect, you and I both think. What about them is deserving of respect? On your theory, at least one answer to this question is that they are sentient beings. Just as, on my theory, human beings deserve respect because they are sentient, so too, on your theory, human beings deserve respect because they are sentient. On your theory, this is not the whole story about them, but it is certainly part of the story about them. In the terminology of i)-iv): just as, on my theory, ' ■ All human beings deserve respect' is a U truth that is made true by other categorical moral truths (plus appropriate non-moral truths), ultimately by the brute truth that ■ all sentient beings deserve respect, so too, on your theory, ' ■ All human beings deserve respect' is a U truth that is made true by other categorical moral truths (plus appropriate non-moral truths), ultimately by the brute truth that ■

221

all sentient beings deserve respect. On your theory, this is not the whole story about ' ■ All human beings deserve respect', but it is certainly part of the story about it. And this is enough to make ' ■ All human beings deserve respect', even on your theory, a truth that fits the description. The definitions under i) thus yield that it is a derived, not a brute, truth, which is not the verdict cognitivist friends of the Assumption will want. The definitions under i) must be revised.

Theses i)-iv) fail to yield, about case #3, the verdict cognitivist friends of the Assumption will want. Here is a slightly different way of pointing out the problem with i)-iv). When you and I discuss what deserves respect, the theory you put forward looks like a different theory from the one I put forward; that is, case #3 looks different from case #2. Cognitivist friends of the Assumption should be able to draw this distinction, and so to say that ' ■ All human beings deserve respect' is a derived truth in one case but a brute truth in the other. Theses i)-iv), however, do not allow them to say this: ' ■ All human beings deserve respect' is a truth that fits the description in both cases, so i) yields that it is a derived truth in both cases. Theses i)-iv) thus do not allow cognitivist friends of the Assumption to distinguish between case #3 and case #2, between the theory you put forward and the one I put forward.

Case #3 and case #2 provide just one instance of this problem. Here are a couple of other instances, borrowed directly from the philosophical literature, in which i)-iv) are blind to a distinction that cognitivist friends of the Assumption will want to be able to draw. Consider the standard criticism made against J.S. Mill's *Utilitarianism* of endorsing both of two incompatible theories of (intrinsic) goodness. On the one hand, Mill endorses hedonism, according to which there is only one sort of things that are good, only one principle about goodness, i.e. that pleasures are good. On the other, he famously puts forward his doctrine of higher pleasures, according to which "some *kinds* of pleasure are more desirable and more valuable than others"

222

(ch. 2, emphasis in the original), namely those produced by engagement in the activities and pursuits that exercise our higher faculties: think, for example, of writing a novel, pursuing a scientific discovery, or leading a life characterized by a significant amount of self-examination. But, so goes one way of pressing the charge against Mill, if some kinds of pleasure are better than others, then there must be something else that is good about the former, in addition to their being *pleasures*, namely their being pleasures of that kind: their being *higher pleasures*. Higher pleasures, so must Mill think, are good for two reasons: not only (as any hedonist would say) because they are pleasures, but also because they are higher pleasures. That is, the doctrine of higher pleasures commits Mill to a *further* principle about goodness, i.e. that higher pleasures are good, in addition to the principle that pleasures are good, thus contradicting the hedonist claim that the latter is the *only* principle about goodness[3].

Without taking a stand on whether Mill's views are actually inconsistent, here at any rate are two theories of goodness discussed in the literature: the hedonist theory and the theory that, according to this criticism, Mill is committed to by the doctrine of higher pleasures. Cognitivist friends of the Assumption will want to be able to distinguish between them. On both theories, they will want to say, ' ■ All pleasures are good' is a brute truth; the two theories differ in that, on the first of them, ' ■ All higher pleasures are good' is a derived truth, while on the second, it is a brute truth. But i)-iv) do not allow them to draw this distinction, since in both cases ' ■ All higher pleasures are good' is a truth that fits the description: it is a U truth that is made true by

---

[3] For a recent discussion of this criticism of Mill, see Sturgeon [2010] (defending Mill from the charge of inconsistency) and Brink [2013], ch. 3 (pressing the charge). My aim in the text is not to discuss the charge, so I pass over various complications mentioned in the relevant literature. In particular, I focus on versions of the charge that grant that 'pleasures' refers, in both 'Pleasures are good' and 'Higher pleasures are good', to mental states, so that what the latter principle talks about is, quite straightforwardly, a subset of what the former does (on other ways of pressing the charge –Brink is a case in point-, this is not so: according to some writers, when he talks about higher pleasures Mill has in mind, not the mental states produced by engagement in certain activities and pursuits, but the activities and pursuits themselves).

other categorical moral truths (plus appropriate non-moral truths), ultimately by the brute truth that ■ all pleasures are good. Thesis i) thus yields that in both cases it is a derived truth.

Or consider G.E. Moore's discussion in *Principia Ethica*. The main target of the substantive (as opposed to meta-ethical) discussion in *Principia* is hedonism. Moore rejects the hedonist claim that there is only one sort of things that are good, arguing instead that there are many different sorts of good things. Chief among them, he famously claims, are "certain states of consciousness, which may be roughly described as the pleasures of human intercourse and the enjoyment of beautiful objects" (§113). Suppose we are convinced by Moore on this score and expand the hedonist list of good things by adding these two items. The resulting theory has three principles: pleasures are good; the pleasures of human intercourse are good; enjoyments of beautiful objects are good. Note that the hedonist will agree with us that the pleasures of human intercourse are good; in her view, however, what is good about them is only that they are pleasures: they are good only *qua* pleasures, she will claim. On our theory, on the other hand, the pleasures of human intercourse are good for two reasons: because they are pleasures and because they are pleasures of human intercourse. Presumably, any enjoyment is a pleasure, so the same holds for enjoyments of beautiful objects.

Is our Moore-inspired theory Moore's own theory? Moore, as I say, rejects the hedonist claim that there is only one principle about goodness. Does he also reject the hedonist principle that pleasures are good? Chapter 3 of *Principia*, which is where most of Moore's critical discussion of hedonism occurs, is not entirely explicit on this point: here Moore is mainly focused on arguing that there are other principles about goodness, in addition to the hedonist principle, and one is left wondering whether he also means to reject the principle that pleasures are good, especially since his arguments against the claim that this is the only principle do not

seem to commit him to denying that this is at least *a* principle about goodness. However, in chapter 6, where he lays out his positive views about goodness, Moore is explicit (§127) that he rejects the principle that pleasures are good. There is a proviso, though. Already in chapter 3 (§§52-53), Moore distinguishes between pleasures we are conscious of and pleasures we are unconscious of, granting that the hedonist may have only the former in mind when she claims that pleasures are good; and in the passage of chapter 6 just mentioned, Moore is equally explicit that, if by 'pleasures' we mean only pleasures we are conscious of, then he accepts the principle that pleasures are good (though only slightly so, if compared to the other sorts of good things there are). If this is what we should mean by 'pleasures' in a discussion of hedonism, then our Moore-inspired theory is Moore's own theory of goodness[4].

Whether or not our Moore-inspired theory is Moore's own theory, here at any rate is another theory of goodness present in the literature. Cognitivist friends of the Assumption will want to be able to distinguish between the hedonist theory and this theory too. While ' ■ All pleasures are good' is a brute truth on both theories, they will want to say, ' ■ All the pleasures of human intercourse are good' is a derived truth on the first of them, but a brute truth on the second. Once again, i)-iv) do not allow them to draw this distinction: ' ■ All the pleasures of human intercourse are good' is in both cases a U truth that is made true by other categorical moral truths (plus appropriate non-moral truths), ultimately by the brute truth that ■ all pleasures are good; that is, it is in both cases a truth that fits the description. Thesis i) thus yields the same verdict in both cases: it is a derived truth. Analogously for ' ■ All enjoyments of beautiful objects are good'.

---

[4] Provided, of course, that we do not take our list of good things to be exhaustive: in Moore's view there are, as I said, many different sorts of good things, and the three items on our list do not cover them all.

We have looked at the two cases I called case #2 and case #3, and at a couple of other analogous pairs; it is easy to generalize from them. For any pair of such cases, i)-iv) are unable to draw the relevant distinction. And it is a serious count against a reading of the Assumption that it yields the result that, in each of these pairs of cases, we do not have two different moral theories involved, but one and the same theory. I take it as a *datum* that you and I, in our discussion about what deserves respect, put forward different moral theories; that someone endorsing the doctrine of higher pleasures (on the reading of it mentioned above) and the hedonist also have different moral theories; and that so do Moore (or at least someone endorsing our Moore-inspired theory) and the hedonist. The relevant parties offer, in the examples above, different accounts of the principles pertaining to the relevant moral property, different answers to the question 'What are the moral principles pertaining to M-ness?', for the relevant moral property M-ness. It is this *datum* that explains why it makes sense to say that you and I disagree with one another; why it makes sense, for someone reading the doctrine of higher pleasures along the lines mentioned above, to charge Mill with inconsistency; and why it makes sense to say that Moore (or someone endorsing our Moore-inspired theory) and the hedonist also disagree with one another.

For each pair of cases, the two relevant theories yield the same categorical moral truths, fair enough; still, they offer different stories about what accounts for them. To illustrate again with one of the examples above: you and I agree that, say, Marta deserves respect. On my theory, however, she does only because she is a sentient being: nothing special about us human beings, in my view. On your theory, on the other hand, she does for two reasons: because she is sentient and because she is human; there is something special about us human beings, according to you. And it would be especially hard for a friend of the Assumption to suggest that this amounts to no real difference: if you think things owe their moral features to their non-moral properties, you

226

had better make room for differences even among people who agree about what moral features things have, namely differences about what non-moral properties things owe those moral features *to*. Cognitivist friends of the Assumption will therefore not want to formulate their reading of the Assumption by means of i)-iv): those theses need to be revised.

When I introduced i)-iv) in section 2.2, I remarked that the Assumption is meant to capture an aspect of the very nature of morality. As applied to the cognitivist reading of the Assumption, the idea, I pointed out, is that the nature of moral properties is such that categorical moral truths (if there are any) *cannot* but be arranged in a certain neat, orderly way. Alternatively put, we can think of the cognitivist reading of the Assumption as a constraint over possible worlds. Thinking of it this way, the nature of the problem with i)-iv) can be described as follows: i)-iv), we saw above, flesh out the Assumption so as to make room only for possible worlds in which, if a truth fits the description, it is derived; cognitivist friends of the Assumption will instead want to flesh it out so as to make room both for possible worlds in which a truth fits the description and is derived (case #1 and those analogous to it; case #2 and those analogous to it) and for ones in which a truth –in some instances, even the same truth- fits the description and is brute (case #3 and those analogous to it)[5].

How can they do that? The problem, we saw above, stems from the definitions offered by thesis i). The lesson from case #3 and those analogous to it is that some truths fitting the description, i.e. some U truths that are made true by other categorical moral truths (plus

---

[5] Consider again case #1 and those analogous to it. Though it is implausible to say, of the actual world or a world similar to it in which ‘ ■ All murders are wrong’ is a brute truth, that ‘ ■ All murders committed on a Wednesday are wrong’ is also a brute truth in that world, it is, I think, a question worth asking whether there are possible worlds – which would, of course, be *extremely* different from any like the actual one- in which ‘ ■ All murders are wrong’ and ‘ ■ All murders committed on a Wednesday are wrong’ are both brute truths; that is, whether there is a case that mirrors case #1 in the same way in which case #3 -and those like it- mirror case #2 -and those like it-. Analogous questions are worth asking (and perhaps have, at least in some instances, a better chance of getting a positive answer) about the cases analogous to case #1.

appropriate non-moral truths), are brute truths. In order to accommodate this lesson, cognitivist friends of the Assumption will want to revise

i)   All categorical moral truths are either <u>derived truths</u>, i.e. they are made true by other categorical moral truths (plus appropriate non-moral truths) or <u>brute truths</u>, i.e. they are not made true by other truths.

