CHAPTER I

Underdetermination in Science

What philosophers of science discuss under the heading of the *underdetermination of theory by evidence* is a mixed bag of ideas, expressed in a variety of theses, which are themselves supported and contested by an even wider array of arguments and counter arguments. Fortunately, we are not interested in the debate about scientific underdetermination for its own sake, but only insofar as it might inform our understanding of ethics. Accordingly, the goal of this chapter is to provide an overview of the discussion with a certain bent. That is, we can focus on those aspects of the underdetermination debate that are of sufficient generality to be of interest for other realms as well and sidestep discussions that are too narrowly concerned with specific issues in science.

Moreover, it is important to keep in mind that the thesis of this book is not that *because* science and ethics are of a sufficiently similar nature, and there is underdetermination in science, it follows that there is underdetermination in ethics. What I am offering is not a companion-in-guilt argument. Instead, to borrow an expression from Leibowitz and Sinclair (2016, p. 4), I employ what might be called a *companions-in-illumination* approach. Since I want that illumination to reach as far as possible, it would not be helpful to simply identify the most plausible version of the thesis in science and apply it to ethics. The nature and scope of the phenomenon might be rather different in the two realms, and some arguments that are more convincing in science might be less so in ethics and vice versa. Instead, the goal is to learn as much as possible from the scientific debate, to let this inform the discussion in ethics.

1.1 Introducing the Idea and Its Earliest Forerunners

Before we get to the systematic issues, I want to provide an intuitive access to the topic of scientific underdetermination. I start with the basic structure. I then look at some examples, ranging from everyday situations,

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

to the standard examples from physics, to examples from other domains of science. Following this, I give some historical context to the idea by discussing its two most important early proponents, Pierre Duhem and W. V. O. Quine. As we will see, many of the distinctions that have structured the later debate have their roots already with these two authors.

The Basic Structure and Some Examples

We can all readily remember situations in which our available evidence was not enough to decide between competing explanations. Imagine you come home after a day at work and, to your surprise, you find your balcony door open (you always close it for fear of rain). Different explanations immediately spring to mind. Maybe you simply forgot to close it; maybe a strong wind forced it open; or maybe your partner arrived earlier from work and opened it. Or consider a person at a train station waiting for a delayed train. While waiting, this person develops several hypotheses on why the train might be delayed, such as problems with the engine, a staff shortage, or a signal failure. Situations like these abound, and they are ideal to illustrate the basic structure of underdetermination.[1]

This structure can be characterized as follows. First, as Douven (2008, pp. 292–293) explains, in the most general sense, underdetermination concerns a specific epistemic relationship between two distinct classes of propositions. One class – the propositions about the evidence – underdetermines a second class – the theories/explanatory propositions – if and only if knowledge of all the members of the first class is insufficient for knowledge of the members of the second class. My observation that the balcony door is open does not tell me whether the correct explanation is that I forgot to close it, a gust of wind opened it, or my wife did so. Second, as Quine (1975, p. 313) observes, typically, the relationship between the two classes has a a directedness: There is an inference relation, but it only goes in one direction. Belief in a theory makes us expect certain observations because the theory entails predictions about observables. In contrast, the observations do not imply the theory, since they are compatible with alternative theories.[2] If the person at the station knew that there was a signal failure, that would make them expect that the train will not arrive. But merely observing the delay does not tell them the cause. Third, underdetermination (at least if it is to be of an interesting sort) requires that

[1] The first example is mine; the second is from Ladyman (2002, pp. 162–163).
[2] Compare also Kosso (1992, p. 87).

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

there is a *radical* conflict between the theories, meaning that they cannot be true at the same time.[3] If I forgot to close the door, then my wife didn't open it when she got home. If there was a staff shortage and the train never left for its destination, then the fault does not lie with a signal failure.

Whereas the first two conditions are relatively easy to assess in a specific case, the third is more difficult. What we want to exclude, if the underdetermination at hand is to be of an interesting variety, are cases where the rival explanations are either explanations at different levels or merely notationally different.[4] Yet, this isn't always easy to establish. There may be highly sophisticated ways in which we can translate the predicates of one theory into those of its rivals, so that seemingly radical disagreements vanish. These translations need not be obvious. Depending on the subject matter, and considering our state of knowledge, we might not be able to reconcile the relevant predicates, although an idealized science would be able to do so. Thus, as we shall see throughout the book, whether two theories are indeed radically different is a topic of much contention. However, it is also of the utmost importance, because the really interesting cases of underdetermination are of the radical variety.

As a first approximation, we can thus state the basic structure as follows. Underdetermination is about a relation between (a) two classes of propositions – one specifying the evidence and the other constituting the theories – which (b) has a directedness, such that knowledge of the latter would make us believe the former but not vice versa, and where (c) there is a conflict between members of the second class that amounts to more than them giving explanations on different levels or being notationally different.[5]

[3] Authors aren't always precise when it comes to what the radical conflict amounts to. Often it is expressed in terms of *logical* incompatibility. Two theories are logically incompatible if they entail propositions that amount to a contradiction if combined. However, it is not necessary that differences are of a logical variety to render the disagreement radical. Theories could also be incompatible in a merely *metaphysical* way (i.e., that they cannot both be true in any possible world) or a *nomological* way (i.e., that they cannot both be true under the laws of nature).

