The Analogy to the Moral Realm

In Chapter 1, I attempted to provide an initial understanding of the discussion of underdetermination in science. The main goal of this book is to investigate what we can learn from this discussion with regard to the realm of moral philosophy. The reason why this should prove to be instructive, I claim, is that there is an analogy to be found: Just as scientific theories can be underdetermined by the empirical data, so moral theories can be underdetermined by our considered judgments or intuitions about particular cases. This chapter takes a closer look at the analogy or, more precisely, the two notions that are substituted in the analogy: *moral theories* for *scientific theories* and *considered judgments or intuitions about particular cases* for the *empirical data*.

Some preliminary remarks are required. First, as already noted in the Introduction, the supposed analogy is a *structural* one. I am not claiming that since the subject matter of science and ethics is somehow similar in nature and there is underdetermination to be found in science, we should also expect to find it in ethics. Instead, the analogy is only concerned with a similarity in the relation between the two sorts of theories and their corresponding relata. As will become clear, I do think that the two sorts of theories share two common functions. However, I am not presupposing that the evidence of science and ethics are epistemically or metaphysically on par *in every respect*. Hence, I will sometimes highlight differences between the realms of science and ethics where they are important for understanding the analogy. Yet, at least until Part III of the book, I will mostly attempt to remain neutral on conspicuous epistemic or metaphysical questions as far as they do not directly pertain to the structural analogy.

Second, I do not claim that mine is the only way to draw an analogy between science and ethics. There are other ways of conceiving of an analogy to scientific theorizing that have their own merits, yet engaging with them would go beyond the scope of this book. Ultimately, the

35

important question is not whether this is the only way to draw the analogy, but whether it is a useful one. This depends on it making us see matters in ethics in a new and illuminating light, which, I hope, is what it does.

Third, and perhaps most importantly, in outlining the analogy, I am not claiming that the depiction given here is the last word on how to think of some of the issues involved. As will become clear in Part III, I am myself rather skeptical about the prospects of the kind of explanation in ethics outlined in this chapter. However, this is a rather revisionary view, and I do not want to frame the whole analogy in such revisionary terms, since that would make it unacceptable to the majority of readers. Instead, I am hoping to outline an analogy that most readers will recognize as more or less accurate with regard to how the issues at hand are standardly conceived of in normative ethics. My own misgivings will have to wait until Part III, and I hope that even though few readers will probably accept all that is to come, they might at least recognize it as not too far away from how many of their peers think of moral theorizing.

2.1   The Underlying Picture of Moral Theories

The first notion that is substituted in the analogy is that of a moral theory for that of a scientific one. The notion of a moral theory is a very basic one in normative ethics. Anyone taking an introductory class on normative ethics will encounter Kantian, consequentialist, and contractualist theories. Nonetheless, if we are going to develop an analogy between scientific and moral theories, we need a more precise depiction of what moral theories are. I start with a basic understanding that identifies two functions of moral theories. I then say more about the two functions, respectively.

Two Functions of Moral Theories

Here is a fairly common view of moral theories. Moral theories serve (at least) two functions. First, they have to yield the correct verdicts about what we ought to do in particular cases.[1] In doing this, they have to account for the considered judgments and intuitions about particular cases that we already have as well as provide guidance about future cases. Second, moral theories have to explain why we have to choose these actions. As it is often

---

[1]   Remember that this is a simplified way of talking about all deontic verdicts, which also include verdicts about what is permissible to do, allowed to do, and so on.

put, moral theories have to give an account of what makes some act right, or of the right-makers.[2]

As attentive readers will already have noticed, these two functions align with the two tenets of the Textbook View from the Introduction. This should not come as a surprise, of course. The Textbook View is simply a specific view regarding the two main functions of moral theories. It states that regarding both functions, the main traditions of moral theories arrive at incompatible results. The Textbook View thus presupposes the picture of the two functions, and if I was right to say that the former is quite widely shared, then so too must the latter be.

One reason for this seems to be that the distinction between the two functions maps onto another distinction that has much initial appeal: the distinction between a *practical* and a *theoretical* side of moral theorizing. An early statement of this can be found in Smith, who states that moral principles serve two roles:

> The first such role may be viewed as *theoretical*. In this role, moral principles specify the characteristics in virtue of which acts possess their moral status. [...] The second function for moral principles may be thought of as *practical*. [They are used as] a standard by reference to which a person can guide his or her own behavior: a standard to help the person choose which acts to perform and which not. (Smith, 1988, pp. 89–90)[3]

Why is it these two roles that moral theories (need to) fulfill? Very plausibly, it has to do with two related purposes of moral theorizing. As Tännsjö explains:

---

[2] An early use of this terminology can be found in Bales (1971). Note that when metaethicists talk about moral explanation, they often understand it in another way, going back to Harman (1977). Harman questioned whether postulating moral facts is indeed part of our best explanation of why people make certain moral judgments. Perhaps facts about, e.g., people's upbringing might suffice and thus provide a more parsimonious explanation. This is not the kind of explanation we are interested in here. Moral facts, in Harman's sense, are appealed to in order to explain something *about* our (moral) practice. They are supposed to explain why we act in a certain way, not why that action is correct or incorrect. We can call this an *external* explanation. In contrast, normative moral theories offer *internal* explanations. They explain why some act is right or wrong *from a moral point of view*.

[3] The thought is echoed by Leibowitz:

> It is common to distinguish between two different roles of moral theorizing: a theoretical role and a practical role. The theoretical role of moral theorizing is to provide an account of the rightness / wrongness of actions. [...] The practical role of moral theorizing is to guide judgment or action. (Leibowitz, 2009, pp. 349–350)

> In our lives, when faced with hard choices, we seek, if we take things seriously, practical guidance from moral theory. We try to find out what we *ought* to do. But, if we are of a reflective bent, we also want to understand *why* we ought to do what we ought to do. (Tännsjö, 2006, p. 213)[4]

Thus, the two functions stem from two purposes of moral theorizing: They tell us how to morally behave, and they explain to us why we should do so. The first demand is for help concerning our actions, while the second is for purposes of explanation.[5]

What is more, and of most importance for our analogy, these functions are also ones that are, on a plausible reading, shared with scientific theories. This fact has not escaped ethicists. Here is one example:

> A scientific or ethical theory can serve two functions. One is to discover new truths, which in ethics means using the theory's principles to arrive at moral judgements about particular cases that we could not otherwise make. The other is to explain truths we do know; here the principles tell us why particular moral judgements we already confidently make are true. (Hurka, 2011, p. 18)

Hurka identifies precisely the two functions outlined above – one having to do with yielding the correct verdicts, the other with explaining those verdicts – and observes that those are shared with scientific theories. Not coincidentally, the same bifurcated picture also underlies Dietrich and List's formal representation of moral theories that leads them to note the phenomenon of moral underdetermination. Their framework (the reason-based representation) distinguishes between two dimensions of moral theories:

---

4   Väyrynen agrees:

> We can distinguish two motivations for constructing ethical theories. One is *practical*: we might want an ethical theory to guide action. Another is *theoretical* and, specifically, *explanatory*: we might want an ethical theory to explain why certain actions are right and others wrong. (Väyrynen, 2006, p. 291)

5   The emphasis on the practical side has one aspect that I gloss over. Some of the quoted authors are also concerned with how specific theories have to be in order to serve as applicable, immediately helpful guides in concrete situations. They are thus interested in the very concrete question of how to *effectively* guide our behavior. This is a question I will largely neglect in what follows. Instead, I prefer to speak, in a more abstract way, about theories' extensions. Theories are practical insofar as they yield sets of verdicts for particular cases. Whether they are optimally suited to be applied by creatures such as us will not concern me. See Bales (1971) for a discussion of this point as it relates to act-utilitarianism, especially p. 264, where he draws a comparison to scientific theory.

> Reason-based representations encode not only a theory's *action-guiding recommendations* (that is, how we should act, according to the theory) but also the *reasons* behind those recommendations (that is, why we should act in that way). (Dietrich and List, 2017, p. 422)

Slightly more technically, they distinguish between a rightness (or permissibility) function, which yields a set of action-guiding verdicts for each situation, and a reasons structure, which encodes the underlying reasons that are being given for why specific acts are right or wrong, permissible, mandatory, or forbidden. The permissibility function answers the question "Which actions are right or wrong?", while the reasons structure answers the question "Why are those actions right or wrong?". Normative moral theories, Dietrich and List insist, include both: a rightness function and a reasons structure.

Overall, the picture of moral theories serving two functions is very plausible and well founded in the literature. But what exactly do these two functions amount to in the moral case? Let us consider each more closely in turn.

Yielding the Correct Verdicts

To fulfill the first role identified above, that of providing practical guidance, a moral theory has to yield the correct verdicts about what we should do. How does it do that? There is something close to a classical understanding to be found in the literature, what I shall call the *deductive principles model*.[6] In this section, I give a quick overview of the model and then consider some criticisms. This will lead me to add some modifications to the model while resisting an alternative picture that is too revisionary. First, however, there are some terminological issues that need to be clarified.