And how can they revise those definitions?

I have already hinted at the answer. Go back to case #2 and case #3, i.e. to the two theories we put forward in our discussion about what deserves respect. There is, I pointed out above, one aspect under which case #2 and case #3 are certainly alike: just as, on my theory, human beings deserve respect because they are sentient, so too, on your theory, human beings deserve respect because they are sentient. However, there is also another aspect under which the two cases are certainly different: as I put it above, whereas on my theory this is the whole story about them, on your theory this is not the whole story about human beings and why they deserve respect. Human beings deserve respect only because they are sentient, says my theory; human beings deserve respect not only because they are sentient, says your theory, but also because they are human.

In the terminology of cognitivist friends of the Assumption: on both theories, ' ■ All human beings deserve respect' is made true by other categorical moral truths (plus appropriate non-moral truths), ultimately by the brute truth that ■ all sentient beings deserve respect. However, on my theory this is the whole story about it: ' ■ All human beings deserve respect' is true *only* because it is made true by other categorical moral truths (plus appropriate non-moral

truths). On your theory, on the other hand, this is not the whole story about it: ' ■ All human beings deserve respect' is *not* true only because it is made true by other categorical moral truths (plus appropriate non-moral truths).

This, I suggest, is the difference to fasten on when defining 'derived truth' and 'brute truth': the difference between being, and not being, true only because made true by other categorical moral truths. The reason why, in case #2, ' ■ All human beings deserve respect' is a derived truth is not just, as i) would have it, that it is made true by other categorical moral truths, but that it is true *only* because it is made true by other categorical moral truths. A truth can be made true by other categorical moral truths and still be a brute truth, provided it is *not* true only because it is made true by other categorical moral truths: this is what happens with ' ■ All human beings deserve respect' in case #3. Analogous considerations apply to the other pairs of analogous cases we looked at above.

The same holds for case #1. Suppose again that ' ■ All murders are wrong' is a brute truth and consider ' ■ All murders committed on a Wednesday are wrong'. The reason why the latter is a derived truth is not just that it is made true by other categorical moral truths, ultimately by the former. It is rather, I suggest, that it is true *only* because it is made true by other categorical moral truths. Analogously for ' ■ All N_1&F's are M', for many, many replacements for 'F', supposing that ' ■ All N_1's are M' is a brute truth[6].

Here then is how I suggest cognitivist friends of the Assumption should revise thesis i):

---

[6] Consider (if you can picture it) a world in which ' ■ All murders committed on a Wednesday are wrong' is made true by other categorical moral truths, but it is *not* true only because it is made true by other categorical moral truths. *If* there is such a world, it is one in which ' ■ All murders committed on a Wednesday are wrong' is a brute truth (and, of course, it is also a world extremely different from any like the actual one).

I)   All categorical moral truths are either <u>derived truths</u>, i.e. they are true only because made true by other categorical moral truths (plus appropriate non-moral truths) or <u>brute truths</u>, i.e. they are (*a*) not made true by other truths or (*b*) made true by other categorical moral truths (plus appropriate non-moral truths) but not true only because made true by other categorical moral truths (plus appropriate non-moral truths).

Once thesis i) is revised this way, theses ii)-iv) also need to be revised along the same lines:

II)   All categorical moral truths that are made true by other categorical moral truths (plus appropriate non-moral truths) are ultimately made true by brute moral truths (plus appropriate non-moral truths).

III)   All categorical moral truths that are made true by other categorical moral truths (plus appropriate non-moral truths) are ultimately made true in only one way, i.e. the relevant way.

IV)   All brute moral truths can be expressed by an appropriately strong universal moral sentence using a non-moral predicate denoting a non-moral property[7].

---

[7] How I use the phrase 'being ultimately made true by' also needs to be revised. So far, I have been using it as follows: a moral truth m is ultimately made true by a moral truth m', plus an appropriate non-moral truth n, iff m is made true by m', plus n, and m' is not made true by other truths. The constraint on m' needs to be weakened: a moral truth m is ultimately made true by a moral truth m', plus an appropriate non-moral truth n, iff m is made true by m', plus n, and m' *either* is not made true by other truths *or* is made true by other moral truths, plus appropriate non-moral truths, but it is not true only because made true by other moral truths, plus appropriate non-moral truths (on the other hand, it remains true that no such constraint is imposed on n: see ch. 2, n. 6). So, for example, on the theory you put forward when we discuss what deserves respect, 'Marta deserves respect' counts as ultimately made true both by ' ■ All sentient beings deserve respect', plus 'Marta is a sentient being', and by ' ■ All human beings deserve respect', plus 'Marta is a human being'. The following difference between i)-iv) and I)-IV) is worth pointing out. Consider a categorical moral truth that is made true by other categorical moral truths (plus appropriate non-

To recap: we looked at a certain kind of cases, namely those involving a truth that fits the description, i.e. a U truth that is made true by other categorical moral truths (plus appropriate non-moral truths). Cognitivist friends of the Assumption will want to say that in some cases of this kind (case #1 and those analogous to it; case #2 and those analogous to it), the truth that fits the description is a derived truth, while in others (case #3 and those analogous to it), it is a brute truth. Given the definitions under i), theses i)-iv) yield that in all cases of this kind, the truth that fits the description is a derived truth. The definitions offered by I), on the other hand, allow theses I)-IV) to yield the right results, the results cognitivist friends of the Assumption will want, about all the cases we looked at.

In addition to accommodating the lesson that some truths fitting the description are brute truths, I)-IV) accommodate just as well as i)-iv) do the other lesson I pointed out earlier in this section, namely that being a U truth is not sufficient for being a brute truth, given that thesis IV) is identical to thesis iv). Cognitivist friends of the Assumption will therefore want to formulate their reading of the Assumption by means, not of i)-iv), but of I)-IV). From now on, it will be the latter set of theses I'll refer to as the reading of the Assumption available to cognitivists.

---

moral truths). Unlike i)-iv), I)-IV) allow for the possibility that the path that can be tracked, starting from such a truth and asking 'What other categorical moral truths make it true?', never reaches an end. How so? Given II), that path sooner or later reaches a brute truth. Given I), this brute truth may be a brute truth of type (*a*); if it is, then the path does reach an end. But it may also be a brute truth of type (*b*), in which case the path continues. Thesis II) applies again, guaranteeing that sooner or later the path, once more, reaches a brute truth, and the same considerations apply again. That is, II) guarantees two things: that the path that starts from the original categorical moral truth sooner or later reaches a brute truth; and that every time that path continues after reaching a brute truth, sooner or later it, once more, reaches a brute truth. But II) does allow for the possibility that the path never reaches a brute truth of type (*a*), hence an end. For certain purposes, this is an important difference between i)-iv) and I)-IV). For our purposes, however, this difference does not matter, so even if there are other ways of revising ii)-iv) that do not allow for that possibility, I'll bracket them in what follows.

Specifically, it is the latter set of theses that the cognitivist proposals about the content (or payoff) of moral sentences should be taken as built on[8].

In this section, we have taken a second look at the cognitivist views I presented in chapter 2. Consideration of cases of a certain kind taught us two lessons, which required revising, respectively, the cognitivist proposals about the content (or payoff) of moral sentences and the cognitivist reading of the Assumption. In the next two sections, we'll turn to Gibbard and ask whether he has the resources to accommodate those two lessons and deal with the relevant kind of cases. I'll argue that he does not.

In my criticism of Gibbard, I'll reverse the order I have followed in this section. In section 2, I'll start from the second lesson and argue that Gibbard lacks the resources to put forward an adequate reading of the Assumption, one that yields the right results about all the cases of the kind I have drawn attention to. Gibbard's proposals about the content of moral sentences are built on his reading of the Assumption, so this is one reason to reject those proposals. In section 3, I'll turn to the first lesson and argue there is a further problem with Gibbard's proposals, one that arises even if one grants that Gibbard has an adequate reading of the Assumption to build those proposals on.

---

[8] For completeness, I should point out that an analog of our second lesson holds for theses i.H)-iv.H), which formulate Hare's reading of the Assumption. In order to distinguish between case #2 (and those analogous to it) and case #3 (and those analogous to it), the definitions of 'inferred belief' and 'direct belief' offered by i.H) need to be revised, drawing a distinction, within the categorical moral beliefs that are inferred from other categorical moral beliefs (plus appropriate non-moral beliefs), between those that are, and those that are not, held only because inferred from other categorical moral beliefs (plus appropriate non-moral beliefs). Theses ii.H)-iv.H) (and my use of the phrase 'being ultimately inferred from') also need to be revised in ways that parallel the revisions just made to ii)-iv) (and my use of the phrase 'being ultimately made true by'). Though able to draw the relevant distinctions when applied to the cases we have looked at in this section, the revised theses are still vulnerable to the objections I raised against i.H)-iv.H) in chapter 3. As I say, I mention the need for this revision only for completeness: unlike the revision -pointed out earlier in this section- to the non-cognitivist proposals built on Hare's reading of the Assumption, I won't get back to this revision to i.H)-iv.H) in my arguments to follow.

2. **First objection to the 'Go by the hyperstates' strategy**

The second of the two lessons from the previous section, I noted above, points to results that a cognitivist friend of the Assumption will want her reading of the Assumption to yield, the results in question being verdicts about cases of the kind I drew attention to. Though in section 1 I was focusing on cognitivist friends of the Assumption, the point is of course entirely general: *any* adequate reading of the Assumption, be it cognitivist or non-cognitivist, has to yield the right results about our kind of cases. In the previous section, I defined the relevant kind of cases and formulated the verdicts about them in cognitivist terms: the cases were those involving a *truth* that fits the description, the verdicts were that certain truths fitting the description are *brute* or *derived* truths. In this section, I'll phrase things in non-cognitivist terms, specifically in the terms of a non-cognitivist who wants to take inspiration from Gibbard, and ask whether thesis (G), which formulates Gibbard's reading of the Assumption, yields, or can be revised so as to yield, the right verdicts about our kind of cases.

By way of reminder, here is thesis (G) (S is a complete belief set and S_{cm} is the set of categorical moral beliefs included in S):

(G)   If S is consistent then, provided S_{cm} is not empty, there is a non-empty subset of S_{cm} –let it be the <u>axiomatic set</u>- such that

i.G)   all the beliefs in S_{cm} that are not in the subset –let their set be the <u>generated set</u>- follow from beliefs in the subset (plus appropriate non-moral beliefs in S) by subsumption;

ii.G) all the beliefs in the subset can be expressed by an appropriately strong

universal moral sentence using a non-moral predicate denoting a non-

moral property.

In order to define the relevant kind of cases in terms applicable to thesis (G), let me

remind the reader that I am using 'being a U belief' as shorthand for 'being a belief expressible

by an appropriately strong universal moral sentence using a non-moral predicate denoting a non-

moral property', and let S* be a complete belief set that is consistent. I'll say that a belief b <u>fits</u>

<u>the description</u> when b meets the following two conditions:

- b is a U belief in S* to the effect that ■ all N_2's are M;

- b follows by subsumption from another U belief in S* to the effect that ■ all N_1's are

M, for some property N_1-ness such that S* includes the belief that ■ all N_2's are N_1.