[4] An example of the former comes from the philosophy of mind. We might not be able to reduce psychological explanations to physiological ones. However, at least according to some respectable philosophies of mind, this does not render the two explanations incompatible. An example of the latter involves the two temperature scales Celsius and Fahrenheit. Although two theories using either Fahrenheit or Celsius will make notationally different predictions about temperatures, it would be far-fetched to think that they disagree in any major sense. In fact, it is not even clear whether we should still speak of the theories being different when it comes to such examples. I borrow this second example from Portmore (2011, pp. 109–110). We will come back to it.

[5] Note that this characterization is very broad. It allows us to characterize some classical philosophical problems, such as Descartes' Evil Demon, as cases of underdetermination. But this is not necessarily a bug of the definition. Instead, it just shows that even though underdetermination has mostly been discussed in the philosophy of science, the epistemic situation underlying it is much more widespread. Compare Ladyman (2002, p. 167) and Stanford (2017, p. 2).

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

The two examples above can help illustrate this structure. However, apart from that, the philosophical interest of such mundane examples is rather limited. They are typically due to one person's epistemic restrictions at a specific time and place, and other people might already know the correct explanation (my partner, the train driver).

What makes for philosophically interesting cases are situations where all the evidence available to the scientific community does not suffice to determine between alternative scientific theories. In principle, such situations can arise in several ways. For one, different theories might make predictions that match equally important but non-overlapping subsets of the evidence. For another, and more realistically, different theories may make predictions that have a significant overlap while remaining on par with regard to the predictions that don't overlap. In both cases, theory choice is underdetermined by the evidence. Yet the situation that has arguably attracted the most attention in the literature is the one where we are facing *empirically equivalent* theories. Theories are empirically equivalent if they make the same predictions about observables. This situation is of particular epistemological significance since it precludes the possibility of new data becoming available that would tip the balance in favor of one of the theories.

Having characterized the structure like this, let us next consider some examples. Most of the classical instances are drawn from the history of physics. A widely used one concerns the rivalry between Copernican and Ptolemaic astronomy. Both views were once thought to be able to account for all the available observations of planetary motions. They made similar predictions about when which celestial body would appear in which region of the sky. Yet they did so by assuming very different trajectories for the planets. More so, they differed fundamentally on the architecture of the solar system: One put the sun at the center, the other put the earth there. Of comparable prominence is the case of particle and wave theories of light. For a considerable time, there was an ongoing debate about the nature of light: Does it consist of discrete particles, or should we think of it as a wave? Observations were inconclusive, until certain experiments in the nineteenth century seemed to conclusively prove the wave hypothesis. Yet the twentieth century saw another change of direction, introducing the idea of a wave-particle duality. Again, the data at different historical stages were insufficient to decide between what are radically different stories about the nature of some phenomenon, in this case the nature of light.[6]

[6] Both of these cases already serve Duhem (1906) as examples for underdetermination.

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

More recently, some authors have called for the scope of investigation to be widened from physics to other sciences. Stanford (2006, pp. 51–79), speaking out against the bias for examples from physics, has looked into the history of biology. He considers theories about inheritance and the generation of organisms. For a long time, the two competing approaches were mechanistic and vitalistic; that is, they saw reproduction, inheritance, or growth as explainable either by mechanical processes or by vitalistic forces irreducible to mechanic interactions. Stanford argues that, at several stages, there were alternatives to the dominant hypothesis of the time, which later theorists accepted as at least equally well supported by the data. Tulodziecki (2013, pp. 3734 ff.) explores the history of medicine, specifically alternative hypotheses on the causes of cholera. Since cholera behaves markedly differently to some other contagious diseases, the hypothesis that it is itself communicable was not generally accepted for a long time. Instead, competing hypotheses included the idea that cholera is caused by the inhalation of decomposing organic matter or that it is heritable. Tulodziecki argues that prior to the discovery of the bacterium *vibrio cholerae*, the data alone were insufficient to adjudicate between those theories. Finally, Bortolotti (2008, pp. 100–102) brings up an example from primatology studies. Researchers in primatology have debated whether primates have a *theory of mind*, that is, whether they ascribe unobservable mental states to other beings. Now consider a test: A chimpanzee is put between two trainers, one facing it and the other turning away. The experiment shows that chimpanzees, when hungry, would consistently make their begging gestures toward the trainer facing them. This was taken to corroborate the theory of mind hypothesis, because it would seem to suggest that the chimpanzees understand that the trainer has to see them in order to make the decision to feed them. However, Bortolotti argues, the chimpanzees might also know from past experience that a trainer who faces them is more likely to feed them. Both hypotheses thus seem compatible with the evidence.

These and similar examples should make clear that underdetermination abounds. It ranges from ordinary life, to physics, to most other areas of science. But how did the topic enter philosophical discussion in the first place? Two authors have played a decisive role.

Two Progenitors of the Idea

Underdetermination is invariably linked to two names: Pierre Duhem and W. V. O. Quine. Such is the importance of these two that there has

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

sometimes been talk of the *Duhem–Quine thesis*.[7] However, as much as this designation is a deserved mark of reverence to the undeniable importance of both these authors' contributions, it is inappropriate when used to refer to a specific thesis. Even the two earliest proponents of the idea of underdetermination understood it in a markedly different way.[8] Looking at them in some detail will help us get a sense of where different strains in later discussions have their origins.