Some Preliminaries

To describe the first function in terms of "yielding the correct verdicts" is ambiguous. It might mean at least three things. First, it might mean that the theory has to lead to verdicts that fit the *considered judgments* or *intuitions* we already hold. The theory thus has the somewhat *passive* function of correctly *accounting* for or *capturing* our judgments and intuitions. Second,

---

[6] Leibowitz (2011) only calls it the *deductive* model. However, it is important for our purposes to make clear that the model includes both the notion of *deduction* as well as that of a *principle*. My choice of a slightly more awkward name is supposed to highlight this. Otherwise, my discussion in this section owes a great deal to Leibowitz.

as the quote from Hurka above makes clear, we might also want our moral theories to provide guidance in situations where we do not already have considered judgments or intuitions. Here, it would be misleading to say that the theory is only accounting for some intuitions or judgments. Instead, a more *active* term like *predicting* or *resulting in* the correct verdicts is more apt. As Hurka further points out, this is not different from scientific theories. In science, theories are assessed both by how accurately they can account for the already given evidence and by how their predictions will or will not be corroborated by future observations. Third, however, we might also think that moral theories, in contrast to scientific ones, are also in the business of *correcting* our verdicts. For example, we might think that the verdicts we presently hold are tainted by our accepting the wrong theories and that the correct moral theories should show us which verdicts we should actually hold. I will tend to this (seeming) disanalogy at the end of this chapter.

Here, I am only interested in how to marshal terminology. One way to avoid the problem related to these distinctions is to follow Dietrich and List's use of the concept of *deontic content*. The deontic content of a theory is the total set of verdicts it yields about what we should do. This includes verdicts on cases about which we have (independent) judgments or intuitions as well as cases about which we do not (yet) have such intuitions. It also makes no mention of whether the verdicts that are yielded (the deontic content) fit with the verdicts we presently hold. In the context of this book, we might then further follow Dietrich and List and say that moral theories are underdetermined by their deontic content, thereby remaining neutral with regard to the distinctions above. Using this formulation is, strictly speaking, misleading. The verdicts that constitute a theory's deontic content are derived from the theory. Claiming that these same verdicts underdetermine the theory itself leads to a viciously circular picture of theory confirmation. What we should say instead is that the theory is underdetermined by how its deontic content compares to our (at least partly) independently given considered judgments or intuitions, present and future. However, since this is quite a mouthful, I suggest that we accept talk of a theory being underdetermined by its deontic content *as long as we are clear* that this is just a shorthand. I will often use this terminology when I consider it irrelevant whether the judgments and intuitions that have to be accounted for are the ones we already hold or ones we should hold.

However, there are also cases where we need to distinguish more precisely. For example, some of the authors I will discuss do indeed think that it is

an advantage of theories if they accord with common sense.[7] Theories, on this view, are measured by how well they account for our presently held judgments. In other cases, authors are concerned with the judgments that *the best* moral theories yield.[8] These need not necessarily be those of common sense. Since the framework of moral underdetermination proves instructive for understanding both of these cases, it will not be helpful to restrict the terminology in one direction. Instead, for reasons of convenience, I will for the most part gloss over these distinctions, but use more precise language when it is called for.

Having cleared away this complication, let us now consider in more detail how theories are supposed to fulfill their function of yielding the correct verdicts.

The Deductive Principles Model

If there is a standard view of how moral theories yield verdicts on specific cases, it is what I call the *deductive principles model*. Here are two early representative statements:

> Ideally a normative 'theory' consists of a set of *general* principles analogous to the axioms of a geometric system. That is, ideally it comprises a set of *correct* or valid general *principles*, as *brief* and *simple* as possible compatibly [sic] with *completeness* in the sense that these principles, when conjoined with true nonethical statements, would logically imply every ethical statement that is correct or valid. Such an ideal for a system must be our guide. (Brandt, 1959, p. 295)

and

> [...] one determines what one should do in a particular situation by reference to certain general principles and rules, which one takes as premises from which to deduce a particular conclusion by a kind of practical syllogism, as Aristotle called it. One takes general principles and applies them to individual situations. (Frankena, 1973, p. 2)

Here, too, is a more recent example:

> If there are true moral principles [...] then we can use them, together with relevant factual information, as premises in a deductive argument leading up to a conclusion about what it is we ought to do. [...] In morality we deduce our moral obligations from factual statements [...] together with moral principles [...]. (Tännsjö, 2006, pp. 213–214)

---

7   Portmore (2011) is the prime example.
8   Parfit (2011b) would be the obvious example.

Two notions are relevant in these statements. First, it is assumed that theorizing in ethics involves reference to general moral *principles*. These principles are supposed to yield correct verdicts on more particular cases and thereby guide our actions. Second, the way that the principles lead to these verdicts is laid out in terms of *deduction*, that is, strict logical entailment from the principles together with the relevant non-moral information.

Perhaps the most intuitive examples of this model are certain consequentialist theories. As an example, take a simple utilitarian version (SU) that includes only one principle specifying that:

> SU: Necessarily, an action is morally right if and only if, and because, it maximizes overall utility.[9]

From this principle, together with the relevant information about the utility of the different outcomes of all relevant choice situations, we could logically deduce the deontic status of all acts. Granted, we likely never have that information about utility, which severely impacts the real-life guidance such a theory could provide. However, as regards the logic of moral theorizing alone, the deductive principles model provides for a neat and tidy picture. If we choose the right principles and are furthermore correct about the non-moral facts, deductive logic alone will tell us what to do.

Its neatness is probably one factor in why the model has found many adherents in ethics. Another reason is that it lends itself to an analogy to science. Leibowitz (2011) has recently highlighted that, on closer inspection, the deductive model in ethics is analogous to the *deductive-nomological* model of explanation in science (the DN model, for short). Here is how Leibowitz describes the latter:

> An explanation, on this model, has the form of a deductive argument. The conclusion of the argument is (a sentence describing) the phenomenon to be explained – the *explanandum* – and the premises of the argument – (the sentences describing) the *explanans* – must logically imply the explanandum. Moreover, the explanans must include a statement of a general law, and this law must play an indispensable role in the derivation of the explanandum. (Leibowitz, 2011, pp. 475–476)

The similarities should be obvious. Both the deductive model in ethics and the DN model in science conceive of the connection between theories and

---

9   I will say more about the form of the principle in the next section.

the data in the form of a logical deduction from general laws or principles together with some other information.[10]

Both the deductive principles model's theoretical simplicity and its structural similarity to science might seem to make it an ideal candidate for use in the proposed analogy between ethics and science. Alas, the similarity to the model in science is not good news for ethicists. Instead, it should prompt us to modify the model.

Giving up Deduction, Holding on to Principles

The problem that the comparison with the DN model unearths, as Leibowitz (2011, pp. 472–473) argues, is that the DN model has generally fallen out of favor in the philosophy of science, and this should lead ethicists, who still subscribe to a similar model, to amend their views as well. I think that this is basically correct. However, I only agree with Leibowitz about one of two major revisions that he urges ethicists to make as a consequence of this.

The first revision concerns the notion of deduction in the model. As we saw above, perhaps some consequentialist theories might indeed conform to a deductive model. However, other theoretical frameworks accord less with such a mechanistic picture of moral reasoning. Within those frameworks, applying a principle to a particular case is more complicated.

As an example, take W. D. Ross's pluralist moral theory. Ross (1930) thought that instead of just one moral principle, there are a number of morally relevant considerations. These give rise to a plurality of duties, which often come into conflict. For example, Ross (1930, pp. 17 ff.) was convinced that keeping promises was one of the main imperatives of morality. Yet he also saw that keeping promises in all cases has disastrous consequences. Say that by breaking a minor promise, one could prevent a major accident. Ross could not fathom how morality would stand in the way of preventing the accident in such a case. At the same time, he did not want to rescind the initially plausible requirement of keeping the promise. His well-known solution to this problem was that the conflicting duties had to be framed as *prima facie* duties. Prima facie duties, contrary to what their name suggests, do not just *seem* to be duties.[11] Instead, roughly, prima

---

[10]   Indeed, another early commentator on moral explanation, Ladd (1952, p. 499), explicitly points to one of the founding texts of the DN model, Hempel and Oppenheim (1948), as an example of a more precise account of explanation.

[11]   The phrase *prima facie* invites misunderstandings because of its unhappy epistemic connotation. It suggests that prima facie duties only seem to be duties at first sight, but might turn out not to be on closer inspection. Nothing could be further from Ross's intentions. If anything, Ross (1930,

facie duties are such that they would prescribe an action as long as they are the only morally relevant consideration that is applicable to a specific case. More often than not – indeed Ross (1930, pp. 19 ff.) thinks perhaps always – this is not the case, and more than one prima facie duty is incumbent on an agent. In such a case, one prima facie duty outweighs the other(s) and becomes the *duty proper* (or the actual duty).