Intuitively, a belief that fits the description is a U belief, in a complete and consistent belief set,

that can be subsumed under another U belief in the same belief set. We'll look at cases involving

a belief that fits the description and ask whether (G) yields, or can be revised so as to yield, the

right verdicts about them, the verdicts Gibbard will want [9].

---

[9] If a belief b fits the description, then (given the first condition) b is a universal moral belief in S* and (given the second condition) there is a U belief in S*, different from b, from which b follows by subsumption. It is not built into the second condition that there be *exactly one* such U belief in S*: as we saw in 4.1, thesis (G) allows for the possibility that there be, for one categorical moral belief in the generated set of a complete and consistent belief set, multiple beliefs in the axiomatic set such that the belief in question follows by subsumption from each of them, so the second condition only requires that there be *a* U belief in S*, different from b, from which b follows by subsumption. As we also saw in 4.1, (G) allows for the possibility that a universal moral belief in the generated set of a complete and consistent belief set follow by subsumption, not from a *single* belief in the axiomatic set, but from a *series* of different beliefs in the axiomatic set, so the same remark I made when discussing the cognitivist in section 1 (see n. 2) applies here as well: the definition of the kind of cases I want to draw attention to should, strictly

To illustrate, consider case #1. Suppose a user of moral language holds the principle that murders are wrong. Thesis (G) is about complete belief sets that are consistent and, we saw in chapter 4, one way to think of a belief set with these features is as capturing the views of a user of moral language who has become fully decided, with the beliefs in the axiomatic set capturing the general principles held by such a speaker. Our user of moral language need not be fully decided, so her views are modeled, not by a single complete belief set that is consistent, but by a set of such belief sets: those whose axiomatic set contains the belief that ■ all murders are wrong; or, more accurately, those whose axiomatic set contains the belief that ■ all murders are wrong and that include the other beliefs held by our speaker. Suppose that among those other beliefs is (unsurprisingly) the belief that ■ all murders committed on a Wednesday are murders, and so also the belief that ■ all murders committed on a Wednesday are wrong. If S* is one of the complete belief sets that are consistent modeling the views of our speaker, S* then includes both the belief that ■ all murders are wrong and the belief that ■ all murders committed on a Wednesday are wrong; and, we have said, the former belief belongs to its axiomatic set. Does the latter belief also belong to the axiomatic set of S*? I take it that saying yes is counterintuitive and implausible, and Gibbard will not want to say it. If we say it, the number of beliefs in the axiomatic set of S* explodes, with a lot of them mentioning quite irrelevant properties (remember what we said about the cognitivist take on case #1, when we asked whether, supposing that ' ■ All murders are wrong' expresses a brute truth, it is plausible to say that ' ■

---

speaking, be broadened. The second condition should be: b follows by subsumption from either (*a*) another U belief in S* to the effect that ■ all N_1's are M, for some property N_1-ness such that S* includes the belief that ■ all N_2's are N_1 or (*b*) a series of different U beliefs in S*, each of which differs from b, to the effect that ■ all N_1's are M, ■ all N_3's are M, …, where N_1-ness, N_3-ness, … are properties such that S* includes the belief that ■ each N_2 is either N_1 or N_3 or … . However, just as I did in the previous section, in what follows I'll rely on the definition in the text, since it is enough to make all the points I want to make with regard to Gibbard.

All murders committed on a Wednesday are wrong' also expresses a brute truth). The belief that

∎ all murders committed on a Wednesday are wrong belongs, Gibbard will want to say, to the

generated set of S*. Does (G) yield this result[10]?

Case #1 is just one example: analogous considerations apply to analogous cases. If the

belief that ∎ all N_1's are M belongs to the axiomatic set of S* -where S* is one of the complete

belief sets that are consistent modeling the views of a user of moral language-, then for many,

many replacements for 'F', Gibbard will want to say, the belief that ∎ all N_1&F's are M belongs

to the generated set of S*. Does (G) yield this result?

---

[10] One of the lessons I drew in the previous section from case #1 is that being a U truth is not sufficient for being a brute truth, as shown by the truth that ∎ all murders committed on a Wednesday are wrong. I should mention here that an analog of that lesson holds for thesis (G): being a U belief included in a complete belief set that is consistent is not sufficient for belonging to the axiomatic set of that belief set, as shown by the belief that ∎ all murders committed on a Wednesday are wrong. No problem here for (G): sub-thesis ii.G) says only that being a U belief is a necessary, not also a sufficient, condition for belonging to the axiomatic set of a complete belief set that is consistent. In section 1, I assumed it is implausible to say, of the actual world or a world similar to it in which ' ∎ All murders are wrong' is a brute truth, that ' ∎ All murders committed on a Wednesday are wrong' is also a brute truth in that world, while leaving it open (see n. 5) whether there are possible worlds –extremely different from any like the actual one- in which ' ∎ All murders are wrong' and ' ∎ All murders committed on a Wednesday are wrong' are both brute truths. Analogously, what I am assuming here is that it is implausible to say, in the case of many, many speakers with views modeled by complete and consistent belief sets whose axiomatic set includes the belief that ∎ all murders are wrong, that the belief that ∎ all murders committed on a Wednesday are wrong is also included in the axiomatic set in question. I am leaving it open whether there might be speakers with views modeled by complete and consistent belief sets whose axiomatic set includes both the belief that ∎ all murders are wrong and the belief that ∎ all murders committed on a Wednesday are wrong. Writers in the non-cognitivist tradition tend to think that it is not, in large part, built into the meaning of moral predicates what sort of considerations do or do not matter morally, so that, when it comes to the content of moral principles, the logic of moral predicates imposes few constraints: for example, a proposed principle to the effect that murders committed on a Wednesday are wrong does not, these writers tend to think, violate the logic of moral predicates. The proposed principle is eccentric, they grant, but the problem with it is that it is an *implausible* moral principle, not that it does not count as a *moral* principle: that is, it is a substantive, not a logical, problem. Gibbard might therefore want to claim that, among the complete belief sets that are consistent, are ones whose axiomatic set includes both the belief that ∎ all murders are wrong and the belief that ∎ all murders committed on a Wednesday are wrong, hence that there might be speakers with views modeled by such belief sets. A speaker with such views, Gibbard might want to claim, has a consistent, though implausible, position. What I am assuming in the text is compatible with such claims. I am assuming only (what is granted, indeed stressed, by non-cognitivists) that it is implausible to say, in the case of many, many speakers, that they hold such views: eccentric views are rare. Alternatively put, my claim in the text is about a speaker of the ordinary, non-eccentric variety, not about *every* possible speaker (just as my claim in section 1 was about the actual world and worlds similar to it, not about every possible world). For brevity's sake, I'll omit this qualification throughout my discussion, taking it as understood.

Well, what result does (G) yield about *any* belief fitting the description? A preliminary problem with (G) is that it yields *no* result at all about such beliefs. For any belief b fitting the description, included in a complete and consistent belief set S*, (G) yields neither that b belongs to the axiomatic set of S* nor that it belongs to the generated set of S*: (G) leaves it indeterminate which set b belongs to. This is because, for any complete and consistent belief set S*, (G) leaves it indeterminate *what* set is the axiomatic set of S*.

How so? Let us take a closer look at (G). Thesis (G) guarantees that, if a complete belief set is consistent and includes categorical moral beliefs, there is *a* set with the features under i.G) and ii.G). It does not, however, guarantee that there is *exactly one* such set. And as a matter of fact if there is one such set, then there is more than one, and (G) leaves it indeterminate which, among these different sets, is *the* axiomatic set.

I'll present the intuitive gist of the argument for this conclusion in the text, laying out the argument in a more formal way in a footnote at the end. Here is the intuitive idea. Let S* be a complete belief set that is consistent and assume S* includes categorical moral beliefs. Suppose we want to make a list of the beliefs that belong to the axiomatic set of S*. Given how (G) defines the axiomatic set, we need to look for those, among the U beliefs in S*, under which all the other categorical moral beliefs in S* can (given the non-moral beliefs in S*) be subsumed.

Now, among the categorical moral beliefs that can be subsumed under a U belief are, not only individual moral beliefs or universal moral beliefs that are not strong, but also other U beliefs. A belief that fits the description is precisely a belief like this: a U belief in S* that can be subsumed under another U belief in S*. But note: when making our list, it does not matter whether, given a belief b like this, i.e. given a U belief in S* that can be subsumed under another U belief in S*, we add b to our list *or* we leave b out of our list and add the subsuming belief,

237

call it b', instead. If our aim is to list those, among the U beliefs in S*, under which all the other categorical moral beliefs in S* can be subsumed, it can do no harm if we add b to the list: b is, after all, a U belief in S*; nor can it do any harm if we leave b out of the list and add b' instead: anything that can be subsumed under b can, after all, be subsumed under b' as well. But of course, depending on which way we go, we end up with a different list: specifically, one that does, or does not, include b.

We therefore have many different ways of constructing our list: there are many different sets with the features under i.G) and ii.G). Which among them is the axiomatic set of S*? Thesis (G) leaves this question unanswered. For any complete and consistent belief set S*, (G) thus leaves it indeterminate what set is the axiomatic set of S*. Specifically, for any belief b in S* that fits the description, (G) leaves it indeterminate whether b belongs to the axiomatic or the generated set of S*[11].

---

[11] More formally, let S* be a complete belief set that is consistent and assume S* includes categorical moral beliefs. Thesis (G) guarantees that there is a non-empty subset of the categorical moral beliefs included in S* that has the features under i.G) and ii.G): call this set A. Note that (G) guarantees also something else, i.e. that S* includes beliefs that fit the description. Why? Because it guarantees that S* includes U beliefs: those making up A. Consider any belief in A: say, the belief that ∎ all N_1's are M. Since it includes this belief and it is both complete and consistent, S* includes also the belief that ∎ all N_1&N_3's are M, which is a belief that fits the description. So (G) guarantees that S* includes beliefs fitting the description. Consider then any belief b in S* that fits the description; let b' be the other U belief in S* from which (given the second of the two conditions defining 'belief that fits the description') b follows by subsumption. Either b belongs to A or it does not. Suppose first that b does not belong to A. Let A* be the set that results from A after adding b to it. Note two things about A*. (*1*) If A has the features under i.G) and ii.G), then so does A*. As for i.G): anything that follows by subsumption from beliefs in A follows by subsumption from beliefs in A* as well, and A* actually includes one more categorical moral belief, of those in S*, than A does. As for ii.G): the only belief A* adds to A, namely b, is a U belief. (*2*) A* differs from A: b belongs to A* but does not belong to A. Suppose now that b belongs to A. Let A* be the set that results from A after taking b out of it and (if b' does not belong to A) adding b' to it. Once again, two things are true of A*. (*1*) If A has the features under i.G) and ii.G), then so does A*. As for i.G): since b follows by subsumption from b', anything that follows by subsumption from beliefs in A (which include b) follows by subsumption from beliefs in A* (which include b'); there is one categorical moral belief, of those in S*, that is included in A but not in A*, namely b, but it also follows by subsumption from beliefs in A*, since it follows by subsumption from b'. As for ii.G): the only belief (if any) that A* adds to A, namely b', is a U belief. (*2*) A* differs from A: b belongs to A but does not belong to A*. So either way, there is another set, in addition to A, that also has the features under i.G) and ii.G), and which differs from A in that b belongs to one of the two sets but not to the other. Thesis (G) leaves it indeterminate which, between these two sets, is the axiomatic set of S*, so for any belief b in S* that fits the description, (G) leaves it indeterminate whether b belongs to the axiomatic or the generated set of S*.