Duhem's First Treatment of the Idea of Underdetermination

The *locus classicus* for the underdetermination idea, if there is one, is commonly thought to be Pierre Duhem's 1906 book *La Théorie Physique: Son Object, Sa Structure*.[9] Duhem, a French physicist and philosopher as well as historian of science, was the first to discuss the phenomenon extensively and to assign it a crucial role in his overall view of science.

A good way to approach Duhem's ideas about underdetermination is to start with his methodology. Duhem (1906, pp. 190 ff.) posits himself in strong opposition to *inductivists*. Adherents of that method, which Duhem calls the *Newtonian Method*, only allow for hypotheses that are directly suggested by generalization from experience. Scientists start by looking at the evidence at hand and, from there, are led to hypotheses, which they test against new evidence.[10] As Duhem (1906, pp. 180–183) acknowledges, this might work for those forms of science that largely deal with commonsensical subject matter, his example being physiology. However, the method is not practicable when theories become symbolic and involve an increasing level of mathematization, such as in the case of physics. The inductive method would shackle such sciences, Duhem thinks, since induction alone does not yield symbolic mathematical theories.[11] Instead, he advocates for the *hypothetico-deductivist* methodology: Scientists should be free to come up with hypotheses on their own; they can generate them out of thin air, so to speak, as long as they check how well the hypotheses accord with the

7 See Harding (1976).

8 Since our primary interests are systematic rather than historical, I do not consider too many exegetical issues. Instead, I focus on both authors' most classical and influential statements of their view and only bring in additional literature where it aids our understanding. In Duhem's case, this means that I am exclusively concerned with what is widely considered the definitive statement of his ideas in his 1906 book *La Théorie Physique: Son Object, Sa Structure*. In Quine's case, the main source is his first treatment of underdetermination in his 1951 paper "Two Dogmas of Empiricism." However, I also consult some of his later texts, since Quine would frequently return to the topic of underdetermination, and his subsequent specifications are points of reference for much of the philosophical debate.

9 See Stanford (2017, p. 3).

10 Compare Ariew (2014, p. 11).

11 Compare Darling (2002, pp. 517 ff.).

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

results of scientific experiments.[12] For Duhem, theories thus have a much more active role in leading experiments, with scientists coming up with ever new hypotheses.

However, as Duhem is quick to notice, this leads to a new problem. What if we come up with different hypotheses that can account equally well for the observations, but are themselves incompatible? This, of course, is the problem of underdetermination.[13] And Duhem thinks that it describes precisely what physics faces. In his view, there is no guarantee that experiments will yield a definite verdict on the theories we have come up with. This is for two reasons.

First, Duhem (1906, pp. 183–190) argues that one can never conclusively falsify a single hypothesis. Instead, he accepts what would later be dubbed a *holistic* view of theory confirmation. According to such a view, we never test single hypotheses in isolation, but only in combination with other hypotheses and additional assumptions. Duhem states that:

> To seek to separate each of the hypotheses of theoretical physics from the other assumptions on which this science rests in order to subject it in isolation to observational test is to pursue a chimera [...]. (Duhem, 1906, pp. 199–200)

Instead, whenever we test a hypothesis, we are at the same time relying on a multitude of auxiliary hypotheses. For example, to interpret what our experiments tell us, we need to rely on theories about how the measuring instruments function. Yet this means that when an experiment fails to corroborate the tested hypothesis, we can always blame one of those auxiliaries. In Duhem's own words:

> The only thing the experiment teaches us is that among the propositions used to predict the phenomenon and to establish whether it would be produced, there is at least one error; but where this error lies is just what it does not tell us. (Duhem, 1906, p. 185)

Second, Duhem also attacks the idea of verification.[14] He objects to direct verification on the basis that the results of experiments are always imprecise, and there is always latitude when it comes to translating

12 This portrayal follows Carrier (2011, pp. 190–191).

13 Carrier (2011, p. 191) elaborates more closely on how adherence to hypothetico-deductivism was instrumental to the rise of underdetermination. He suggests that: "[u]nderdetermination is an unintended by-product of the methodological transition from inductivism to hypothetico-deductivism."

14 This follows Darling (2002, p. 514).

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

them into precise theoretical laws. More prominently, Duhem (1906, pp. 188–190) also opposes indirect verification by elimination through *crucial experiments*, that is, the idea that a single experiment can conclusively verify one theory by contradicting all of its alternatives. His reasoning here is that we cannot logically exclude all possible alternatives by means of one experiment:

> Unlike the reduction to absurdity employed by geometers, experimental contradiction does not have the power to transform a physical hypothesis into an indisputable truth; in order to confer this power on it, it would be necessary to enumerate completely the various hypotheses which may cover a determinate group of phenomena; but the physicist is never sure he has exhausted all the imaginable assumptions. (Duhem, 1906, p. 190)

Combining these two arguments with the hypothetico-deductivist methodology all but ensures that we run into problems of underdetermination. The hypothetico-deductivist methodology allows us to come up with ever more candidate theories. Yet Duhem thinks that we cannot be sure that they are right, nor can we ever be sure that the theories that are apparently contradicted by evidence are indeed false. Problems of theory choice thus abound.