We can already see how this deviates from a purely deductive model, which makes no mention of weighing competing claims. Of course, one could try to come up with an extensive manual specifying the relative weights of such duties, which could be fitted into a deductive model.[12] However, Ross (1930, p. 19) himself thinks that only extensive reflection on a case-by-case basis will allow us to decide between conflicting prima facie duties. Of course, these decisions could still take place within a more-or-less systematic decision procedure. For example, Smith (1988, pp. 95–102) discusses in some detail the use of auxiliary principles that facilitate the transition from very general principles to concrete cases.[13] Still, what is clear is that these authors, and many contemporary ethicists with them, think that more complex skills are involved in moral theorizing than a merely deductive model can do justice to. If we do not want to exclude these views as non-starters, we should follow the lead of philosophers of science. We do not need to insist that explanation in ethics must have the form of a deductively valid argument.

However, we should resist a further, much more consequential, revision that Leibowitz calls for. This is to give up altogether the idea that moral principles are *exceptionless* in that they identify characteristics that all and only the correct acts share. Leibowitz thinks that the belief that moral theories are in the business of coming up with one or several principles that should account for all correct verdicts is misguided. He takes the fact that no such principles have yet been identified to speak strongly against their existence. Leibowitz is, of course, not the first to argue in this vein. Frankena, directly following the quote above, notes:

> In all fairness, however, we must observe at this point that some moral thinkers have a different view of the logic of moral deliberation. [...] act-deontologists and other proponents of "situation ethics" take particular

---

pp. 31 ff.) thinks that whereas our prima facie duties are evident, we can never have certainty about our actual duties in specific situations.

[12] Alternatively, Stratton-Lake (2011) proposes that by eliminating what he considers *derivative* duties we arrive at a regime that does not allow conflicting duties.

[13] Compare also Smith (2012, pp. 373–378) and Frankena (1973, p. 3), who acknowledges that we cannot always solve a moral problem just by appealing to rules, but have to determine which of the rules takes precedence.

judgments to be basic in morality, rather than general ones, which they regard as inductive generalizations from particular cases, if they recognize the existence of general rules at all. (Frankena, 1973, pp. 2–3)

More recently, Dancy (2004) has helped to revive such a view, putting *moral particularism* firmly on the map. There is no space to discuss his and other particularists' arguments in detail here. However, I do want to reply to one line of argument that Leibowitz puts forward since it directly appeals to the analogy to science. Leibowitz (2011, p. 473) suggests that the reason why ethicists are still holding on to a deductive model, including exceptionless principles, is that there has been surprisingly little uptake of the lessons from the philosophy of science in ethics. The latter part may be true. However, I am not convinced that failing to heed the lessons from science is the only or even the main reason that has led ethicists to resist the turn away from exceptionless principles. The reason, I assume, is a different one. The move away from the deductive-nomological model in science has, at least to some degree, dovetailed with the recognition that in most of the special sciences, we simply do not find such exceptionless principles. As Leibowitz (2011, p. 479) recounts, Hempel and Oppenheim were themselves aware of this and even illustrated it with an example from economics.[14] Even though this ultimately did not prompt them to modify their own model, it is easy to see why other philosophers of science did. The turn away from a model that assumes exceptionless laws was supported by the actual science in a bottom-up way.

The same is not true for ethics. Much of moral theorizing still takes place within one of the main traditions, which include exceptionless general principles. Ross's view is perhaps the clearest counter example.[15] Yet it is clearly an outlier. Hence, we cannot say that a turn away from principles in ethics would be supported in the same bottom-up way that the turn away from exceptionless laws in science was. Instead, particularism as a meta-theory of normative theorizing is a much more revisionary view. It implies that most of the experts in the field are mistaken. Granted, this is not a sufficient argument to refute the view, and if we wanted to argue for

---

[14] They considered the explanation of a sudden price drop in cotton as a consequence of one large speculator selling off their stock, leading to more sales by smaller speculators. As Hempel and Oppenheim make clear, economics does not provide anything close to an exceptionless law that could explain such an episode.

[15] Yet note that even Ross assumes that although principles only have a prima facie status, they apply without exception. Prima facie duties do not cease to exist when they are outweighed by another prima facie duty. Only one of them will exert its action-guiding force, but the other's moral relevancy is not diminished in the process.

generalism in a more detailed form, we would have to offer a positive case for it.[16] However, this cannot be done here. Particularism poses a challenge for (almost) everyone (one that is arguably greater than the one posed by moral underdetermination), and it is for another book to try to answer that challenge. In what follows, I will thus hold on to the idea that general principles have an important role to play in our theorizing, even if we have to give up the idea that we can always deduce their implications for more particular cases in a straight forward, deductive way.

Explaining the Verdicts

One function of moral theories is to yield the correct deontic verdicts. The other is to explain why these are the correct verdicts. *That* moral theories are indeed in the business of providing such explanations will seem fairly obvious to most ethicists:[17] offering an account of why some acts are right or wrong is an integral part of what moral theorizing is about, and ethicists have a good intuitive grasp of how different theories attempt to do so. We are familiar with consequentialists referring to the outcomes of acts, Kantians giving special weight to autonomy or universalizability, and contractualists using the idea of an agreement in an initial position, when they attempt to explain what we ought to do.

Still, it will be helpful to say a bit more about this function. What does moral explanation consist in? This question has recently attracted a great deal of attention after several authors argued that we should understand moral explanation in terms of *grounding*, a notion that has itself been occupying metaphysicians for quite some time now.[18] Much in this literature is hotly debated, so we need to proceed with some caution.[19] Yet, the debate is instructive in two respects. First, it provides us with a model of moral explanation that will allow us to structure the discussion in Part II. Second, it helps us to be clear about where the metaethically controversial issues lie. Since I attempt to postpone such issues to Part III, it will be important to take note of them and avoid them if possible. Yet, as I shall also argue, we may not be able to remain completely neutral if

---

[16] Compare McKeever and Ridge (2006) for such a positive argument.
[17] Berker (2018, p. 15) goes so far as to claim that: "[i]t is almost a truism that we want our moral theories not merely to be extensionally adequate, but moreover to be properly explanatory."
[18] Compare Rosen (2017), Berker (2018, 2019), Enoch (2019), and Fogal and Risberg (2020).
[19] For more critical assessments of whether the grounding notion is indeed helpful to understanding moral explanation, compare Baker (2021) and Väyrynen (2013).

we want to be faithful to how most normative ethicists (and probably most metaethicists) think of the subject matter.[20]

The Grounding *Model of Moral Explanation*

Let us start with the less controversial parts of the model. These have their origins in a simple observation. When observing how ethicists talk and write about moral explanation, some highly distinct formulations catch the eye. Those include phrases such as "feature F *making* act A obligatory," "act A being wrong *in virtue of* its consequence C," or "the rightness of act A *depending* on feature F." How are we to understand these phrases? Several philosophers hold that such talk is quite naturally amenable to an analysis in terms of *grounding*.

Grounding is a notion that has attracted much attention in metaphysics lately.[21] An example might be the best way to introduce it. Consider the following claim:

> [...] there is a labor strike *due to the fact* that the truck drivers are refusing to work and instead picketing outside their workplace. (Bliss and Trogdon, 2016, p. 1)

What kind of claim is this? It seems to pick out some kind of dependency relation, in which the fact that the truck drivers refuse to work determines that a strike obtains.[22] But what kind of dependency relation? It will be useful to start with a negative characterization. The grounding relation can be distinguished from several other dependency relations. First, it is not a causal claim. The two events (the labor strike and the truck drivers refusing to work) are taking place simultaneously. If we wanted to give a causal explanation of the strike, we would have to refer to some antecedent event, like the refusal of the truck company to increase wages, for example. Second, the claim is not a linguistic or conceptual one. It is not being claimed that striking just means that some truck drivers are picketing.

---

[20] I emphasize this because I myself will argue for a much more revisionary understanding of moral explanation in Part III. However, this understanding is emphatically *not* how I take most ethicists to understand the notion. Hence, when trying to outline an analogy that should be acceptable to most ethicists and thus not negatively impact what will be argued for with regard to normative ethics, it is important to work with a less revisionary notion.

[21] The philosopher who is credited most for (re-)introducing the term into contemporary metaphysics is Fine (1994, 2001).

[22] I gloss over the fact that some grounding proponents do not think of grounding as a relation. I do not think that much that is being said in what follows depends on this issue. I also speak of facts here, as do, e.g., Audi (2012) and Berker (2019), but will assume that what's being said could be translated into frameworks that don't see the grounding relation as one obtaining between facts.