It follows that (G) fails to yield, about case #1 and those analogous to it, the results Gibbard will want. In these cases, we saw above, Gibbard will want to say that the belief that fits the description belongs to the generated set of the relevant belief set. But (G) fails to yield this result, since it fails to yield any result at all about any belief fitting the description. Thesis (G) therefore needs to be revised.

The problem with (G) is that it fails to yield any result about the beliefs that fit the description, so the way to fix the problem is by adding a clause to (G) providing instructions on how to deal with those beliefs. The obvious way to do that is by adding a clause to (G) that assigns any belief fitting the description, included in a complete and consistent belief set, either to the axiomatic set or to the generated set of that belief set.

The first option is, of course, a non-starter. If revised this way, (G) does yield a verdict about case #1 and those analogous to it: in these (as in all) cases, the belief that fits the description belongs to the axiomatic set of the relevant belief set. But this is not the verdict Gibbard will want. So let us turn to the second option.

The second option amounts to revising (G) as follows (as usual, S is a complete belief set and S_{cm} is the set of categorical moral beliefs included in S):

(G_{Rev}) If S is consistent then, provided S_{cm} is not empty, there is a non-empty subset of
S_{cm} –let it be the <u>axiomatic set</u>- such that
i.G)   all the beliefs in S_{cm} that are not in the subset –let their set be the
<u>generated set</u>- follow from beliefs in the subset (plus appropriate non-
moral beliefs in S) by subsumption;

ii.G) all the beliefs in the subset can be expressed by an appropriately strong universal moral sentence using a non-moral predicate denoting a non-moral property;

iii.G) no belief in S that fits the description belongs to the subset.

That is, the second option amounts to adding a clause to (G) that leaves all beliefs fitting the description out of the axiomatic set, thereby (given i.G)) assigning them all to the generated set.

This way of revising (G) is intuitively more satisfactory. If the axiomatic set is meant to provide the list of U beliefs under which all the other categorical moral beliefs can be subsumed then, granted that there is, strictly speaking, more than one such list, it is only natural to go as general as one can, in determining what U beliefs to include in the set. This is exactly what iii.G) does: a belief that fits the description is a belief that, though a U belief, can be subsumed under another U belief, so iii.G) leaves it out of the axiomatic set. This way of revising (G) is also more promising. Like the first one, it does yield a verdict about case #1 and those analogous to it; unlike the first, it yields the verdict Gibbard will want: in these cases, the belief that fits the description belongs to the generated set of the relevant belief set[12].

---

[12] Though Gibbard does not explicitly discuss the beliefs that fit the description, a case could be made (based on some remarks in his [2003]) that he means to accept the constraint on membership in the axiomatic set put forward by iii.G). I won't, however, try to make such a case: whether or not Gibbard means to accept it, the constraint is one way to address the problem with (G) that we are discussing, so we need to look into it. One question about iii.G) comes out as follows. According to (G), the axiomatic set is non-empty. (Let me briefly remind the reader why. Consider two points. One, focusing on (G) itself: as we saw in chapter 4, (G) fleshes out the Assumption that actions, people, institutions, etc. owe their moral features to their non-moral properties in terms of the idea that use of moral language commits one to general moral principles, where the principles are precisely the beliefs included in the axiomatic set, which had therefore better not be empty. Two, focusing on the use of (G) to put forward an accommodationist proposal about moral explanatory claims, and more generally an account of the content of moral sentences: as we also saw in chapter 4, the semantic proposals built on (G) work, briefly put, by replacing a moral sentence's apparent reference to a moral property with a reference to the non-moral properties picked out by the beliefs included in the axiomatic set, which had therefore, once again, better not be empty.) If (G) is revised by adding iii.G) to it, however, doesn't the axiomatic set end up being empty? That is, isn't *any* U belief in S a belief that fits the description, or one that can be subsumed under another U belief in S, so that there is no limit to how general one can go? Consider any U belief in S: say, the belief that ■ all N_1's are M. If 'F' is a non-moral predicate

However, (G_{Rev}) yields this verdict about *all* cases of our kind: it yields that, in all cases of the kind defined above, the belief that fits the description belongs to the generated set of the relevant belief set, just as theses i)-iv) –we saw in the previous section- yield that, in all cases of our kind, the truth that fits the description is a derived truth. So (G_{Rev}) must be rejected for the same reason why we rejected i)-iv): it is unable to account for case #3 and those analogous to it.

Remember how case #2 and case #3 go. Case #2 is provided by the theory I put forward when you and I discuss what deserves respect. On my theory, there is only one principle when it comes to what deserves respect: that sentient beings deserve respect. Human beings do deserve respect, I believe, but only because they are sentient beings. Case #3 is provided by the theory you put forward in our discussion. On your theory, there are two principles when it comes to what deserves respect: that sentient beings deserve respect and that so do human beings. Human beings, you believe, deserve respect for two reasons: because they are sentient and because they are human.

As I noted above, thesis (G), and so thesis (G_{Rev}), are about complete belief sets that are consistent, which can be thought of as capturing the views of users of moral language who have become fully decided, with the beliefs in the axiomatic set capturing the general principles held by such speakers. Though I am, we are supposing, decided enough to have settled on a specific

---

denoting a non-moral property, so is also 'F&~F' (think of it as the contradictory predicate); let us abbreviate the latter expression as 'F_C'. Since S is a complete belief set that includes the belief that ■ all N_1's are M, if it is consistent it includes also the belief that ■ all $N_1 \vee F_C$'s are M, and the latter is a U belief under which the former can be subsumed. Or consider the belief that ■ all $N_1 \& F_T$'s are M, where 'F_T' is shorthand for 'F$\vee$~F' (think of it as the tautological predicate): an analogous reasoning yields that, if S is consistent, it includes this belief, and this too is a U belief under which the belief that ■ all N_1's are M can be subsumed. Doesn't this show that, if S is consistent, then any U belief in S is one that can be subsumed under another U belief in S, hence that, if we accept iii.G), the axiomatic set of S ends up being empty? That is, doesn't this show that accepting iii.G) is really a bad idea? I believe this is an interesting and complex question, well worth pursuing (even more so if one adopts the broader definition of 'belief that fits the description' mentioned in n. 9). In what follows, however, I'll grant that the answer is negative, on the ground that the belief that ■ all $N_1 \vee F_C$'s are M -or that ■ all $N_1 \& F_T$'s are M- is not clearly different from the belief that ■ all N_1's are M, so the reasoning above does not clearly show that, if S is consistent, the latter belief is one that can be subsumed under *another* U belief in S, hence one that iii.G) leaves out of the axiomatic set of S.

theory about what deserves respect, I need not be fully decided about every moral and non-moral issue, so my views are modeled, not by a single complete belief set that is consistent, but by a set of such belief sets, each of which includes the beliefs that, we are supposing, I hold. If S^*_{Me} is one of the complete and consistent belief sets modeling my views, then its axiomatic set contains, Gibbard will want to say, only one belief: that ■ all sentient beings deserve respect. Though included in S^*_{Me}, the belief that ■ all human beings deserve respect belongs, Gibbard will want to say, to its generated set. No problem here for (G_{Rev}), which yields precisely this result. The belief that ■ all human beings deserve respect is a belief in S^*_{Me} that fits the description -it follows by subsumption from another U belief in S^*_{Me}, namely that ■ all sentient beings deserve respect-, so (G_{Rev}) yields that it belongs to the generated set of S^*_{Me}.

Turn, however, to your views. If S^*_{You} is one of the complete and consistent belief sets modeling your views, its axiomatic set contains, Gibbard will want to say, two beliefs: that ■ all sentient beings deserve respect and that ■ all human beings deserve respect. But (G_{Rev}) does *not* allow him to say so. It is just as true of S^*_{You}, as it is of S^*_{Me}, that it includes another U belief - that ■ all sentient beings deserve respect- from which the belief that ■ all human beings deserve respect follows by subsumption. That is, the belief that ■ all human beings deserve respect is a belief in S^*_{You} that fits the description. Thesis (G_{Rev}) thus yields that it belongs to the generated, not the axiomatic, set of S^*_{You}, which is not the result Gibbard will want.

Thesis (G_{Rev}) fails to yield, about case #3, the verdict Gibbard will want. To put it in a slightly different way, (G_{Rev}) is blind to a distinction that Gibbard will want to be able to draw. The theory you put forward differs from the one I put forward; case #3 differs from case #2. Gibbard should be able to draw this distinction, and so to say that the belief that ■ all human

beings deserve respect belongs to the generated set of the belief sets modeling my views but to the axiomatic set of the belief sets modeling yours. Thesis (G_{Rev}), however, does not allow him to say this, so it does not allow Gibbard to distinguish between case #3 and case #2, between the theory you put forward and the one I put forward.

Case #3 and case #2, we saw in the previous section, provide just one instance of the problem: the same reasoning yields that (G_{Rev}) does not allow Gibbard to distinguish between someone endorsing the doctrine of higher pleasures (on the reading of it mentioned in section 1) and the hedonist, or between Moore (or at least someone endorsing the Moore-inspired theory presented in section 1) and the hedonist; it is easy to generalize from these cases. For any pair of such cases, (G_{Rev}) is unable to draw the relevant distinction. And this is a serious count against a reading of the Assumption, since it is a *datum* that, in each of these pairs of cases, we have two different moral theories involved, rather than one and the same theory.

Let us take stock. We are looking at cases involving a belief that fits the description and asking whether Gibbard's reading of the Assumption yields, or can be revised so as to yield, the verdicts about them that Gibbard will want. We started with thesis (G): a preliminary problem with it is that it yields no verdict at all about the beliefs that fit the description. This is, as I say, a preliminary problem: we need to note it, but it is not hard to fix it. The obvious way to do that is by moving from (G) to (G_{Rev}). Thesis (G_{Rev}) does yield a verdict about our cases; moreover, it yields, about case #1 (and those analogous to it) and case #2 (and those analogous to it), the verdicts Gibbard will want. The problem with it is that it fails to yield the right verdict about case #3 (and those analogous to it); alternatively put, it is unable to distinguish between case #3 (and those analogous to it) and case #2 (and those analogous to it). *This* problem, I claim, is much deeper: it is, I now want to argue, a problem that Gibbard lacks the resources to solve.

If we think of Gibbard's reading of the Assumption as a constraint over complete and consistent belief sets, the nature of the problem can be described as follows: (G_{Rev}) fleshes out the Assumption so as to make room only for complete and consistent belief sets such that, if a belief in them fits the description, it belongs to their generated set; what Gibbard needs is instead a way of fleshing out the Assumption that makes room both for complete and consistent belief sets such that a belief in them fits the description and belongs to their generated set (case #1 and those analogous to it; case #2 and those analogous to it) and for ones such that a belief –in some instances, even the same belief- in them fits the description and belongs to their axiomatic set (case #3 and those analogous to it)[13].