What is the scope of the problem in Duhem's view? As a physicist, Duhem focuses on the realm of physics. Physics, as Duhem argues, is fundamentally different from a field like physiology due to its high level of symbolization and mathematization. What the physicist observes in their experiments always has to be translated into the symbolic language of their theories via the use of numerical measurements. However, our measuring instruments only allow this to be precise to some degree, in principle allowing for different translations. Physics is thus especially vulnerable to underdetermination. However, as Needham (2000, p. 116) points out, it is probably excessive to read Duhem's claims as a purely empirical thesis about the peculiarity of physics.[15] Even though Duhem himself is almost exclusively concerned with physical theories and restricts his observations to them, his arguments are at least in principle open to expansion to other fields that are equally dependent on instruments and have a comparable level of mathematization and symbolization. Thus, underdetermination is already a rather far-reaching phenomenon in Duhem's view. But with our next author, this idea would be taken to an entirely new level.

[15] See also Ben-Menahem (2016, p. 262), who thinks that Duhem argues "[...] for the underdetermination of science at large."

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

Quine's Introduction of the Idea to the Analytic Debate

Duhem rightly holds a distinguished position in the history of theorizing about underdetermination. However, the author who made underdetermination a household name in philosophy is W. V. O. Quine. He introduced the idea to a wide audience and went on to give crucial impulses to the debate over the course of several decades. Still, Quine's role is often much more critically evaluated than Duhem's because he went much further with the idea.

The starting point of any analysis of Quine's views on underdetermination has to be his 1951 paper "Two Dogmas of Empiricism" (TDE for short), more precisely, the last section of it. Here we find his earliest and arguably most influential remarks on the topic. Yet for all the importance that has rightly been placed on those remarks, it is difficult to distill from them a clear statement of a thesis. Instead, underdetermination, as Norton (2008, p. 22) explains, entered the modern literature as a quiet by-product of Quine's analysis of empiricism.

Quine's clearest point of connection with Duhem lies in their shared holism, which Quine (1951, p. 41) expresses in the famous line: "[...] our statements about the external world face the tribunal of sense experience not individually but only as a corporate body."[16] Holism, as Quine later defines it, is the doctrine that:

> [...] [S]cientific statements are not separately vulnerable to adverse observations, because it is only jointly as a theory that they imply their observable consequences. (Quine, 1975, p. 313)

This view is epitomized by a metaphor. Quine (1951, p. 42) invites us to think of the totality of our beliefs as constituting a single web. At the periphery, the web impinges on experience.[17] Here, we find beliefs like "There are brick houses on Elm Street." Such beliefs are directly linked to experience; they are *closest* to observation. In contrast, beliefs at the center are only indirectly linked to observation. What connects them to the periphery are logical interconnections, such as entailment or negation.

Importantly, despite the difference in relative distance to experience, there is no difference of kind between beliefs that are closer to the center and those that are closer to the periphery. Quine (1951, p. 43) holds that

[16] Quine (1991, p. 269) later relates that he was only made aware of Duhem's work after having already published an earlier version of TDE in the *Philosophical Review*, which is why the reference does not appear in that earlier version.

[17] Quine does not explain how this impinging on experience works, but much of his later monograph *Word and Object* is concerned with laying this out in more detail. Compare also Quine (1969).

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

no beliefs are immune to revision, down to the rules of logic.[18] He also thinks that we could hold on to any belief at the periphery *come what may* if we are ready to make enough amendments in other parts of the web of beliefs. Quine cites the pleading of hallucination or the adjustment of linguistic rules as examples of how this could be done. As Quine (1951, p. 44) explains, all our beliefs are measured according to how well they function as a: "[...] device for working a manageable structure into the flux of experience." None of them is sacrosanct.

Underdetermination enters the picture precisely *because* the web is only connected to experience at the periphery. Although we must react in some way to recalcitrant experience, it is not predetermined *how* we should react:

> [...] [T]he total field is so underdetermined by its boundary conditions, experience, that there is much latitude of choice as to what statements to reevaluate in light of any single contrary experience. No particular experiences are linked with any particular statements in the interior of the field, except indirectly through considerations of equilibrium affecting the field as a whole. (Quine, 1951, pp. 42–43)

Since those beliefs closer to the center are not directly connected to experience, we can always structure the beliefs of the periphery in different ways, amending or giving up alternative beliefs in the center. We thus arrive at different theories that contain incompatible beliefs at the center, yet between which choice would be underdetermined, because they do not differ in their ability to account for experience. Quine holds that, in principle, this option is always open to us: We can always construe alternatives to our theories by making amendments within the web, while preserving empirical adequacy.[19]

These ideas, especially the one that we can hold on to any belief if we are willing to make enough amendments to the rest of our theory, are more radical than what we find in Duhem. Duhem's argumentation is informed by detailed examples from the history of science; underdetermination arises as a result of the specific mathematical and symbolical nature of physical theories. Even if these arguments should generalize to other fields of science, their limits are still more or less clear. In contrast, Quine relies more heavily on general epistemological, logical, and linguistic considerations. Both authors think that when we test a hypothesis, there are often auxiliaries at play that could in principle be blamed for the tests failing to corroborate

18 Quine (1966c, p. 232) later doubles down on this claim, when he states that: "[i]n science all is tentative, all admits of revision - right down [...] to the law of the excluded middle."