Third, it goes beyond some familiar logical terms such as *necessary co-extension* or *supervenience*. What is being claimed is not only that for the strike to cease to exist, there would (necessarily) have to be a change in the drivers' behavior. The claim is rather that it is the fact that the truck drivers are striking that makes it the case that a strike obtains and not vice versa. Grounding claims are said to be *hyperintensional*. They are sensitive to how we pick out the grounding and the grounded facts, even if the two are necessarily co-extensive. Whereas necessary co-extension is *symmetric* and *reflexive*, the grounding relation is standardly thought of as both *asymmetric* and *irreflexive*. This makes immediate sense if we think of grounding as itself an explanatory notion or at least closely linked to explanation.[23] If we claim that the fact that the truck drivers are striking explains why there is a strike, we are not simultaneously claiming that the fact that a strike obtains explains why the truck drivers are striking. Furthermore, we are not going to accept as an explanation for why a strike obtains the fact that a strike obtains.[24]

How does this translate to ethics? Berker (2018, p. 14) provides a simple scheme. He contends that normative theories typically advance claims of the following structure:

> "Necessarily, an action is morally right if and only if, and because,___."

The gap is filled differently by the rival traditions. A consequentialist finishes it using some phrase such as "it maximizes utility," a contractualist with a phrase such as "it can be agreed upon by everybody," and so on. Yet the *because*, in all theories, stands for the same relation, that of moral grounding.[25]

That relation can be negatively characterized in an analogous way to what I have just outlined. First, the explanations that figure in normative ethical theorizing are not plausibly thought of as causal explanations. The consequentialist does not claim that an act's wrongness is caused by it leading to suboptimal consequences; nor does the Kantian think of the Categorical Imperative in terms of a cause. Second, with some prominent

---

[23]   There is a discussion as to whether grounding is itself an explanatory notion or whether it is only tightly linked to explanatory notions. Compare Berker (2019, p. 907). Again, I'll assume that most of what is being said can be framed in one way or the other.

[24]   For a discussion of these last points, compare Audi (2012).

[25]   In the literature, there is a debate about whether the notion of grounding in ethics is the same as the one in metaphysics and, hence, whether grounding is a unitary notion. Berker (2018) thinks it is, but see also Fine (2012) for a different view. Be that as it may, both sides agree that what is established here is a highly specific relation of priority or dependency between one feature (e.g., the fact that some act maximizes utility) and another feature (e.g., the fact that this act is morally wrong).

exceptions (Bentham perhaps being the most notorious), ethicists have not thought of proponents of other traditions as conceptually confused when making rival explanatory claims. They don't consider questions regarding moral explanation to be (merely) conceptual ones but rather substantive ones about what makes acts right or wrong. Third, most ethicists do not want to only make claims of supervenience or necessary co-extension. For example, they don't just want to make claims to the effect that there can be no difference concerning the rightness of an act without there also being a difference in its consequences. Instead, they want to claim that it is the consequences that render or make an act right or wrong.[26] This is why Berker adds the "because" clause to the "if and only if." In sum, explanatory claims in ethics neither attempt to establish a (merely) causal, nor a (merely) conceptual, nor a (merely) logical link. Instead, they specify the features that ground an act's deontic status and thereby explain it.

I take it that what has been said so far will strike many ethicists as uncontroversial. It is also a rather thin description of moral explanation. Still, it has some implications that will prove important in what comes later. The first, and more important, implication is that it renders the disagreements between the different traditions substantial. lf explanation is about the grounds of rightness, our main alternative traditions cannot all be right, since they are picking out different grounds. If one tradition claims that rightness flows from consequences to acts, and another tradition claims that it comes from features of the act itself, these explanations are, on the grounding model, incompatible. At least one of the traditions is wrong about what makes acts right or wrong; it inaccurately describes a relation in the moral domain. This is crucial for our discussion of underdetermination. If moral explanation is hyperintensional, then we can distinguish between different moral traditions even if they turn out to be necessarily co-extensive. Hence, the presumption should be that, pending further argument, there will be disagreements remaining between the main moral traditions even if they come to the same conclusions about what we should do. The resulting underdetermination will thus be, to use the notion from Chapter 1, *radical*. This, to repeat, will probably not be controversial for most ethicists. Indeed, it is perfectly compatible with the Textbook View. The grounding model provides us with a way to understand why different moral traditions disagree regarding their explanations, and

---

[26] Recognition of this fact significantly predates recent discussion of grounding. Compare Depaul (1987).

it further implies that this is something substantially different from these theories disagreeing about what we should do.[27]

The second implication is as follows. Since grounding is a metaphysical notion, normative ethicists, who routinely offer moral explanations, are routinely doing metaphysics. Furthermore, since moral metaphysics is traditionally the domain of metaethics, this would also seem to imply that the traditional distinction between normative ethics and metaethics is put into question. Berker (2018, pp. 40 ff.) welcomes this, arguing that the distinction between normative ethics and metaethics collapses. I do not think that we have to put it in such dramatic terms. Normative ethics, as it has historically grown, is not a neatly confined discipline but involves a variety of issues. When it comes to explanation, it seems, at least at first glance, that there are some metaphysical aspects involved.[28] However, as we have so far described these aspects, they seem rather innocuous. They are metaphysical in the rather boring, disciplinary sense of that term that distinguishes certain topics from others that we consider in, for example, the philosophy of language or the philosophy of science. They do not involve any esoteric-seeming claims. As Berker himself writes elsewhere:

> In recent years, it has become customary to use the term 'grounding' as the chief way of designating the form of dependence at issue here, but don't let that label mislead you. 'Grounding' is not a technical term referring to a recently discovered relation. It is simply a new name for a familiar, commonplace notion that ordinary folk make use of [...]. (Berker, 2019, p. 906)[29]

Additionally, we can still quite easily distinguish between the work of a normative ethicist and that of a metaethicist. Most importantly, normative ethicists must not hold, much less defend, an elaborate theory about any of the metaphysically controversial issues concerning moral explanation to which we turn next.

The More Controversial Aspects
Two issues regarding the grounding model deserve some more space because they have metaethical implications that potentially impact how

---

27 As we shall see, matters are more complicated because some of the theories we will consider are not *standard* versions of their traditions, and the question of whether they nonetheless radically disagree with some deontically equivalent version of another theory thus remains to be decided.

28 Indeed, Berker (2018, pp. 5–7) notes that some of the earliest texts that show at least a partial awareness of the distinct relation of grounding are by ethicists: Hare (1952) and Dancy (1981). Compare also Fogal and Risberg (2020, p. 178).

29 Compare also Dasgupta (2017, pp. 75–76) and Trogdon (2013, p. 99) who calls this the *quotidian* understanding of grounding.

widely acceptable the proposed model of moral explanation is. I will argue that one of these issues can be put aside without taking a side. The other, however, requires me to take a stance that will be controversial for at least some metaethicists.

First, the less problematic issue. So far, I have claimed that moral theories have an explanatory side and that this side can be understood in terms of the grounding relation. However, moral principles can be understood to be explanatory in two ways. One is that they just outline or appeal to an explanation. They point to the features of acts that make these acts right or wrong. The other is that the principles themselves figure irreducibly in metaphysical explanations. That is, they do not just identify the grounds but are themselves (part of) these grounds. Berker (2019, pp. 904–905) calls the former sense "explanation-involving" and the latter "explanation-serving."

For those not well acquainted with the grounding literature, this might seem arcane. Yet the distinction is important because it has repercussions for many metaethical debates. For example, according to one way of drawing the distinction between metaethical naturalism and non-naturalism, naturalist views are those that claim that moral facts are fully grounded by non-moral facts, whereas non-naturalists deny this. Non-naturalists thus have to provide an account of what, besides non-natural facts, grounds the moral facts. One especially promising candidate for this can be found precisely in moral principles. Moral principles, on this view, serve as the laws that, together with the non-moral facts, ground an act's being right or wrong. At least one intuitive way to argue for metaethical non-naturalism thus presupposes that moral principles are themselves parts of explanations, that is, explanation-serving.[30]

Fortunately, whether moral explanations are explanation-serving is much more controversial than whether they are explanation-involving. Both naturalists and non-naturalists can accept that moral principles are explanatory in the explanation-involving sense.[31] This sense is all that our discussion of moral underdetermination requires. As I remarked above, one of the crucial aspects of the grounding model, for our purposes, is that it renders the moral explanations that different traditions advocate radically incompatible. Yet, this is true regardless of whether the explanations simply capture different morally relevant grounding features or whether they are

---

[30] Compare Rosen (2017), Enoch (2019), and Fogal and Risberg (2020).
[31] Even though non-naturalists will deny that the principles themselves are part of what grounds the rightness of an act, they can accept that the principles point to features that do provide these grounds (or can be reduced to other grounds). Compare Berker (2019, p. 908).

themselves (non-natural) parts of them. This fact is presumably also the reason why the way that normative ethicists formulate explanatory claims is mostly mute in this regard.[32] This is how it should be. The metaphysically controversial aspects are not normative ethicists' primary business, and formulating, for example, the Categorical Imperative, should not prejudge the issue between metaethical naturalists and non-naturalists.