How can Gibbard revise his reading of the Assumption so as to achieve this result? Well, how do cognitivist friends of the Assumption do it? As we saw in section 1, they face an analogous problem, framed in terms of possible worlds and of truths fitting the description that may be either derived or brute truths. They can solve it, I suggested, by distinguishing, within the truths that fit the description, i.e. within the U truths that are made true by other categorical moral truths, between those that are, and those that are not, true only because made true by other categorical moral truths; only the former, they can say, are derived, whereas the latter are brute. Can't Gibbard make an analogous move? He should distinguish, within the beliefs in a complete and consistent belief set S* that fit the description, i.e. within the U beliefs in S* that follow by subsumption from another U belief in S*, between those that are, and those that are not, held

---

[13] Consider again case #1 and those analogous to it. In n. 10 I noted that I am leaving it open whether there might be eccentric speakers with views modeled by complete and consistent belief sets whose axiomatic set includes both the belief that ■ all murders are wrong and the belief that ■ all murders committed on a Wednesday are wrong; that is, whether there is a case that mirrors case #1 in the same way in which case #3 mirrors case #2. As I pointed out in the same footnote, the standard position in the non-cognitivist tradition leads to the conclusion that there is such a case. If so, it too is a case that (G_{Rev}) fails to allow for.

only because they follow by subsumption from another U belief in S*; only the former, he should say, belong to the generated set of S*, whereas the latter belong to its axiomatic set.

Let me illustrate this suggestion by means of case #2 and case #3. In order to distinguish between these two cases, so goes the suggestion, Gibbard should say that you and I differ in the following respect. If I were to develop my views and become fully decided (while preserving the beliefs we are supposing I hold) then, Gibbard should say, I would hold that ■ all human beings deserve respect *only* because this belief follows by subsumption from another U belief I hold -in our example, that ■ all sentient beings deserve respect-. On the other hand, if you were to develop your views and become fully decided (while preserving the beliefs we are supposing you hold) then, Gibbard should say, you would *not* hold that ■ all human beings deserve respect only because this belief follows by subsumption from another U belief you hold –in our example, that ■ all sentient beings deserve respect-. This difference, Gibbard should say, is the reason why the belief that ■ all human beings deserve respect belongs to the generated set of the complete and consistent belief sets modeling my views, but to the axiomatic set of those modeling yours, despite being in both cases a belief that fits the description.

Gibbard, however, cannot make this move. That is, he cannot make it without giving up what is distinctive, and attractive, about his way of fleshing out the Assumption, as compared with the other non-cognitivist reading of the Assumption available in the literature, namely Hare's. On the suggestion we are considering, whether a belief in a complete and consistent belief set belongs to the axiomatic or the generated set ends up turning on *why* that belief is held: it turns on why a fully decided speaker whose views are captured by the belief set holds the belief in question. But Gibbard's reading of the Assumption is designed to be *silent* about such epistemological matters.

245

As we saw in chapter 4, Gibbard's reading is a logical claim about how categorical moral beliefs are logically related to one another in a complete and consistent belief set, one that is silent on the question why users of moral language, even fully decided users, hold their categorical moral beliefs. To illustrate with case #2: the suggestion is that Gibbard should say that, if fully decided, I would hold that ■ all human beings deserve respect only because this belief follows by subsumption from my other belief that ■ all sentient beings deserve respect; Gibbard's reading of the Assumption, however, says *nothing at all* about why, if fully decided, I would hold that ■ all human beings deserve respect, not even that I would hold it because it follows by subsumption from my other belief that ■ all sentient beings deserve respect.

And this, we also saw in chapter 4, is an *advantage* of Gibbard's reading over Hare's, which does formulate an epistemological view about why users of moral language hold their categorical moral beliefs, since it allows Gibbard's reading to avoid the problems I raised in chapter 3 for Hare's. One of those problems was that, on the epistemological view formulated by Hare's reading of the Assumption, one's categorical moral beliefs are dependent on one's moral principles: on that view, as I put it in chapter 3, evidential support within one's moral belief system ultimately flows only from one's principles to the other categorical moral beliefs those principles subsume. This, I argued, is false, since it is at odds with the fact that, as I pointed out looking at the way in which we reason about our moral views, those other categorical moral beliefs are (at least to a degree) independent of one's moral principles.

But note: the suggestion we are considering is equally at odds with that fact. Let me illustrate with case #2 again. The suggestion makes, about case #2 and my belief that ■ all human beings deserve respect –supposing I become fully decided-, the same claim that Hare's reading of the Assumption makes about my categorical moral beliefs more generally, i.e. that

that belief is dependent on my moral principles: if fully decided, I would hold that belief *only* because it follows by subsumption from my principle that ■ all sentient beings deserve respect. If Gibbard adopts the suggestion we are looking at to solve the problem we are discussing in this section, he therefore ends up vulnerable to the same problem I raised in chapter 3 for Hare, and his reading of the Assumption turns out not to be much of an improvement over Hare's[14].

If not along the lines of the suggestion we have just discussed, how can Gibbard revise his reading of the Assumption so as to achieve the result he needs? If he had better not do it, as proposed by that suggestion, by adding some epistemological claim to his reading, can he do it in terms of the resources distinctive of his way of fleshing out the Assumption, i.e. using only logical materials? No: my examples show that he cannot.

Gibbard needs to distinguish between case #2 and case #3, and so to say that the belief that ■ all human beings deserve respect belongs to the generated set of the complete and consistent belief sets modeling my views, but to the axiomatic set of those modeling yours. Assume he can say what he needs to say about me. If he can, then he is committed to saying the same about you. Why?

Because your theory and mine yield the same categorical moral beliefs, when coupled with the same set of non-moral beliefs. This means that, for any complete and consistent belief set modeling your views, there is one modeling mine that contains exactly the same beliefs[15].

---

[14] Indeed, if Gibbard adopted the suggestion we have been discussing, his reading of the Assumption would come quite close, in the respects here relevant, to the revision of Hare's reading that I mentioned in n. 8. As I noted in that footnote, though that revision is able to deal with the kind of cases we are looking at in this chapter, it is still vulnerable to the problems I raised against Hare's original reading of the Assumption in chapter 3.

[15] Strictly speaking, what is true is the following: your theory, *plus* the non-moral belief that human beings are sentient beings, and my theory, *plus* the non-moral belief that human beings are sentient beings, yield the same categorical moral beliefs, when coupled with the same set of non-moral beliefs. Since it is built into the description of case #2 and case #3 that you and I hold the non-moral belief in question, in the text I drop the qualifications.

Now, insofar as Gibbard's way of fleshing out the Assumption employs only logical materials, whether a belief in a complete and consistent belief set belongs to the axiomatic or the generated set is a matter of how that belief is logically related to the other beliefs in the belief set by means of the logical relation 'following from'. But note: if the beliefs in the belief set modeling your views and those in the belief set modeling mine are exactly the same, *the logical relations among them will also stay the same*.

If the belief that ■ all human beings deserve respect belongs –as we are assuming Gibbard can say- to the generated set of the complete and consistent belief sets modeling my views, it will therefore belong –Gibbard is committed to saying- to the generated set of the complete and consistent belief sets modeling your views as well. So, Gibbard *cannot* say what he needs to say about you. He cannot distinguish between case #3 and case #2, between your theory and mine. Nor can he distinguish between the cases making up my other pairs. These results are unacceptable.

The reason why Gibbard is committed to these results is that what is distinctive about his way of fleshing out the Assumption is that it employs logical materials, but such resources are not enough to draw the distinctions we need to draw. The logical relations employed by Gibbard's way of fleshing out the Assumption cannot discriminate between theories yielding the same set of categorical moral beliefs, because given a set of categorical moral beliefs, the logical relations among them are fixed. The metaphysical relations employed by the cognitivist, on the other hand, can discriminate between theories yielding the same set of categorical moral truths, because given a set of categorical moral truths, the metaphysical relations among them are *not*

fixed: the same set of truths can be arranged in different ways in terms of the relation 'being made true by', resulting in different sets of brute truths[16].

The source of this difference is that, as I remarked when introducing the cognitivist reading of the Assumption (in section 2.2) and Gibbard's reading (in section 4.1), the instantiation of the metaphysical relation 'being made true by' amounts to something more than the instantiation of the logical relation 'following from': as it is sometimes put, the logical relation 'following from' is a lighter relation than the straightforwardly metaphysical one 'being made true by'. Remember my example about Dave, one of the students in my same year in the philosophy graduate program. The belief expressed by 'Dave is a good person' follows from the beliefs expressed by 'All the people in my year are good people' and 'Dave is in my year'. However (even supposing that the three sentences in question are all true), saying that the truth expressed by the first sentence is made true by the truths expressed by the other two amounts to adding something else -something that is in fact false-.

Employing the lighter relation 'following from' allows Gibbard to flesh out the Assumption without appealing to the relation 'being made true by', which is not available to non-cognitivists. But precisely because it is lighter, the relation 'following from' allows us to draw less fine-grained distinctions than does the relation 'being made true by'.

In particular, we saw above, it does not allow us to distinguish between theories yielding the same set of categorical moral beliefs, such as those involved in my pairs of cases. But we do need to draw those distinctions: it is a *datum* that each of my pairs of cases involves two different

---

[16] Epistemological relations also can discriminate between theories yielding the same set of categorical moral beliefs: the same set of beliefs can be arranged in different ways in terms of the relation 'being inferred from'. This explains why (as I noted in n. 8) there is a way to revise Hare's reading of the Assumption that is able to draw the relevant distinctions when applied to the cases we are looking at in this chapter. Hare's way of fleshing out the Assumption must be rejected, not because it is unable to deal with these cases, but for the reasons I gave in chapter 3.

moral theories, rather than one and the same. Gibbard's attempt to flesh out the Assumption using logical materials is therefore bound to be inadequate. The idea that actions, people, institutions, etc. owe their moral features to their non-moral properties cannot be captured in light logical terms.

Nor, I argued in chapter 3, can that idea be captured in epistemological terms: Hare's attempt to flesh out the Assumption using epistemological materials also fails. The idea that actions, people, institutions, etc. owe their moral features to their non-moral properties can be adequately captured only in metaphysical terms. The most natural way of fleshing out the Assumption, by means of a straightforwardly metaphysical picture describing how categorical moral truths are arranged in terms of the metaphysical relation 'being made true by', turns out to be the only adequate way of fleshing it out: there is no adequate non-cognitivist reading of that idea. If you like the Assumption, you had better be a cognitivist.

At the end of the previous chapter, after presenting Gibbard's semantic proposals and showing they avoid all the problems I raised in chapter 3 for the proposals built on Hare's reading of the Assumption, I pointed out that the following questions remained open: in the first place, whether Gibbard's proposal about moral explanatory claims succeeds in accommodating our practice with regard to such claims; secondly, whether his 'Go by the hyperstates' strategy – of which his accommodationist proposal is a particular application- provides an adequate account of the content of moral sentences quite generally, and one that succeeds in mimicking the cognitivist.

The argument in this section for the inadequacy of Gibbard's reading of the Assumption has immediate consequences for his accommodationist proposal about moral explanatory claims:

Gibbard's proposal is built on his reading of the Assumption, so once that reading is rejected as inadequate, the proposal must be rejected too.

The point here is entirely general. *All* the accommodationist proposals we looked at in the previous chapters are built on (the relevant reading of) the Assumption: it is the idea that things owe their moral features to their non-moral properties that is meant to make it plausible that (as all the accommodationist proposals suggest) what we discuss, when we discuss our moral explanatory claims, is the causal efficacy of *non*-moral properties, i.e. the non-moral properties that things owe their moral features to. If (the relevant reading of) the Assumption goes, so must the accommodationist proposal built on it.

Let me expand on this general observation, taking the details of Gibbard's accommodationist proposal into account. As we saw in section 4.2, what is meant to make Gibbard's proposal plausible is (bracketing complications that are not relevant here) the following reasoning: start with Gibbard's idea that the content of a moral explanatory claim 'M-ness causes E' can be accounted for in terms of the complete and consistent belief sets that include the corresponding belief; *given thesis (G)*, it is plausible to represent each such belief set by means of a conjunction of the form 'N-ness causes E and ■ all N's are M' (with 'N' a non-moral predicate denoting a non-moral property), and so to represent the content of 'M-ness causes E' by means of the big disjunction of conjunctions mentioned by Gibbard's proposal. Once thesis (G) is rejected and no adequate way of revising it is found, there is no reason anymore to think of that disjunction as a plausible way of capturing the content of 'M-ness causes E'.