19 Compare also Quine (1966b, p. 241).

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

the thesis. However, whereas Duhem clearly has in mind other scientific hypotheses when talking of auxiliaries, Quine proposes that the rules of logic themselves could be revised if needed.[20] Alternative theories, Quine (1951, p. 43) thinks, can be defended by citing hallucination, reinterpreting words, or changing the laws of logic. Yet these means are available in all fields of knowledge. Quine is, of course, aware of this. He does not limit his claims to any particular theory, or field of theories, such as physics. Instead, when he introduces the metaphor of the web, he explicitly talks of the totality of our knowledge or beliefs. This is not surprising considering his background. Whereas Duhem draws almost exclusively on his own work as a physicist, Quine's background is in philosophy, with an emphasis on logic.[21] Under his treatment, underdetermination thus becomes a ubiquitous problem.

To sum up, although there is sometimes talk of the *Duhem–Quine thesis* in honor of its two pioneers, these authors actually held very different views about most aspects of underdetermination. These early differences laid the ground for a diversification in the ways that later philosophers would argue for underdetermination as well as what form of underdetermination they would argue for. In the next sections, I will consider the most common strategies that have been used to argue for underdetermination followed by an overview of the different forms that the thesis of underdetermination has taken on.

1.2 Three Paths to Underdetermination

In arguing for the claim that scientific theories are underdetermined, philosophers have used at least three broad strategies. Two we have already encountered when discussing the pioneers; the third entered the scene only later. These strategies are markedly different and, as we shall see in the last section, they are employed in the service of arguing for very different theses, too. Yet since our interest does not concern the debate in the philosophy of science for its own sake, I will not try to assess which of these strategies ultimately proves to be most promising. Instead, my hope is that analyzing these strategies will afford us insights into some of the arguments that, contrary to what their proponents might think, lead to the moral version of underdetermination.

20 Compare Pietsch (2012, pp. 88–92) and Needham (2000, p. 110).
21 Quine (1951, p. 44) accordingly refers to himself as a lay physicist.

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

The Inductive Strategy

The most generic way to argue that underdetermination is a thing in science is to simply point to a host of concrete examples and make a convincing case that we can generalize from them. This is one of the strategies Duhem uses when going through a number of case studies from the history of science, especially physics. Duhem tries to show that many of the important theories in physics have rivals that can equally well account for all the evidence. Considering these cases, we are supposed to conclude that they do not merely represent specific problems for specific theories but are rather indicative of a general predicament in which physics finds itself. Adapting this approach to the whole of science yields what we can call the *inductive* strategy. It consists in enumerating important cases of underdetermination from many of the special sciences and then generalizing from these.

Although all inductive arguments share this basic structure of extrapolation from a sample, they need not be that simple, of course. An example of a more sophisticated recent attempt is what Stanford (2006, pp. 17–26) calls the *New Induction*.[22] Stanford argues that we can strengthen the case for underdetermination by taking into consideration the historical record. When we go back through history, underdetermination seems to have been the rule rather than the exception. From this we can construe a meta-induction: Just as underdetermination has been widespread historically, we should assume that it will be so in the future, too. This means that even though many cases of underdetermination are resolved as new evidence becomes available by advancements in the sciences, it is nevertheless rational, based on the historical record, to expect to detect new examples in the future.

What's more, Stanford (2006, pp. 27–37) thinks that we do not have to content ourselves with alternatives to theories that were known to the proponents of the original theories at the time they were underdetermined. Instead, he suggests that we can also refer to alternatives that were *unconceived* at the relevant time. As long as we can convincingly show that the actual choice of theory was objectively underdetermined by some alternative at some time, we do not have to request that someone actually thought of the alternative. It suffices that scientists could have thought of the alternatives and, having done so, they would not have had the means to break the ensuing underdetermination.

This second move considerably widens the base for examples of underdetermination. Still, how do we come up with such examples? Stanford thinks

[22] Compare also Stanford (2001, pp. 7–12).

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

that there is no way around detailed analyses of specific cases. Examples of underdetermination have to be hard-earned, one by one, and Stanford (2006) spends a considerable amount of time leading by example. This work ethic is admirable. However, it also has some obvious restrictions. First, it is simply not clear whether we can find enough examples to support a convincing inductive case. As we saw, Duhem himself was well aware that his arguments are mostly from physics, and accordingly he thought that they only support a thesis for that realm.[23] Second, the present (and past) cases of underdetermination might well establish that the phenomenon is real, but how should we know whether Stanford's meta-induction is valid? Many of the best-known examples from the history of science have already become unusable through additional evidence. Even if there is still a broad sample of examples in the present, which is a contested point, what tells us that when those are eventually solved, they will once more make way for new cases?

Stanford happily accepts these limitations to his thesis. Others, however, have not been so sanguine. Instead, they have come up with strategies that are less piecemeal.

The Holistic Strategy

One such strategy is what we have already encountered in the last section when discussing *holism*. Indeed, the fact that both Duhem and Quine relied so heavily on this strategy in their classical formulations of the underdetermination thesis has led to some authors confusing the strategy with the thesis of underdetermination itself. Yet we need to distinguish the argumentative strategy from what is argued for.[24]

The strategy starts with an observation about hypothesis-testing. Holists are wary of the notion of a single hypothesis tested in isolation – a point that, as we saw, Duhem and Quine repeatedly stress. Instead, the test is always of a hypothesis and a whole battery of other hypotheses and background beliefs. When facing recalcitrant data, the evidence itself does not tell us whether the mistake lies with the hypothesis or one of the additional assumptions. Accordingly, we can make different adjustments in our theories to account for the recalcitrant evidence. This is where underdetermination gains its foothold. Our different reactions to recalcitrant data

[23] Kitcher (2001, p. 195) doubles down on this, pointing out that examples are really mostly from a subset of physics and therefore even less representative of science per se.