The second issue is trickier. Baker (2021) has recently argued that the grounding model itself (not just the explanation-serving interpretation of it) is unduly biased against certain metaethical positions, specifically, expressivist ones. The grounding model assumes that when we make "because" claims in ethics, we are referring to some metaphysical dependence or priority relation. However, on Baker's understanding:

> The expressivist denies that normative properties reduce to or can be realized by the natural. She also denies that there is a sui generis relation of metaphysical determination between the natural and the normative. The normative, then, isn't grounded in any familiar way at all [...]. Baker (2021, p. 3)

Baker, instead, proposes an alternative account of moral explanation modeled on certain understandings of explanation in the philosophy of science and mathematics. On this account, roughly, an explanation is a unifying generalization. We explain by showing how a particular injunction is a specific instance of a more general one. For example, when we answer the question "Why is lying wrong?" with "Because it treats people as mere means," we subsume the act-type of lying under the more general act-type of treating people as mere means. Since we (implicitly) assume that the latter act-type is wrong, we thereby show the former to be wrong, too. This, Baker claims, is explanatory because it provides us with a specific kind of information: that there is nothing *special* about *this* act-type (lying). Instead, it is just an instance of a more general act-type (treating people as mere means), of which *any* instance is wrong.

The unificationist model of explanation is a serious contender in the philosophy of science, and there is no a priori reason to think that it couldn't be a good model of ethical explanation as well.[33] I cannot discuss its merits here. At this stage, I will simply insist that the unificationist model is not the most natural reading of moral explanatory claims in ethics. Talk about "feature F *making* act A obligatory" or "the rightness of act A *depending* on

---

[32] Compare Fogal and Risberg (2020, p. 174) who argue that the common-sense idea that moral principles are explanatory is neutral with regard to this issue.
[33] Compare Kitcher (1989).

feature F" seems to fit more easily with the grounding model. Furthermore, most metaethicists can effortlessly make room for the grounding model. Both naturalists and non-naturalist realists, as we have seen, can be very naturally understood as making grounding claims. Error theorists, too, can accept that we are making such metaphysical claims; they will just insist that such claims are summarily wrong (at least when the claims are positive). Only the expressivist cannot accept this model of explanation. However, that is not (primarily) a problem for the model. In this regard, I agree with Fogal and Risberg's assessment:

> It might be objected that focusing on metaphysical explanation prejudges certain metaethical disputes, such as whether expressivism is true. Insofar as there's a conflict, however, the problem lies with those views. After all, everyone needs a story about the metaphysical-seeming judgments we make about moral matters, including judgments about what makes actions right, wrong, etc. If expressivists can provide such a story, the seeming conflict disappears. If they can't, that's a problem for their view. (Fogal and Risberg, 2020, p. 173)

In the end, I don't think that what is assumed here speaks against expressivism in any decisive way. The expressivist project in metaethics is best not understood as the claim that there are no metaphysical-looking phrases in moral theorizing. It is rather the claim that we can account for these claims in a much more innocuous way than by postulating metaphysical properties (such as grounds). If such an account were to be successful, the challenge to expressivism would be averted. Yet, at least for now, I will assume that no such account has proved more successful than the grounding model, and since the latter fits our moral discourse in a more straightforward way, I am going to stick with it (at least in the less controversial, *explanation-involving*, sense).

In other words, I propose that the grounding model is a plausible account of the semantics of moral explanation. It is a plausible reconstruction of how many normative ethicists and metaethicists understand and talk about the explanatory side of moral theorizing. This is independent of other misgivings that one might have with it. In what follows, I shall thus often use the grounding model to illustrate how to understand moral explanation.

2.2 The Evidence of Moral Theories

So far, we have been concerned with the first substitution in the analogy, moral theories for scientific theories. We now turn to the second

substitution. What takes the place of the empirical data in the moral realm? I will argue that there are two natural candidates: our *considered judgments* and our *intuitions*. Both ways of conceiving of the analogy between science and ethics have seemed natural to many philosophers, and I will argue that, at least for the purposes of this book, we do not need to decide between the two. Thus, when it comes to the question of what *kinds of entities* figure in the analogy, we can take an ecumenical stance. However, there is a second issue, which is not about what entities figure in the analogy but about these entities' *level of generality*. Here, I will take a more steadfast position and argue that we should consider our intuitions or considered judgments about *particular cases* to be the analog to the evidence in science.

Intuitions and Considered Judgments

Whether there are analogs to the data of science in ethics has been discussed by moral philosophers independently of the issue of moral underdetermination. There seem to be two prime candidates: *considered judgments* and *intuitions*.

The idea that the natural correlate to the data of science is our intuitions has a long history. Indeed, the whole idea of searching for some equivalent to the data of science in ethics might have its strongest historical background in the tradition of intuitionism. One of intuitionism's main advocates states this in exemplary fashion:

> We have no more direct way of access to the facts about rightness and goodness and about what things are right and good, than by thinking about them; the moral convictions of thoughtful and well-educated people are the data of ethics just as sense-perceptions are the data of a natural science. (Ross, 1930, pp. 40–41)

This idea has lived on in the intuitionist tradition, with some modern-day intuitionists expanding it beyond ethics. In this spirit, Audi writes:

> Intuitions are important not only in intuitionist ethical theories but also for philosophy in general. Indeed, even many philosophers who do not speak of intuitions theorize as if they were in part seeking to provide an account of shared intuitions, say about knowledge or obligation. [...] The main point is that philosophy, like any theoretical enterprise, must have data, and intuitions are crucial philosophical data. (Audi, 2008, p. 476)

What makes intuitions such a popular candidate for the position of analog of data in science? There are several reasons, but perhaps the most

important one is suggested by a structural feature of the analogy to science. In science, we typically assume that the way we can test our theories is by checking how the predictions that they make fit the available data. Thus, we have three elements: theories – predictions – data. If we want to draw an analogy to ethics, it is natural to think that the deontic verdicts that moral theories yield take the place of the predictions in science. Yet that gives us only two elements. What's missing is the element against the background of which we test whether the deontic verdicts our theories yield are the correct ones. This is where intuitions come into play.

Here is Boyd, describing the analogy in just this way:

> It appears that in moral reasoning, moral intuitions play the same role which observations do in science: we test general moral principles and moral theories by seeing how their consequences conform (or fail to conform) to our moral intuitions about particular cases. (Boyd, 1988, pp. 184–185)

Intuitions, by making this analogy to science work, have seemed to many ethicists to play an important role in theory choice.

The second candidate is what are called *considered judgments*. This notion goes back to Rawls (1951, 1971) and his outline of the method of reflective equilibrium. Roughly, the method states that in order to justify our moral beliefs, we start from what we think about the matter in a pre-theoretic stage and try to account for it by means of a theory. We do this by mutually adjusting both what we initially thought about the subject matter and the theoretical principles themselves, until an equilibrium is reached.

But not just any judgments we might come up with on a whim are supposed to figure in the process. Instead, only considered judgments qualify. Rawls (1951, pp. 181 ff.) provides a list of conditions that must be met in order for a judgment to count as considered, such as that whoever makes the judgment is not to gain or lose from making it, that they have thought the case through, and that they feel a certain level of certainty about the judgments. Finally, the judgments also have to be such that they should not be the (conscious) result of an application of a principle. As Rawls explains:

> The reason for this restriction will be evident if one keeps in mind the aim of the present inquiry, namely, to describe a decision procedure whereby principles, by means of which we may justify specific moral decisions, may themselves be shown to be justifiable. Now part of this procedure will consist in showing that these principles are implicit in the considered judgments of competent judges. It is clear that if we allowed these judgments to be determined by a conscious and systematic application of these principles,

> then the method is threatened with circularity. We cannot test a principle
> honestly by means of judgments wherein it has been consciously and
> systematically used to determine the decision. (Rawls, 1951, p. 183)

This, once more, ties in well with what I have said about the three
elements in science: theories – predictions – data. When Rawls speaks about
"judgments that are determined by a conscious and systematic application
of the principles," we can think of these as the analogs to the predictions
of scientific theories. These, as Rawls argues, cannot serve as the data (in
the process of Reflective Equilibrium) since that would make the process
of justification circular. Instead, we need a third entity that can take that
place: *considered judgments*.

An Ecumenical Solution

This leaves us with two suggestions for what should take the position
of evidence in ethics: intuitions and considered judgments. Which one
should we prefer? Basically, I don't think that we have to make a decision
because the two suggestions can be made compatible. Yet how exactly to
render them compatible depends on how we understand the notion of an
intuition. Since there is notorious disagreement about this even among
defenders of the notion, we need to look at it in more detail.

The original intuitionists were not always clear about what kinds of
entities they had in mind when talking about intuitions, but it seems that
nowadays, three main competitors have established themselves.[34]

The first understanding of intuitions poses no problems at all when
it comes to compatibility with the notion of considered judgments. On
this understanding, intuitions are just (firmly held) beliefs, that is, doxastic
states with a propositional content. This way of understanding intuitions
renders the difference between them and considered judgments incon-
sequential, since the latter are also just firmly held doxastic states with
a propositional content. Thus, on this understanding, our considered
judgments are intuitions and the two suggestions amount to the same
thing.