One might still find it plausible that the content of a moral explanatory claim can be accounted for in terms of the complete and consistent belief sets that include the corresponding

belief, fair enough. But without Gibbard's reading of the Assumption, that idea will not lead to a proposal on which the content of 'M-ness causes E' is captured by a sentence that ascribes no causal power to moral properties, and *this* is what is crucial to the accommodationist project. Without Gibbard's reading of the Assumption, that idea will not lead to an *accommodationist* proposal.

It will be obvious that analogous considerations apply to Gibbard's 'Go by the hyperstates' strategy too, so I can be brief here. All the strategies for assigning content (or payoff) to moral sentences that we looked at in the previous chapters, be they cognitivist or non-cognitivist, are built on (the relevant reading of) the Assumption: as I pointed out when introducing those strategies, what is meant to make each of them a plausible way to account for moral sentences is (the relevant reading of) the Assumption. Once Gibbard's reading is rejected as inadequate, so must his general account of the content of moral sentences. Moreover, since the reason I have given in this section for rejecting that account does not, as we saw in section 1, apply to the cognitivist, Gibbard's account of moral sentences fails to mimic the cognitivist's. Despite all that, as we saw in the previous chapter, his 'Go by the hyperstates' strategy achieves for the non-cognitivist in terms of the mimicking project, Gibbard's attempt to mimic the cognitivist proves ultimately unsuccessful.

3. Second objection to the 'Go by the hyperstates' strategy (and one more objection to the 'Go by the speaker's standards' strategy)

In section 1, I drew attention to a certain kind of cases and highlighted two lessons to be learnt from them. The second lesson points to verdicts about those cases that a friend of the Assumption *will* want her reading of the Assumption to yield. In section 2, I argued that

Gibbard's reading of the Assumption fails to yield those verdicts and must therefore be rejected as inadequate. Since Gibbard's accommodationist proposal about moral explanatory claims, and more generally his account of moral sentences, are built on his reading of the Assumption, this is one reason to reject the proposal, and the general account.

The problem I raised in the previous section for Gibbard's semantic proposals is thus rooted in his reading of the Assumption. In this section, I'll raise a different problem. I'll bracket the argument from the previous section and grant Gibbard that he has an adequate way of fleshing out the Assumption. *Even so*, I'll argue, there is a further problem with his accommodationist proposal, and with his general account, one that comes out if we attend to the first of the two lessons I pointed out in section 1.

Indeed, the problem I want to raise in this section has a broader scope than the one in section 2: it is a problem, not only for Gibbard's semantic proposals, but also for the semantic proposals built on Hare's way of fleshing out the Assumption, which we looked at in chapter 3. It is actually easier to introduce the problem starting with the latter proposals.

By way of reminder, the first of the two lessons I highlighted in section 1 points to verdicts about cases of our kind that a friend of the Assumption will *not* want her reading of the Assumption to commit her to. The lesson is (as I put it in section 1, in terms applicable to the cognitivist reading of the Assumption) that being a U truth is not sufficient for being a brute truth, or (as I also put it in section 1, in terms applicable to Hare's way of fleshing out the Assumption) that being a U belief is not sufficient for being a direct belief. This is the lesson I drew from case #1 and those analogous to it: think again of ' ■ All murders are wrong' and ' ■ All murders committed on a Wednesday are wrong'. Quite generally put: supposing that ' ■ All N_1's are M' expresses a brute truth, or a direct belief held by a given speaker, ' ■ All N_1&F's are

253

M' does not, for many, many replacements for 'F', also express a brute truth, or a direct belief held by that speaker (even if we assume that the speaker does, unsurprisingly, believe that ∎ all N₁&F's are M).

As we saw in section 1, accommodating this lesson does not require making any revisions to either the cognitivist or Hare's reading of the Assumption. It does, however, require revising the semantic proposals built on those readings, since it means that the properties to 'go by', when doing the semantics of moral sentences, cannot be read off the U truths, or the U beliefs of a given speaker. The substitution operations relied on by the cognitivist proposals are defined in terms of the n's, and accommodating this lesson requires narrowing the scope of this expression, reading 'the n's' as shorthand for 'the non-moral properties that can be denoted by some non-moral predicate 'N' such that ' ∎ All N's are M' expresses a *brute truth*', not merely 'a *truth*'. Analogously for the proposals built on Hare's reading of the Assumption: accommodating the first lesson requires reading 'the n_S's' as shorthand for 'the non-moral properties that can be denoted by some non-moral predicate 'N' such that ' ∎ All N's are M' expresses a *direct belief* held by S', not merely 'a *belief* held by S'.

For my discussion in this section, it will be helpful to remember that, as I pointed out in section 2, an analog of the first lesson holds for Gibbard's reading of the Assumption: being a U belief included in a complete and consistent belief set is not sufficient for belonging to the axiomatic set of that belief set[17]. It will also be helpful to use 'U sentence' (in conformity with my usage of 'U truth' and 'U belief' in the previous sections) as shorthand for 'appropriately strong universal moral sentence using a non-moral predicate denoting a non-moral property'.

---

[17] See n. 10.

In order to raise the problem I want to press in this section, we need to go back once again to the cognitivist proposals that the non-cognitivist is trying to mimic, specifically to the cognitivist proposals that rely on enriched singular substitution. By way of reminder, here is the cognitivist proposal about moral explanatory claims that relies on enriched singular substitution (supposing there is only one brute truth ascribing M-ness, say that ■ all N_1's are M):

(B) The content, or payoff, of 'M-ness causes (does not cause) E' is that N_1-ness causes (does not cause) E and N_1-ness is the n[18].

What is distinctive about the cognitivist proposals that rely on enriched singular substitution is that they put forward, as capturing the content (or payoff) of a moral sentence, a conjunction of *two* sentences: in addition to the first conjunct, which replaces the moral sentence's apparent reference to the moral property of being M with a reference to a non-moral property, the second conjunct (*a*) expresses the relevance of the non-moral property in question – what makes *that* property the property to replace for being M-, namely that it is the M-making property, the property picked out by the brute moral truth there is. And, we saw in chapter 2, the advantage of this type of proposals over those relying on non-enriched singular substitution, which omit the second conjunct, is that they yield that the original moral sentence (or the sentence capturing its payoff) has moral content, given that the second conjunct (*b*) has moral content: 'N_1-ness is the n' is shorthand for 'N_1-ness is the non-moral property that can be

---

[18] The problem I want to press can be raised about both moral explanatory claims and categorical moral sentences: throughout this section, I'll use moral explanatory claims to illustrate it. As we saw in chapter 2, the cognitivist proposals specify the content (or payoff) of 'M-ness causes E' in different ways, depending on whether there is only one, or more than one, brute truth ascribing M-ness: for simplicity's sake, in my discussion I'll consider the case in which there is only one brute truth ascribing M-ness; analogous considerations apply to the case in which there is more than one such truth.

denoted by some non-moral predicate 'N' such that ' ■ All N's are M' expresses a brute *truth*, so it entails that ■ all N₁'s are M, hence it has moral content.

Let us look now at the non-cognitivist proposals, built on Hare's way of fleshing out the Assumption, that rely on enriched singular substitution. By way of reminder, here is the non-cognitivist proposal about moral explanatory claims that relies on enriched singular substitution (for the case of a user S of moral language who has made up her mind about what direct beliefs to accept, for a given moral predicate 'M', and holds only one such belief, say that ■ all N₁'s are M):

(B.H) The content of 'M-ness causes (does not cause) E' (as uttered by S) is that N₁-ness causes (does not cause) E and ■ all N₁'s are M[19].

The non-cognitivist proposals that rely on enriched singular substitution are designed to mimic the cognitivist proposals of the same type. They also put forward, as capturing the content of a moral sentence, a conjunction of two sentences: the first conjunct mentioned by (B.H), like the first mentioned by (B), replaces a moral explanatory claim's apparent reference to the moral property of being M with a reference to a non-moral property; the second conjunct mentioned by (B.H), like the second mentioned by (B), is designed to (*a*) express the relevance of the non-moral property in question –what makes that property the property to replace for being M-, in this case that it is the property singled out by S's standard, the one picked out by the direct moral

---

[19] As we saw in chapter 3, the non-cognitivist proposals built on Hare's way of fleshing out the Assumption specify the content of 'M-ness causes E' in different ways, depending on whether the relevant speaker -who has made up her mind about what direct beliefs to accept- holds only one, or more than one, direct belief: for simplicity's sake, in my discussion I'll consider the case of a speaker who holds only one direct belief; analogous considerations apply to the case of a speaker who holds more than one such belief.

belief held by S. And here the non-cognitivist faces at least a *prima facie* problem: I mentioned it briefly in chapter 3, saying I would get back to it in this chapter; we now need to take a closer look at it.

In order to express the fact that N_1-ness is the property picked out by the brute moral truth there is, the cognitivist has the second conjunct be 'N_1-ness is the *n*'. On the other hand, I noted in 3.2, in order to express the fact that N_1-ness is the property picked out by the direct moral belief held by S, the non-cognitivist cannot have the second conjunct be 'N_1-ness is the n_S'. Why? Because the second conjunct is designed to (*b*) have moral content, but 'N_1-ness is the n_S' lacks any moral content: it is a non-moral remark about what S believes. It is shorthand for 'N_1-ness is the non-moral property that can be denoted by some non-moral predicate 'N' such that ' ∎ All N's are M' expresses a direct *belief* held by S', so it does *not* entail that ∎ all N_1's are M, and indeed it has no moral content at all. The advantage of enriched over non-enriched singular substitution -the whole point of adding a second conjunct- is precisely that enriched substitution preserves the moral content of the moral sentence being analyzed, so the non-cognitivist cannot frame the second conjunct as 'N_1-ness is the n_S'.

How, then, can the non-cognitivist mimic the cognitivist here? How can she frame the second conjunct? Well, what does have moral content are the sentences expressing a speaker's standards -' ∎ All N_1's are M', in the case of our speaker S-, so the non-cognitivist can solve her problem if she can frame the second conjunct in terms of such sentences; and the idea, we saw in 3.2, is that she can do it as follows.

Look again at condition (*a*): the second conjunct is designed to expresses the fact that N_1-ness is the property picked out by the direct moral belief held by S. To say the same thing in different words, it is designed to express the fact that S has made up her mind about what direct

beliefs to accept and holds only one, that ∎ all N_1's are M, while rejecting all others. Therefore, so goes the idea, the non-cognitivist can have the second conjunct be '∎ All N_1's are M and not (∎ all N_2's are M) and not (∎ all N_3's are M) and …'. This big clause, of course, meets condition (*b*) as well: it clearly has moral content. So the idea is that the non-cognitivist can solve her problem by framing the second conjunct in terms of the sentences expressing the standards the speaker accepts and those she rejects, which is how the second conjunct mentioned by (B.H) is formulated (as I noted in 3.2 and elsewhere, in my formulations of the non-cognitivist proposals I have usually shortened the big clause to '∎ All N_1's are M', dropping the rest and taking it as understood).