[24] Quine (1975, p. 313) does make the distinction, when he states that: "[t]his doctrine of empirical underdetermination is not to be confused with holism." See also Needham (2000, pp. 116 ff.).

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

lead to different theories that might, during the process of adjustment, become incompatible.

A commitment to holism does not necessarily lead to a belief in underdetermination. At its core, it is only a rejection of the idea that we can test hypotheses *in isolation*. This alone says nothing about there being alternatives to our best theories, let alone (radically) incompatible ones. Maybe we can test sets of hypotheses successfully to the point where no alternatives remain.[25] In practice, however, holists have typically accepted underdetermination. This has to do with the fact that holists often mistrust our abilities to rationally choose between rival theories as long as they stay true to the data. As long as we save the phenomena, there is much latitude, in their view, as to which part of our theory we can amend.

The difference from the inductive strategy should be plain to see. According to the inductive strategy, it has to be shown how the choice between two existing theories turns out to be underdetermined. In contrast, according to the holistic strategy, underdetermination arises when one theory is developed in different ways. Nowhere in the former strategy do we find a commitment to a specific picture of hypothesis-testing. For all we know, a proponent of the inductive strategy could be opposed to holism and, at least in principle, a friend of crucial experiments. In contrast, holists need not care too much about existing examples, since their view of theory confirmation ensures cases of underdetermination.[26]

This makes holism appealing to proponents of underdetermination. Holism provides a more principled reason for thinking that underdetermination is ubiquitous than generalizations from a few scattered theories can. If scientific methodology itself is unable to adjudicate between rival reactions to recalcitrant data, we have reason to think that underdetermination is a pervasive phenomenon.

The Algorithmic Strategy

Although the holistic strategy promises a more sweeping defense of underdetermination than the inductive variant, it still does not deliver what some more radical defenders of underdetermination want. Unless we follow

25 Adeel (2010, pp. 20 ff.) comments on this point. He observes that holism is basically just a rejection of the empiricist dogma of reductionism. It does not commit one to underdetermination.

26 Stanford (2017, pp. 3–4) elevates this distinction into one about different forms of underdetermination; *holistic* and *contrastive*. However, we need to be careful here. Although the argumentative strategies are different, the forms of underdetermination which they support do not necessarily need to be.

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

Quine in holding that *any* belief can be held on to no matter what (and conversely any belief can be given up), there is no certitude that holism will provide alternatives to all or even most of our theories. This has prompted some philosophers to take matters into their own hands, so to speak. If science does not provide such alternatives, why not provide them ourselves? The idea is to come up with algorithms that produce empirically equivalent, yet theoretically incompatible, alternatives for every imaginable theory. One example is the following by Kukla:

> Given theory T, construct T2 which asserts that T holds when somebody is observing something, but that when there's no observation going on, the universe follows the laws of some other theory T′. (Kukla, 2001, p. 23)

T2 and T are obviously empirically equivalent, since they agree on anything ever observed. At the same time, they are not compatible on a theoretical level since T2 contains an additional theory T′ whose whole purpose is to contradict T. Furthermore, since T2 and T′ are formulated in the most open way, the algorithm is maximally flexible, thus promising to deliver all that a proponent of underdetermination may want.

Still, many commentators have been skeptical. Some criticisms have been directed at the specific algorithm Kukla advances. As Bonk (2008, p. 171) writes, Kukla's alternative theory has an odd element of observer dependence. Whereas the original theory makes no mention of what is and what is not observed, the algorithmically produced alternative does so without providing a reason for it.[27] But not all algorithms are prone to this kind of criticism. Consider the following one, which Kukla attributes to van Fraassen (1980):

> Given any theory T, construct the rival T′ which asserts that the empirical consequences of T are true, but that T itself is false. (Kukla, 1998, p. 59)

This algorithm, too, provides an empirically equivalent alternative for any scientific theory, which also explicitly contradicts the original theory. Yet nothing about this algorithm depends on the observer's stance.

Other criticisms are of a more general kind, however. Stanford (2001, pp. 11–12), himself a friend of underdetermination, as we have seen, has criticized algorithmic attempts for striking a *devil's bargain*. He holds that such artificially construed theories like T2 and T′ would not be considered serious alternatives by any working scientist. Others have tried to identify more substantial criteria for excluding algorithmically produced theories.

[27] Compare also Norton (2008, p. 26).

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

Laudan and Leplin (1991, p. 463) complain that the algorithmic procedure amounts to *logico-semantic trickery*.[28] They suggest two criteria on the basis of which algorithmic alternatives might be excluded.[29] The first one is *parasitism*. According to Laudan and Leplin (1993, p. 68), a theory is parasitic upon another theory if it depends wholly on the explanatory and predictive mechanisms of the original theory. For whatever the algorithmic alternative is supposed to explain or predict, it needs to make reference to some other theory. The second criterion, *superfluity*, is tightly connected to the first one. As Laudan and Leplin (1993, p. 13) explain it, a theory is superfluous if whatever is added by the algorithm to ensure that the new theory contradicts the original one does not have any empirical consequences. It thus seems uncalled for, since it does not even attempt to advance our knowledge.