Things are a bit more complicated when it comes to the two other
understandings. Both agree that we should not understand intuitions in
terms of doxastic states, yet they disagree about what non-doxastic states
are at stake. Some have suggested that intuitions should be described on
the model of emotions (or, more precisely, they cannot be divorced from

---

[34] I follow Tropman (2014, pp. 178–185) in this classification.

emotions). Others have brought in the distinct category of what they call *intellectual seemings*.[35]

These two accounts obviously stand in contrast to the suggestion of considered judgments due to their non-doxastic status. However, there are at least two ways to reconcile the different notions for our purposes. The first route is to follow Lillehammer (2011, pp. 181–185) and Audi (2015, pp. 62–65), who propose a disjunctive account of intuitions that includes doxastic as well as non-doxastic states. On a disjunctive account, considered judgments can be subsumed under the heading of intuitions, together with non-doxastic forms of intuitions. Everyone could thus agree to take intuitions as the evidence of moral theories, while setting aside the intramural debates within the intuitionist tradition for the purpose of this book. The second way to reconcile the two suggestions is to work with a two-level framework. On that framework, we state the analogy in terms of considered judgments but leave it open for proponents of non-doxastic understandings of intuitions to amend the analogy. The reason this works is that, in principle, everyone can accept the notion of considered judgments. Intuitionists need not deny that we all have considered judgments of the doxastic form. They will simply want to point out that this is not the whole story when it comes to where our justification for these judgments comes from. Thus, regarding how we should understand moral underdetermination, everyone can accept the framing that theory choice can be underdetermined by our considered judgments. Intuitionists will just add that if we dig deeper, we see that the underdetermination actually stems from the fact that these considered judgments are themselves underdetermined by the non-doxastic intuitions that justify our holding these judgments in the first place.

In the rest of the book, I will mostly rely on this second way of reconciling the two suggestions, and there are two reasons for my doing so. First, I agree with Smith (2010, pp. 83–84) who argues that many theorists prefer the notions of moral judgments and verdicts when they are searching for the analogs to the data of scientific theory precisely because they take these notions to be very commonsensical, whereas the notion of intuitions is philosophically more disputed. Putting the analogy in terms of considered judgments is thus less contentious. Second, the notion of judgments is more in sync with much of the literature I am engaging with in this book. Three projects are especially important to my argument: the *consequentializing* and *deontologizing* projects and Derek Parfit's *On What*

---

[35] Compare Roeser (2006) for the former and Huemer (2008) and Stratton-Lake (2016) for the latter.

Matters. All these projects and the surrounding debates are mostly framed in terms of judgments and verdicts, not intuitions. It would thus be highly inconvenient each time to translate the results of these projects into the language of intuitions. Still, my decision to go with considered judgments is entirely pragmatic, and I do not think that much hinges on it. Readers who prefer to put the analogy in terms of intuitions are welcome to amend it, and I am positive that all that will be said can easily be made compatible with this way of putting things.

Why the Particular Level

Having settled on the kind of entity that figures in the analogy – considered judgments or intuitions – we can ask a second, separate question: At which level of generality should these intuitions and judgments be?

We have seen that both Hurka and Boyd take the analogs of the data of science to be verdicts or intuitions about *particular cases*. Talk of *cases* might suggest that what is meant are verdicts or intuitions about specific *act-tokens*. This is somewhat of a simplification, however. Although ethicists sometimes discuss concrete cases, most discussions are arguably not like this. Think about the infamous *trolley cases*. These are often used as a standard example of how rival theories differ with regard to particular cases. However, they are almost never about specific situations. That is, they do not concern a specific situation involving some particular person and some innocent bystanders and a trolley at 8 am on June 16, 1987 in rural Minnesota.[36] Rather, what we usually discuss are *act-types*, albeit of a very high level of specificity (i.e., letting one (overweight) person fall on the tracks to stop the trolley from killing five). Applied to the analogy, it seems that what stands in for the data of scientific theories are not considered judgments (or intuitions) about specific cases but more or less specific descriptions of general scenarios. Still, I think that Boyd's and Hurka's usage is rather innocuous, and we can put these complications aside for the rest of the book. The distinction between intuitions and considered judgments about very specific act-types on the one side, and what we classically understand as mid-level and highest-level principles on the other side, is clear enough. The question thus resurfaces: Why opt for the more specific intuitions and judgments as the analog to the data of science?

---

[36] The complication and the example were brought to my attention by Philip Stratton-Lake in conversation.

There is one way in which this seems to be suggested by the analogy to science itself. It is quite natural to think that it is our observations of particular spatiotemporal events that constitute the scientific data. Analogous reasoning would then lead us to think of particular judgments as the evidence in ethics. However, as Rawls reminds us, this is too hasty. In his early outline of the method of reflective equilibrium, Rawls (1951, pp. 182–183) took particular judgments to be the evidence of theories. Later, however, he changed his mind and proposed to include judgments of any level of generality as data:

> [...] [O]ne does not count people's more particular considered judgements, say those about particular actions and institutions, as exhausting the relevant information about their moral conceptions. People have considered judgements at all levels of generality, from those about particular situations and institutions up through broad standards and first principles to formal and abstract conditions on moral conceptions. (Rawls, 1974, p. 8)[37]

Rawls is surely right that we have considered judgments (and intuitions) about more than just the particular level. Thus we need further reasons if the particular level is to be prioritized in the analogy.

I propose that there are two reasons for this. The first has to do with *ambiguity*. Intuitions about mid- and highest-level principles are ambiguous in a crucial sense. A proposition such as "Lying is wrong" can be interpreted in two ways. First, it might just be a generalized deontic claim, covering all instances of lying. In this sense, the sentence tells us that we are not to do any of the particular acts that fall under this generalization. Second, the proposition might be thought to provide an explanation for why a particular case of not telling the truth should be avoided – "Why should I avoid this act of telling something wrong?" "Because lying is wrong." Because of this ambiguity, when we are presented with the considered judgment (or intuition) "Lying is wrong," we might not always know which it is – a generalized deontic claim or an explanatory claim. Both can be expressed with a claim that seems to have the same propositional content. However, this poses a problem for the idea that our more general judgments can confirm or disconfirm our theories. If the judgments that are supposed to confirm or disconfirm the explanatory claims of our theories are themselves already explanatory claims, then we cannot say that the explanatory claims have received some kind of independent confirmation. Instead, this would be as if our scientific theories received support not by the empirical data but by part of the explanation itself. This is not

---

37 For a discussion, compare Daniels (1979, p. 258).

how we standardly think of theory confirmation in science. Instead, we think of our empirical observations as confirming or disconfirming our explanations. Similarly, philosophers like Hurka and Boyd probably think that it is our considered judgments (or intuitions) about what we should do that confirm or disconfirm our theories about why we should do so. In other words, it is our judgments about the deontic that confirm or disconfirm our moral explanations. Yet, if this is so, then focusing on the particular deontic judgments as the evidence of our moral theories is more straightforward. Particular deontic judgments are in no danger of being mistaken for explanatory judgments since, well, they simply aren't explanatory. To avoid the confusion that can arise due to the fact that more general intuitions or considered judgments can be either deontic or explanatory, it makes sense to consider only considered judgments on the particular level to be the data.

This immediately poses the question: What about our intuitions about explanatory claims themselves? If Rawls is correct, and he surely is, we have intuitions about these claims as well. I will take up this question in the next section. The point here is just that *if* we think of the analogy in terms of the deontic underdetermining the explanatory, then it is less confusing to present the deontic content in the form of particular judgments or intuitions (be they about act-tokens or very specific act-types), since these are not at risk of being confused with explanatory judgments or intuitions.

The second reason to prefer particular intuitions or judgments has to do with *determinateness*. Particular considered judgments are determinate about what we should do in a way that more general principles are not necessarily. Take the above example about lying again. If I accept a particular considered judgment, for example, do not lie to your mother about the the math grades you received, this tells me what to do. In contrast, having a considered judgment or intuition about a more general principle – say "You ought not to lie" – does not necessarily tell me what I ought to do. Why? Because I might at the same time have an intuition about another principle that can override this principle in the specific case. Similarly, having the considered judgment that one ought always to maximize the good does not tell one what to do unless one also has a full theory of the good available. Thus, in most cases, only particular considered judgments are determinate insofar as when we are certain of them, we know what to do.[38]

---

[38]   It has been pointed out to me by an anonymous referee that this might not be true. For example, if a particular verdict entails that I repay a debt to my friend today, I still have many ways to do this (paying cash or using PayPal, doing it in the morning or in the evening, and so on). The particular

Now, this is not to say that *theories* that involve mid- and highest-level principles are not determinate about what we should do. The principle "Maximize the good" together with a full theory of the good does, at least in principle, yield maximally determinate verdicts about what we should do in particular cases. To put it in Dietrich and List's terminology, in principle, the complete deontic content of a person's considered judgments (or intuitions) could be encoded in a set of mid-level principles together with some auxiliary information, such as the person's theory of the good. Thus, the point is not that theories cannot be determinate; the point here is just that the intuition that we ought to maximize the good *alone*, as with the maxim that one ought not to lie, is usually not determinate in this way. In contrast, intuitions and considered judgments about particular cases *are* determinate in this way. This is the second reason to consider judgments about particular cases to be the evidence of moral theories.