In my discussion so far, I have granted that this idea works, but we are now finally in a position to see that it really won't do[20]. The second conjunct, I have just said, is designed to express the fact that S has made up her mind about what direct beliefs to accept and holds only one, that ∎ all N_1's are M, while rejecting all others. However, it does not follow from this, as the idea would have it, that the non-cognitivist can frame the second conjunct in terms of the big clause mentioned above. Why? Because the big clause is made up of U sentences and negations of U sentences, but S need not reject all U beliefs other than the belief that ∎ all N_1's are M: she may well accept some of the U sentences that are preceded by 'not' in the big clause. Remember the first lesson: being a U belief is not sufficient for being a direct belief, so the fact that S holds only one direct belief, that ∎ all N_1's are M, does not mean she does not hold other U beliefs as

---

[20] In 3.2, I already pointed to a reason for thinking that this idea won't work, which goes as follows: the idea can only work if the non-cognitivist can account for what it is to reject a standard, but she cannot do that, since non-cognitivists cannot deal with negation. As I said there, however, though I find the worry about negation compelling, in the present work I am granting that non-cognitivists can answer it, so I'll bracket this problem for the idea we are looking at. What I am about to raise in the text is a further problem for that idea, one that arises *even if* one grants that non-cognitivists can deal with negation (this is the further problem I promised I would raise for that idea in ch. 3, n. 15).

258

well. For example, S may (unsurprisingly) believe that ■ all N_1&F's are M (for some replacement for 'F'). Indeed, it is utterly unlikely that S does not hold, in addition to her direct belief, other U beliefs like this, which follow straightforwardly from her direct belief.

Alternatively put, the big clause is off target: at most it expresses, not that the belief that ■ all N_1's are M is S's only direct belief, but that it is S's only U belief. This is utterly unlikely to be true. Even more importantly, it is *not* what the non-cognitivist needs. To say it one more time: the second conjunct is designed to (*a*) express the fact that N_1-ness is the property picked out by the *direct belief* held by S. The big clause fails to do that: at most, it expresses that N_1-ness is the property picked out by the *U belief* held by S. The big clause thus meets condition (*b*), fair enough, but it fails to meet condition (*a*). The non-cognitivist cannot use it as her second conjunct.

In order to make sure that her second conjunct is on target, the non-cognitivist can change the big clause to a clause to the effect that '■ All N_1's are M' expresses a *direct belief* held by S and '■ All N_2's are M' does not express a *direct belief* held by S and '■ All N_3's are M' does not express a *direct belief* held by S and … . But this amounts (in the terminology I have been using) to changing it to a clause to the effect that N_1-ness is the n_S, so we are back to where we started. Such a clause does meet (*a*), but it fails to meet (*b*): it lacks moral content. The *prima facie* problem for the non-cognitivist is a real one.

So far I have focused on the non-cognitivist proposals built on Hare's way of fleshing out the Assumption: those that rely on enriched singular substitution, I have argued, turn out on closer inspection not be available to the non-cognitivist. What about Gibbard's proposals, which rely on general substitution? They are not available to the non-cognitivist either. The argument above applies, *mutatis mutandis*, to those proposals as well, given that –as we saw in chapter 4-

what they put forward, as capturing the content of a moral sentence, is the disjunction of all the conjunctions of the sort mentioned by the non-cognitivist proposals discussed so far.

By way of reminder, here is Gibbard's proposal about moral explanatory claims:

(C.G') The content of 'M-ness causes E' is that

(N_1-ness causes E and ■ all N_1's are M) or

(N_2-ness causes E and ■ all N_2's are M) or … or

(N_1-ness causes E and N_2-ness causes E and ■ all N_1's are M and ■ all N_2's are M) or

(N_3-ness causes E and N_4-ness causes E and ■ all N_3's are M and ■ all N_4's are M) or … or

…[21].

Remember how Gibbard's semantic proposals work. The content of a moral sentence is accounted for in terms of the set of complete and consistent belief sets that include the corresponding belief. Such belief sets are represented by means of conjunctions: those among them whose axiomatic set contains only one categorical moral belief, that ■ all N_1's are M, are represented by the first conjunction mentioned by (C.G'); those among them whose axiomatic set contains only one categorical moral belief, that ■ all N_2's are M, are represented by the second

---

[21] As we saw in chapter 4, Gibbard's proposal about moral explanatory claims comes in two varieties, depending on whether the operation of general substitution relied on is multiple or disjunctive: for simplicity's sake, in my discussion I'll consider the proposal relying on multiple general substitution; analogous considerations apply to the one relying on disjunctive general substitution.

conjunction; and so on. The set of all such belief sets, hence the content of 'M-ness causes E', is then represented by means of the disjunction of those conjunctions.

Let us focus on the first conjunction mentioned by (C.G'): 'N_1-ness causes E and ■ all N_1's are M'; considerations analogous to those I am about to make apply to all the other conjunctions as well. To repeat, that conjunction is designed to represent those, among the relevant complete and consistent belief sets, whose axiomatic set contains only the belief that ■ all N_1's are M. Moreover, since Gibbard's proposals are (like the other proposals we have looked at in this section) designed to preserve the moral content of the moral sentence being analyzed, that conjunction is designed to have moral content.

The first conjunct of the conjunction we are considering replaces a moral explanatory claim's apparent reference to the moral property of being M with a reference to a non-moral property: it is designed to be a non-moral sentence. In order for the conjunction to meet the conditions above, its second conjunct is therefore designed, we saw in chapter 4, to (*a*) represent the complete and consistent belief sets whose axiomatic set contains only the belief that ■ all N_1's are M, and (*b*) have moral content. The second conjunct of the conjunction we are looking at is thus subject to constraints analogous to those on the second conjunct mentioned by (B.H) above. And, I noted in 4.2, the idea behind the way in which Gibbard's proposals are formulated is that it can meet those constraints in an analogous way.

The conjunct we are looking at, I have just said, is designed to (*a*) represent the complete and consistent belief sets whose axiomatic set contains only the belief that ■ all N_1's are M. One way to think of such belief sets is as modeling the views of a speaker who has settled on a specific set of standards, holding only one, that ■ all N_1's are M, while rejecting all others. Therefore, so goes the idea, Gibbard can frame the conjunct in question as ' ■ All N_1's are M and

261

not (■ all N_2's are M) and not (■ all N_3's are M) and …' (as I noted in 4.2 and elsewhere, in my formulations of Gibbard's proposals I have usually shortened this big clause to ' ■ All N_1's are M', dropping the rest and taking it as understood). And, of course, if framed this way the conjunct in question meets (*b*) as well, since this big clause clearly has moral content.

In my discussion of Gibbard's proposals, I have so far granted that this idea works, but here too we are finally in a position to see that it really won't do[22]. The big clause above fails to (*a*) represent the complete and consistent belief sets whose axiomatic set contains only the belief that ■ all N_1's are M. The reason is that that clause is made up of U sentences and negations of U sentences, but an analog of the first lesson holds for Gibbard's reading of the Assumption: being a U belief included in a complete and consistent belief set is not sufficient for belonging to the axiomatic set of that belief set. It follows that the belief sets the big clause is designed to represent may include other U beliefs as well, in addition to the belief that ■ all N_1's are M: they may include beliefs expressed by some of the U sentences preceded by 'not' in the big clause. Indeed, they are guaranteed to: since they are complete and consistent, they are guaranteed to include, for example, the belief that ■ all N_1&F's are M. The big clause is therefore off target: at most it represents, not the complete and consistent belief sets such that the belief that ■ all N_1's are M is the only *belief belonging to their axiomatic set*, but those such that that belief is the only *U belief included in them*.

In order to have a conjunct that meets (*a*), Gibbard can change the big clause to a clause to the effect that ' ■ All N_1's are M' expresses a belief that belongs to the axiomatic set and ' ■

---

[22] Once again, in 4.2 I already pointed to a reason for thinking that this idea won't do, rooted in the worry about negation. Since, however, in the present work I am granting that non-cognitivists can answer that worry, I'll bracket this problem. The problem in the text is a *further*, independent problem for the idea we are looking at (this is the further problem I promised I would raise for that idea in ch. 4, n. 16).

All N$_2$'s are M' does not express a belief that belongs to the axiomatic set and ' ■ All N$_3$'s are M' does not express a belief that belongs to the axiomatic set and … . Such a clause, however, obviously fails to (*b*) have moral content. Gibbard therefore does not have a way to frame the conjunct in question. The semantic proposals he puts forward turn out on closer inspection not to be available to him[23].

In this section, I have raised a problem for the non-cognitivist proposals that rely on either enriched singular substitution or general substitution. This problem is avoided by the cognitivist whom the non-cognitivist is trying to mimic. I have already shown this about the cognitivist who relies on enriched singular substitution; it is easy enough to see that the same holds of the cognitivist who relies on general substitution.

By way of reminder, here is the cognitivist proposal about moral explanatory claims that relies on general substitution (supposing there is only one brute truth ascribing M-ness):

(C) The content, or payoff, of 'M-ness causes E' is that the *n* causes E[24].

What is distinctive about the cognitivist proposals that rely on general substitution is that they replace the moral expression occurring in a moral sentence with the definite description 'the *n*', i.e. with an expression still containing moral terminology. Like those relying on enriched singular substitution, such proposals also yield that the original moral sentence (or the sentence

---

[23] In n. 20 and n. 22 I pointed out that my argument against the non-cognitivist proposals we have looked at in this section is not rooted in the worry about negation. My conclusion, however, is very much in the spirit of the conclusion drawn by those who find that worry compelling, namely that there is something fishy about non-cognitivist semantics: non-cognitivists are not really in a position to advance the semantic proposals they put forward. My argument shows that the worry about negation is only *one* way to press this charge, so that even if non-cognitivists can answer that worry, they still have work to do if they want to avoid the charge.

[24] As I pointed out in n. 18, in discussing the cognitivist proposals I'll consider the case in which there is only one brute truth ascribing M-ness; analogous considerations apply to the case in which there is more than one such truth.

capturing its payoff) has moral content. Unlike those relying on enriched singular substitution, however, they do so without the need for an additional conjunct designed to secure this result. Therefore, the issues discussed in this section, which center around the conditions that such an additional conjunct has to meet, just do not arise with regard to the cognitivist proposals that rely on general substitution.

The argument in this section provides a further reason, in addition to those I offered in chapter 3, to reject the non-cognitivist semantic proposals, built on Hare's way of fleshing out the Assumption, that rely on enriched singular substitution. It is a reason that applies both to the accommodationist proposals about moral explanatory claims and to the proposals about categorical moral sentences. Moreover, since my argument does not, as I have just noted, apply to the corresponding cognitivist proposals, it provides a further reason to think those non-cognitivist proposals fail to mimic the cognitivist's. The arguments in chapter 3 grant that the proposals in question are available to the non-cognitivist and point to specific problems they face. The argument in this section claims there is a prior, general problem with those proposals, namely that they are not available to the non-cognitivist in the first place: the non-cognitivist is not even in a position to advance them.

The argument in this section also provides my second reason, in addition to the one I offered in the previous section, to reject Gibbard's semantic proposals, which rely on general substitution. It is, once again, a reason that applies both to Gibbard's accommodationist proposals about moral explanatory claims and more generally to his account of moral sentences; and a further reason to think that Gibbard's attempt to mimic the cognitivist proves ultimately unsuccessful. In the previous section, I argued that Gibbard's proposals must be rejected because the reading of the Assumption they are built on is inadequate. In this section, I have granted

264

Gibbard that he has an adequate way of fleshing out the Assumption and argued that the proposals he sets out to build on it are not in any case available to him: he is not in a position to advance them either.