As Laudan and Leplin are aware, we need to be careful not to overstress these criteria, so as not to discard some legitimate theories in the process. Certainly, legitimate theories can include parts that are void of empirical content. We can even imagine scenarios in which we want to hold on to a theory due to its apparent success in prediction, while knowing that its explanation is wrong.[30] Thus, the two criteria should probably not be seen as providing sufficient reasons to exclude theories. Still, there is a lesson to be learned: Even though we might not be able to get rid of all theories that meet one or both of the criteria, parasitism and superfluity remain clear disadvantages. If some algorithm produces new theories that *summarily* suffer from these defects, this should make us suspicious.

Norton (2008, pp. 33 ff.) goes a step further. Not only are algorithms a suspicious form of proving underdetermination; Norton thinks that the whole strategy is self-refuting. In his opinion, any such argument that is short enough to be expanded within the usual length of a journal article must miss its goal. The reason for this is that if it takes so little space to make the argument, we cannot preclude that the theories that we come up with are merely notationally different versions of the same theory rather than real alternatives. Norton (2008, p. 40) probably speaks for many critics when he states that: "[a]nything that easy and powerful seems too good to be true." Algorithms, although powerful in producing cases of underdetermination, should thus be treated with caution.

Summing up, we have seen three different strategies to make the case for underdetermination. All of them have their upsides and downsides, and

28 The sentiment is shared by Hoefer and Rosenberg (1994, p. 603).
29 My discussion here follows Kukla (1998, pp. 66–72).
30 See Kukla (1998, pp. 68–77) for these replies.

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

they have been used to argue for markedly different theses. This has led to an increasingly multifaceted picture of scientific underdetermination. In the last section, I want to close the introduction of scientific underdetermination by giving a brief overview of this picture.

1.3 Varieties of Underdetermination

Since "Two Dogmas of Empiricism," the discussion of underdetermination has grown extensively, as even a superficial survey of the literature shows.[31] I will therefore have to restrict my overview of the varieties of underdetermination to what I consider to be the most crucial distinctions. Again, the aim is not to take sides but rather to lay the groundwork for our discussion of moral underdetermination.

Scope

The first distinction concerns the number of theories for which underdetermination is postulated. Let us return for a moment to Duhem and his discussion of singular historical examples. The initial claim that he aims to establish with these examples is that at least for some theories at some time, the data were insufficient to determine which is correct. As we have seen, Duhem then extrapolates from these examples to physics in general. However, what if this extrapolation proves untenable? We can still claim that *some* theories are or have been underdetermined. This gives us the most modest version of the underdetermination thesis:

> *The existence version*: For some scientific theories, there is an alternative theory, which underdetermines theory choice.

In contrast, if the extrapolation from a few theories proves successful, then *all* physical theories are underdetermined. If the scope can furthermore be broadened to other areas of science, as many proponents have claimed, then underdetermination applies to *all* scientific theories. Put another way, there will be no scientific theory that does not have an alternative that underdetermines theory choice. Laudan (1990, p. 271) accordingly dubs this:

> *The non-uniqueness version*: For every scientific theory, there is an alternative theory, which underdetermines theory choice.

Although already very broad in scope, there are still more ambitious versions. The non-uniqueness version only claims that there is at least one

[31] For an excellently concise overview, see Park (2009).

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

alternative to every theory. Yet, as Laudan (1990, p. 271) informs us, there is an even broader version:

> *The egalitarian version*: The choice between all scientific theories and *any* of their empirically adequate rivals is underdetermined.

Whereas the non-uniqueness form is very common in the debate, it is less clear who actually subscribes to this egalitarian form. Laudan (1990, p. 271) thinks that Quine was at some stage committed to the egalitarian version, due to his conviction that we might hold on to any belief come what may.[32]

Local versus Global

The second distinction does not concern the scope of the underdetermination thesis but the scope that the relevant theories themselves have. So far, we have always talked about particular scientific theories, such as theories of light or planetary motion, being underdetermined. This is in accordance with how both Duhem and Quine initially state their views. It is called:

> *The local version*: Some (or all) specific scientific theories are underdetermined by the data.

This is arguably the more common form in which philosophers of science think about underdetermination. However, Quine (1975, pp. 313 and 327) at a later stage introduces a different version. Instead of, or besides, particular theories, might not our whole system of the world, that is, the totality of science itself, be underdetermined? If so, then underdetermination also comes in:

> *The global version*: The totality of our scientific worldview is underdetermined by the data.

The two versions are not mutually exclusive. It might be the case that some, or all, of our local theories are underdetermined *and* the whole system of the world is, too. Which thesis is more relevant is also an issue of debate. Hoefer and Rosenberg think that:

> The problem of underdetermination of theory by evidence is, in its most acute (and therefore most interesting) form, a problem for global theories, or total science [...]. (Hoefer and Rosenberg, 1994, p. 592)

[32] Laudan (1990, pp. 288–291) also thinks that the egalitarian version tacitly underlies many radically social-constructivist claims by some sociologists of science as well as some postmodern philosophers.

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

Okasha counters that:

> Much of the interest of underdetermination stems from the thought that the actual theories currently accepted by the scientific community might be strongly underdetermined; so it is natural to feel disappointed when we are told that the argument only applies to 'global' theories. (Okasha, 2002, p. 313)

Okasha (2002, pp. 309 ff.) further argues that the most promising answers to some of the objections to the idea of underdetermination only work for either the global or the local form, making the overall case for underdetermination more difficult. Interestingly enough, we shall see later that it is not at all clear if and how this distinction can be drawn in ethics. It will be helpful to keep it in mind, nevertheless, since the disanalogy proves instructive as well.