In sum, if we think that our intuitions and judgments about what we should do confirm or disconfirm our explanations of why we should do so, then it is more useful to focus on judgments and intuitions of the particular level because they are *less ambiguous* and *more determinate*. This is what I shall do in what follows. However, before I can do that, two general worries about the analogy need to be addressed.

2.3   Complications

The Ethicist's Third Tool

In the last section, I argued that we should think of particular deontic judgments (or intuitions) as the data of moral theories. However, as the quote from Rawls showed, we have intuitions about the explanatory side of moral theories as well. This might be thought to threaten the analogy put forward in this chapter for the following reason. In science, we typically assume that there are two sets of criteria that can be appealed to in theory choice. First, we can ask which theory fits better with the empirical data. If one theory does a better job at this, that is a big advantage that will often be decisive in theory choice. Second, we can ask about the theoretical virtues

---

verdict is thus not determinate in the sense that it fully specifies what I have to do (and when I have to do it). I think that we can grant as much. However, this does not seem to be a serious problem for my claim of determinateness. We might simply reformulate this claim to the effect that particular verdicts are determinate *insofar* as morality is determinate. If morality is quiet about the exact payment modalities, then arguably no more determinateness is required here.

of theories themselves. If one theory is much *simpler* than another, or more *fruitful* in its predictions, that, too, might grant it a decisive advantage.[39]

Both these sets of criteria are presumably also relevant in ethics. If one theory accords better with our considered judgments and intuitions about particular cases, that is an advantage. If one of two theories offers a much simpler explanation about why these judgments are correct, that is a second advantage. So far, the case is analogous. However, in ethics, we might think that there is a third set of criteria that we can appeal to. That third set consists of precisely the kind of mid- and highest-level intuitions about the more general principles that Rawls points to. This fact seems to provide ethicists with a *third tool* in theory choice.[40] If true, this constitutes a disanalogy with science and, more problematically, it could undermine the whole idea that underdetermination of moral theories by our considered judgments is something we should care about. For why should we care if moral theories are underdetermined by our considered judgments about particular cases if we also have direct intuitions about the general principles of such theories?

I think that this worry hints at something true. There is indeed a disanalogy to be found here. However, first, the disanalogy is not as fundamental as one might be tempted to think. Second, there is a way to do justice to this topic while at the same time upholding the analogy to science to a sufficient degree.

The reason why the disanalogy is less deep than one might think is that the case is not so clear cut in science either. Consider, as an example, the famous dictum by Einstein (1926) that "[...] He [God] does not play dice." Einstein is here expressing his misgivings with an indeterministic interpretation of quantum mechanics. But what kind of misgiving is being expressed here? One might think that it is about a formal feature of (an interpretation of) quantum mechanics, the fact that it gives too much room to probability. However, putting it like that is somehow misleading. Einstein's complaint is not formal in the way it would be if he objected that the theory is not simple enough, or too ad hoc, and so on. Rather, it expresses a thought that this theory just cannot be an accurate description of the underlying structure of the world. Thus, at least on one plausible reading, Einstein is here objecting to a scientific theory on the basis that

---

39 We will consider a dissenting view about this when we discuss Bas van Fraassen's view in Part III. For now, I will assume that most philosophers of science are willing to grant theoretical virtues at least some evidential force.

40 I am indebted to an anonymous referee for pushing me to say more about this issue as well as suggesting the term "third tool."

the way it explains the world just does not fit with some intuition he has about how the world could possibly be. In other words, Einstein seems to think that there is a third tool in theory choice available to him.

If that is the case, the disanalogy to ethics does not run as deep as one might think. The difference here is not one of kind. Scientists, too, have direct intuitions about what constitute right (or wrong) explanations, rendering the situation more strictly analogous to ethics than it seemed at first glance. That said, I do think that there remains a disanalogy of degrees regarding the preponderance and strength of such direct intuitions about explanations. Many of Einstein's peers and many philosophers of science do not share Einstein's intuition and arguably do not profess to have such intuitions about quantum mechanics or most other scientific theories at all. This might make us hesitant to give too much credence to intuitions like Einstein's. Even if we occasionally find people expressing very strong intuitions about scientific explanations, the fact that others seem not to have such intuitions casts doubt on the reliability of such intuitions. In contrast, most ethicists (like most people in general) do have very strong intuitions about what makes acts right or wrong. In this sense, the third set of criteria, direct intuitions about what makes an explanation plausible or not, seems to be much more common. Arguably, this means that such intuitions should also carry more weight in theory choice in ethics. We might thus worry that, even though the example of Einstein shows that the analogy remains formally correct, the substance of the analogy is still threatened because intuitions about explanations occupy a much more important place in ethics than in science, and the analogy does not do justice to them.

I am positive that these worries can be dealt with, for two reasons. However, since these reasons involve issues that can only be outlined in more detail at a later stage, I will only be able to provide a rough sketch of them here.

The first reason has to do with the fact that most ethicists do not just have primitive intuitions about explanation. Instead, when pressed to give reasons *why* their explanations are supposed to be better, ethicists will point to formal virtues of these theories, such as *simplicity* or *fruitfulness*. This, as I have already pointed out, is perfectly analogous to science. Indeed, underdetermination in science is often seen as a phenomenon that prompts scientists to consider what factors other than empirical adequacy guide their choice of theories.[41] Yet, typically, we do not consider our intuitions about

---

[41] Compare Carrier (2011) for this point.

what would constitute the best scientific explanation part of the data itself. Instead, we bring in considerations about virtues of theories to distinguish between good and bad explanations of the data. In the same way, Section 5.2 will provide an in-depth look at the theoretical virtues of moral theories. Thus, as far as intuitions about what constitutes a good explanation can be backed up with theoretical virtues, we can shoehorn our discussion of these intuitions into the discussion of the theoretical virtues of moral theories. This is just how underdetermination is discussed in the philosophy of science and thus poses no threat at all to the analogy.

Yet what about intuitions that ethicists might have about what the correct explanations are that cannot be captured in terms of theoretical virtues? This brings us to the second reason. Some ethicists might insist that they have independent and direct knowledge about what constitutes the correct explanations in ethics. What are we to do with these intuitions? Here, I think that a firmer stance is called for. In a nutshell, I think that given persistent peer disagreement in ethics about which explanations from which moral traditions are right, insisting that one knows the correct answer *without* being able to defend this by pointing to either deontic or theoretical advantages of that explanation simply does not cut it. Such unsubstantiated intuitions or considered judgments should not be taken to have any evidential import. This, of course, is a highly controversial claim, and I will have to do much more to defend it in Part III of the book. For now, however, I hope that what I have said is enough to convince the reader that the idea of the ethicist's third tool does not undermine the analogy to science.

Theory-Ladenness

The second complication concerns one of the most vigorously discussed topics in the philosophy of science for several decades. I have so far relied on a very strong distinction between scientific or moral theories on the one side, and the empirical data or our considered judgments on the other. The latter confirm or disconfirm the former. For this to be possible, one might think, the latter need to be given to us in an independent way or else we face a threat of circularity. Yet many philosophers of science have come to doubt this simple picture. They have instead argued that the evidence itself is *theory-dependent*, our observations *theory-laden*.[42] If this is the case, the relation between evidence and theory is not so straightforward.

---

[42] Compare Boyd and Bogen (2021) for an in-depth treatment of the topic.

Empirical observations cannot serve as an independent arbiter to confirm or disconfirm theories, because they are tainted by theoretical assumptions.

It is easy to see how the same complication might occur in ethics. Whether we think that a particular act is right or wrong might depend on our more general moral beliefs, such as whether we think that consequences are all that counts or that we may never treat a person as a mere means, and so on. If that is the case, we might wonder whether putting things in terms of underdetermination even begins to make sense. My answer to this challenge has two parts.

First, it is important to note that the idea of theory-ladenness of observation is not *in principle* opposed to the idea of underdetermination. On the contrary, the former might even support the latter. As we learned from Douven (2008, p. 293) in Chapter 1, on the most basic understanding of underdetermination, one set of propositions underdetermines another set of propositions if knowledge (or justified belief) of the first set is insufficient to provide knowledge (or justified belief) about the second set. Yet theory-ladenness might lead to exactly such a situation. Since the data is tainted by theoretical assumptions, it cannot confirm or disconfirm a theory, and the theory is thus underdetermined by the data. Considerations about theory-ladenness might thus be brought in to support claims of underdetermination. Indeed, that is what Duhem does when he argues against crucial experiments on the basis that many observations in science involve the use of instruments that can only be interpreted while making theoretical background assumptions. Similarly, van Fraassen (1980, pp. 80–83), who will play a central role in Part III, treats theory-dependence as an indisputable matter of fact without thinking that this poses any problems for his own treatment of underdetermination. In the same fashion, proponents of underdetermination in ethics might marshal considerations about theory-ladenness in support of their arguments. The fact that our normative judgments about particular cases is dependent on which theory we accept makes it difficult to confirm (or disconfirm) moral theories. If we think that our justification to believe in moral theories depends on this support, the worse for moral theory.