In the present work, I have distinguished three different types of semantic proposals put forward by cognitivists and non-cognitivists, depending on whether the substitution operation relied on is non-enriched singular substitution, enriched singular substitution, or general substitution. The argument offered in this section applies only to the non-cognitivist proposals of the second and third type: the non-cognitivist proposals relying on non-enriched singular substitution are immune to it.

By way of reminder, here is the non-cognitivist proposal about moral explanatory claims that relies on non-enriched singular substitution (for the case of a user S of moral language who has made up her mind about what direct beliefs to accept, for a given moral predicate 'M', and holds only one such belief, say that ■ all N_1's are M):

(A.H) The content of 'M-ness causes E' (as uttered by S) is that N_1-ness causes E[25].

The proposals that rely on non-enriched singular substitution omit the additional conjunct that is distinctive of those relying on enriched singular substitution, so the argument in this section does not apply to them. This is of little help to the non-cognitivist, however, since the arguments I offered in chapter 3 apply to *all* non-cognitivist proposals built on Hare's way of fleshing out the Assumption, including those that rely on non-enriched singular substitution.

---

[25] This is Blackburn's accommodationist proposal about moral explanatory claims. As I pointed out in n. 19, in discussing the non-cognitivist proposals built on Hare's way of fleshing out the Assumption I'll consider the case of a speaker who has made up her mind about what direct beliefs to accept and holds only one direct belief; analogous considerations apply to the case of a speaker who has made up her mind and holds more than one such belief.

This is of little help to the non-cognitivist also for another reason, worth pointing out explicitly. Precisely because they omit such an additional conjunct, the proposals that rely on non-enriched singular substitution yield that the moral sentence being analyzed lacks moral content. This is a very implausible result in itself; moreover, it is one that the non-cognitivist has strong, independent reasons for wanting to avoid, since her account of moral agreement and disagreement presupposes that moral sentences do have moral content. And indeed –as we saw in chapter 3-, the non-cognitivist proposals that rely on non-enriched singular substitution are put forward by the non-cognitivist as adequate only for a limited class of peripheral uses of moral sentences, *not* for their central, typical use[26]. Insofar as a non-cognitivist accommodationist proposal, and more generally a non-cognitivist account of moral sentences, are to be minimally adequate, they have to rely, by the non-cognitivist's own standards, on either enriched singular substitution or general substitution, thereby falling within the scope of the argument in this section.

Here then is a different way of putting my conclusion in this section, and in the present work as a whole. If you like the Assumption, I argued in section 2, you had better be a cognitivist: there is no adequate non-cognitivist reading of that idea. And *even if* non-cognitivists have an adequate way of fleshing out the Assumption, I argued in this section, they cannot in any case use it to build their semantic proposals on it, insofar as those proposals are designed to meet the minimal requirement of preserving the moral content of moral sentences. If you want to build your semantic proposals on the Assumption, you had still better be a cognitivist. All in all, the prospects of your accommodationist project, and more generally of your account of moral sentences, look a lot brighter if you are a cognitivist.

---

[26] As I noted in 3.5, Blackburn –who does rely on non-enriched singular substitution in his accommodationist proposal about moral explanatory claims- is in this respect the exception, and a surprising one at that, not the rule.

BIBLIOGRAPHY

Alwood, Andrew. 2010. "Imperative clauses and the Frege-Geach problem." <u>Analysis</u> 70: 105-17.

Armstrong, David. 1983. <u>What is a Law of Nature?</u> Cambridge: Cambridge University Press.

Audi, Robert. 1991. "Moral epistemology and the supervenience of ethical concepts." <u>Southern Journal of Philosophy</u> 29 (Supplement): 1-24.

Audi, Robert. 1993. "Ethical naturalism and the explanatory power of moral concepts." In Wagner, Stephen and Richard Warner, eds. 1993. <u>Naturalism: a Critical Appraisal</u>. Notre Dame: University of Notre Dame Press, pp. 95-115.

Ayer, Alfred Jules. 1936. <u>Language, Truth, and Logic</u>. London: V. Gollancz.

Bennett, Karen. forthcoming. <u>Makings Things Up</u>. New York: Oxford University Press.

Blackburn, Simon. 1984. <u>Spreading the Word</u>. New York: Oxford University Press.

Blackburn, Simon. 1991a. "Just causes." <u>Philosophical Studies</u> 61: 3-17.

Blackburn, Simon. 1991b. "Reply to Sturgeon." <u>Philosophical Studies</u> 61: 39-42.

Blackburn, Simon. 1996. "Securing the nots: moral epistemology for the quasi-realist." In Sinnott-Armstrong, Walter and Mark Timmons, eds. 1996. <u>Moral Knowledge?</u> New York: Oxford University Press, pp. 82-100.

Boyd, Richard. 1988. "How to be a moral realist." In Sayre-McCord, Geoffrrey, ed. 1988. <u>Essays on Moral Realism</u>. Ithaca: Cornell University Press, pp. 181-227.

Brink, David. 1989. <u>Moral Realism and the Foundations of Ethics</u>. Cambridge: Cambridge University Press.

Brink, David. 2013. <u>Mill's Progressive Principles</u>. Oxford: Oxford University Press.

Copp, David. 1990. "Explanation and justification in ethics." <u>Ethics</u> 100: 237-58.

Geach, Peter. 1965. "Assertion." <u>The Philosophical Review</u> 74: 449-65.

Gibbard, Allan. 1990. <u>Wise Choices, Apt Feelings</u>. Cambridge, MA: Harvard University Press.

Gibbard, Allan. 2002. "Normative explanations: invoking rationality to explain happenings." In Bermudez, Jose Luis and Alan Millar, eds. 2002. <u>Reason and Nature. Essays in the Theory of Rationality</u>. Oxford: Oxford University Press, pp. 265-82.

Gibbard, Allan. 2003. <u>Thinking How to Live</u>. Cambridge, MA: Harvard University Press.

Griffin, James. 1992. "Values: reduction, supervenience and explanation by ascent." In Charles, David and Kathleen Lennon, eds. 1992. <u>Reduction, Explanation and Realism</u>. Oxford: Oxford University Press.

Hare, Richard. 1952. <u>The Language of Morals</u>. Oxford: Oxford University Press.

Harman, Gilbert. 1977. <u>The Nature of Morality</u>. New York: Oxford University Press.

Harman, Gilbert. 1986. "Moral explanations of natural facts – Can moral claims be tested against moral reality?" <u>Southern Journal of Philosophy</u> 24 (Supplement): 57-68.

Hooker, Brad and Margaret Little, eds. 2000. <u>Moral Particularism</u>. Oxford: Oxford University Press.

Jackson, Brendan. 2007. "Truth vs. pretense in discourse about motion (or, why the sun really does rise)." <u>Nous</u> 41: 298-317.

Jackson, Frank and Phillip Pettit. 1990. "Program explanation: a general perspective." <u>Analysis</u> 50: 107-17.

Mackie, John. 1977. Ethics. Harmondsworth: Penguin.

McDowell, John. 1979. "Virtue and reason." The Monist 62: 333-50.

Mill, John Stuart. 1871. Utilitarianism.

Moore, George Edward. 1903. Principia Ethica. Cambridge: Cambridge University Press.

Quinn, Warren. 1986. "Truth and explanation in ethics." Ethics 96: 524-44.

Railton, Peter. 1986a. "Moral realism." The Philosophical Review 95: 163-207.

Railton, Peter. 1986b. "Facts and values." Philosophical Topics 14(2): 5-31.

Railton, Peter. 1989. "Naturalism and prescriptivity." Social Philosophy and Policy 7: 151-74.

Railton, Peter. 1998. "Moral explanation and moral objectivity." Philosophy and Phenomenological Research 58: 175-82.

Sayre-McCord, Geoffrey. 1988. "Moral theory and explanatory impotence." In Sayre-McCord, Geoffrrey, ed. 1988. Essays on Moral Realism. Ithaca: Cornell University Press, pp. 256-81.

Schroeder, Mark. 2008. Being For. New York: Oxford University Press.

Shafer-Landau, Russ. 2003. Moral Realism: a Defense. Oxford: Oxford University Press.

Slote, Michael. 1971. "The rationality of aesthetic value judgments." The Journal of Philosophy 68: 821-39.

Stevenson, Charles. 1944. Ethics and Language. New Haven: Yale University Press.

Sturgeon, Nicholas. 1985. "Moral explanations." In Copp, David and David Zimmerman, eds. 1985. <u>Morality, Reason and Truth</u>. Totowa, NJ: Rowman and Allanheld, pp. 49-78.

Sturgeon, Nicholas. 1986a. "Harman on moral explanations of natural facts." <u>Southern Journal of Philosophy</u> 24 (Supplement): 69-78.

Sturgeon, Nicholas. 1986b. "What difference does it make whether moral realism is true?" <u>Southern Journal of Philosophy</u> 24 (Supplement): 115-41.

Sturgeon, Nicholas. 1991. "Contents and causes." <u>Philosophical Studies</u> 61: 19-37.

Sturgeon, Nicholas. 1992. "Nonmoral explanations." <u>Philosophical Perspectives</u> 6: 97-117.

Sturgeon, Nicholas. 1995a. "Evil and explanation." In Nielsen, Kai and Jocelyn Couture, eds. 1995. <u>On the Relevance of Metaethics</u>. Calgary: University of Calgary Press, pp. 155-85.

Sturgeon, Nicholas. 1995b. "Critical study of *Wise Choices, Apt Feelings*." <u>Nous</u> 29: 402-24.

Sturgeon, Nicholas. 1998. "Thomson against moral explanations." <u>Philosophy and Phenomenological Research</u> 58: 199-206.

Sturgeon, Nicholas. 2002. "Ethical intuitionism and ethical naturalism." In Stratton-Lake, Philip, ed. 2002. <u>Ethical Intuitionism: Re-evaluations</u>. Oxford: Oxford University Press, pp. 184-211.

Sturgeon, Nicholas. 2006a. "Ethical naturalism." In Copp, David, ed. 2006. <u>The Oxford Handbook of Ethical Theory</u>. New York: Oxford University Press, pp. 91-121.

Sturgeon, Nicholas. 2006b. "Moral explanations defended." In Dreier, James, ed. 2006. <u>Contemporary Debates in Moral Theory</u>. Oxford: Blackwell, pp. 241-62.

Sturgeon, Nicholas. 2009. "Doubts about the supervenience of the evaluative." In Shafer-Landau, Russ, ed. 2009. <u>Oxford Studies in Metaethics</u> vol.4. Oxford: Oxford University Press, pp. 53-90.

Sturgeon, Nicholas. 2010. "Mill's Hedonism." <u>Boston University Law Review</u> 90: 1705-29.

Thomson, Judith. 1990. <u>The Realm of Rights</u>. Cambridge, MA: Harvard University Press.

Thomson, Judith. 1996. "Moral objectivity." In Harman, Gilbert and Judith Thomson. 1996. <u>Moral Relativism and Moral Objectivity</u>. Oxford: Blackwell, pp. 67-154.

Thomson, Judith. 1998. "Reply to critics." <u>Philosophy and Phenomenological Research</u> 58: 215-22.

Williams, Bernard. 1985. <u>Ethics and the Limits of Philosophy</u>. Cambridge, MA: Harvard University Press.

Wright, Crispin. 1992. <u>Truth and Objectivity</u>. Cambridge, MA: Harvard University Press.

Zangwill, Nick. 2006. "Moral epistemology and the because constraint." In Dreier, James, ed. 2006. <u>Contemporary Debates in Moral Theory</u>. Oxford: Blackwell, pp. 263-81.