However, and this is where the second part of my answer starts, theory-ladenness might challenge the importance of moral underdetermination in another way. Here is why. There are a variety of reasons why one might think that we can't be sure which moral theory is the correct one. One might be a particularist, thinking that there are no true general claims in ethics (deontic or explanatory). One might think that there is too much disagreement regarding our particular moral beliefs such that no theory can

account for them all. One might even think that there can be no knowledge about particular judgments themselves. Moral underdetermination is an independently interesting phenomenon, or so I think, because it does not rely on any of these (rather controversial) assumptions. Instead, it assumes that even if we all agreed about particular cases, could have knowledge about particular cases, and could even come up with general explanatory principles, there would still be the problem that we do not know which explanation is correct. Theory-ladenness might call into question whether this case is indeed as special and interesting as I am making it out to be. Since the moral data is tainted by theory anyway, one might argue, it makes no difference whether theories agree or disagree about what we should do. We are unable to choose between theories in any case. Put more simply, in outlining the challenge from moral underdetermination, I make a big deal about what a problem it is for ethicists if it turns out that they arrive at the same conclusions about what we should do and we therefore cannot decide between the theories on the basis of their conclusions. Defenders of theory-ladenness might shrug their shoulders and reply that due to the dependence of moral verdicts on theories, the idea that we could confirm or disconfirm theories based on the moral data was an illusion anyway, and we needn't have gone through the whole underdetermination argument to know that.

Thus, for the purposes of this book, it is important to show that theory-ladenness does not undermine the importance of considerations about underdetermination. Fortunately, the philosophy of science offers some help in this regard because it can clarify the issue as well as identify what exactly it is about theory-ladenness that threatens the described picture of theory confirmation.[43] A first issue concerns the fact that theory might guide us when it comes to deciding what to investigate. Depending on our theory, we might focus on data that supports our framework and neglect data that contradicts it. This is a real issue when it comes to developing research agendas (and policies). However, it arguably does not pose a deep theoretical problem for the following reason. It might well be that Kantians or utilitarians prefer to discuss particular scenarios that are more congenial to their theories. For example, utilitarians might prefer discussing so-called *emergency cases* where the consequences at stake are so high that the importance of duties like promise-keeping and truth-telling seems to vanish in comparison. Kantians, by contrast, might want

---

[43] My discussion here closely follows the treatment of the topic by Godfrey-Smith (2021, pp. 203–209).

to focus on scenarios where the outcome with the overall higher utility involves some despicable act, such as the torturing of an innocent person. In this way, their theoretical frameworks might bias ethicists. However, this is clearly not a deep theoretical problem, and in a combative discipline like academic ethics, we can rest assured that theorists from opposing traditions will keep each other honest by posing counter examples and thus forcing each other to take all our considered judgments and intuitions seriously.

A second issue with theory-ladenness concerns language. The way in which we describe our observations often already presupposes a theoretical framework. Thus, theories not only guide us in where to look; they also influence us in how we describe our perceptions. This seems more threatening. How should we choose between rival theories if the language in which we describe the observations that should decide between the theories is already biased? However, it depends on the details. First, this kind of theory-dependence is uniquely threatening for philosophers who assume that we can neatly distinguish between theoretical parts of language and mere observational ones. That is why theory-ladenness has come to be seen as a grave problem for logical positivists. Yet many philosophers nowadays find the logical positivists' semantic program untenable anyway and have developed alternative accounts of scientific language. This includes modern empiricists, such as van Fraassen, who, as already mentioned, take theory-ladenness as a matter of fact. Second, the mere fact that observations are put in theoretical terms does not necessarily inhibit them from speaking for or against a theory. Godfrey-Smith (2021, p. 206) recalls the example of J. B. S. Haldane who, when asked what could convince him to give up the theory of evolution, is said to have quipped "Precambrian rabbits." What the answer illustrates is that an observation can speak decisively against a theory even if it is couched in terms of that very theory.

The case in ethics is analogous. First, as we shall see shortly, proponents of moral underdetermination need not buy into an implausible semantics of moral claims. On the contrary, as I will argue in Part III, moral underdetermination invites us to focus on epistemological questions instead of semantic questions. Hence, proponents of moral underdetermination are in no danger of falling into the analogous trap that logical positivists arguably did. Second, what I have said about the possibility of expressing an observation that refutes a theory in the terms of that very theory seems applicable to ethics as well. If a utilitarian is presented with a case in which the option that has the most overall utility is also clearly wrong, that speaks against their theory. Similarly, if we can come up with scenarios where

treating someone as a mere means seems perfectly harmless, that should make Kantians rethink their commitments. The fact that we use a certain framework to express our commitments does not mean that we cannot be honest about what is entailed about our theories.

The third and final issue is potentially more threatening. It concerns the possibility that the acceptance of a theory directly influences what we perceive in the first place. Standard examples to illustrate this possibility are provided by psychological research such as the famous *Müller–Lyer illusion*, which seems to show that some of our background beliefs or theories can directly influence how we perceive (the length of) some object. This, finally, seems to pose a real theoretical problem. If a theory predetermines what we observe, then it looks viciously circular to refer to these observations in order to confirm or disconfirm *that very* theory. At the same time, this last formulation already points to a possible reply. The problem only arises if our observations are contaminated by *the same* theory that we try to confirm or disconfirm. However, this does not seem to be the case for most advanced scientific theories. What's more, we have a good scientific understanding of how such illusions lead us astray, and they do so in a very predictable and stable way (knowledge of the illusion does not change it!). This means that we can incorporate our knowledge about the workings of such illusions into our overall framework and thus correct our observations insofar as they are leading us astray. Even in its most pernicious form, it is thus not clear at all that theory-ladenness renders observations useless and therefore requires us to completely overthrow our picture of theory confirmation.

How does this translate to ethics? In this final regard, I think that the situation in ethics is less clear. Comparison to the philosophy of science helps us understand why. As we just saw, arguably the most promising way to block the threat from theory-ladenness in its most radical form is by showing that it is not the theories that we are trying to test that influence what we observe. The same is not true in ethics. Here, the very theories that we are trying to test, that is, consequentialism, Kantianism, contractualism, can influence our considered judgments about particular cases. Kantians might find some act reprehensible precisely because they perceive it as an instance of treating someone as a mere means, whereas consequentialists might see no harm in such an act and therefore not perceive it as wrong in the first place. If this is so, we might wonder whether our considered judgments or intuitions can ever serve as independent criteria in theory choice. The problem of theory-ladenness might thus indeed be more serious in ethics than in science.

Yet, even if we accept this, as I am inclined to do, it doesn't follow that theory-ladenness poses an insurmountable problem to the idea of confirmation in ethics. This is for two reasons. First, the situation with regard to our intuitions is not as bad as the previous paragraph makes it seem. Everyday debates in normative ethics attest to this. If our considered judgments and intuitions were wholly determined by the moral theories we hold, we would not expect to see the kinds of arguments we see in ethics all the time. When presented with seeming counter examples, ethicists could always put their feet down and just deny the force of the counter examples. Instead, as we shall see in Part II, ethicists go out of their way to try to accommodate seemingly problematic intuitions. What's even more important is that ethicists often acknowledge that they *feel* the pull of some counter example. They might then decide to bite the bullet or try to explain away the intuition. However, the fact that they acknowledge that intuition in the first place is a clear sign that they do have intuitions about cases that are not predetermined by their theoretical framework. Hence, even though the situation is more complicated in ethics than in science, it is an exaggeration to claim that our initial judgments and intuitions can play *no* role when it comes to deciding which moral theories are correct. To the degree that they do, the problem of underdetermination remains pertinent.

The second reason not to be too worried about theory-ladenness is that there are forms of underdetermination that are less impacted by it. One of those is especially important for my argument: underdetermination between *extensionally equivalent* theories. Theory-ladenness provides a problem for theory confirmation insofar as proponents of different theories might not even agree regarding their considered judgments or intuitions that are supposed to confirm or disconfirm their theories. Depending on which theory one accepts, one might have different intuitions about particular cases, and this disagreement might make it impossible to decide which theory is true. Yet what if there is no such disagreement? As we are about to see, a number of philosophers have recently argued that theorists of different moral traditions can come up with theories that are *deontically equivalent*, that is, yielding exactly the same verdicts about all particular cases. These theories would be underdetermined by their whole *deontic content*. Thus, no matter which initial judgment someone holds based on a particular moral theory, some other moral theory might be able to account for that judgment as well. In such a situation, it arguably does not matter whether our initial judgments or intuitions were influenced by the theory we accept.

Summing up, theory-ladenness is a phenomenon that should be taken seriously in ethics, perhaps even to a higher degree than in science. However, it does not render discussion of moral underdetermination obsolete. On the contrary, I shall now make the case that moral underdetermination is a pervasive and highly interesting phenomenon that deserves more scrutiny.