Temporality

A third issue is temporal. I have noted several times that Quine envisioned a very strong form of underdetermination. Witness a characteristic statement:

> Actually the truths that can be said even in common-sense terms about ordinary things are themselves, in turn, far in excess of any available data. The incompleteness of determination of molecular behavior by the behavior of ordinary things is hence only incidental to this more basic indeterminacy: *both* sorts of events are less than determined by our surface irritations. This remains true even if we include all the past, present, and future irritations of all the far-flung surfaces of mankind, and probably even if we throw in an in fact unachieved ideal organon of scientific method besides. (Quine, 1960, p. 22)

As the last sentence makes clear, Quine expects underdetermination to persist even in the face of all the future data that an idealized science might make available. This has been referred to as:

> *The permanent version*: Some (or all) scientific theories are underdetermined by all present and future data.[33]

However, not all formulations go that far. Sklar (1975) suggests that we should more often think of underdetermination as a phenomenon that

[33] Compare Park (2009, pp. 120–121).

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

holds at some stage in history but that can be dissolved when further evidence becomes available. This can be called:

> *The transient version*: Some (or all) scientific theories are underdetermined by the present, but not all future, data.

This version is much less ambitious, and Sklar (1975, p. 381) is convinced that even philosophers who deny the possibility of permanent underdetermination: "[...] are likely to admit that transient underdetermination is a fact of epistemic life." Still, it is doubtful whether we can draw any strong conclusions from this version.

It is therefore interesting to see that there are also intermediate positions. Stanford (2001, 2006) agrees with Sklar that the most common and accepted forms of underdetermination are of a transient nature. But he adds an additional temporal component, suggesting that underdetermination as a phenomenon of scientific practice can be permanent, even if specific theories are not permanently underdetermined. This is the case if, whenever underdetermination between some theories dissolves when further evidence becomes available, new alternatives to the remaining ones are detected. Stanford calls this:

> *The recurrent version*: At any stage in history, some (or all) scientific theories are underdetermined by the present, but not all future, data.

Deductive versus Ampliative

Finally, there is a distinction that, strictly speaking, does not concern theories themselves but the criteria by which they are chosen.[34] Nonetheless, since this issue is of the utmost importance, it makes sense to treat it separately here.

Proponents of underdetermination often talk about different theories *accounting equally well* for the data. Yet there is a crucial ambiguity in that expression. It might mean that two theories *logically entail* the same claims about the data, that is, we can deduce the same propositions from them about the data. Alternatively, it might mean that the theories are *equally well supported* by the data in a broader sense.

The difference matters. Take the standard theory of the earth's formation several billion years ago. Contrast to it the account of so-called *young earth creationists* who claim that the earth was created in 4004 BCE. This is

34 Comparer Bonk (2008, p. 5).

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

not compatible with some of our data, e.g., much older fossils. However, creationists sometimes claim that God built the earth with the purpose of deceiving us about its real age. In principle, their account can thus be observationally equivalent to the one proposed by standard geology. Logically, it is compatible with the same predictions. Yet no scientists, and few non-scientists, would admit that it is thereby equally confirmed by our observations.[35]

The reason for this is that scientists do not usually restrict their notion of confirmation to deductive logic. To think that one cannot single out the correct theory from its alternatives by means of deductive logic alone is not the same as to hold that the tie cannot be broken by additional forms of reasoning. Even though two theories might logically entail the same set of propositions, there can be a huge difference between how the evidence supports the theories in a fuller sense. To put it in a catchphrase, empirical equivalence does not entail confirmational equivalence.

Laudan (1990, pp. 270 ff.) has forcefully argued that we have to differentiate accordingly between what he calls:

> *The deductive version*: Two theories are underdetermined if they entail the same set of propositions about the data.

and

> *The ampliative version*: Two theories are underdetermined if the data supports them equally strongly.[36]

That theories are *deductively* underdetermined means that we cannot decide between them on the basis of deductive logic. That they are *ampliatively* underdetermined means that we cannot do so even if we bring in the whole canon of scientific methods. It should be easy to guess which form of underdetermination is more difficult to establish.

Summing up, underdetermination is a multifaceted phenomenon that has sparked a rich and ongoing debate. It can be illustrated by everyday situations but is mostly discussed in the context of theories of the special sciences. Its two main progenitors, Duhem and Quine, already laid the ground for many debates about how we should think of it. Whereas Duhem defended a limited thesis of underdetermination regarding physics, Quine widened that thesis to all of our knowledge claims and brought to

[35] The example is from Norton (2008, p. 28).
[36] Compare also Devitt (2002, p. 28).

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press

bear considerations from epistemology, the philosophy of language, and logic that go beyond Duhem's historically informed examples. Following Duhem and Quine, proponents of underdetermination have made use of at least three strategies: the inductive, the holistic, and the algorithmic. Each strategy has its own pros and cons. Finally, we have seen that there are many distinctions that mark different versions of the thesis.

Since it has not been my aim to contribute to the debate in science, I have tried not to take sides on any controversial issues. Instead, I have attempted to introduce the reader to what I think are the most general issues that have relevancy beyond the philosophy of science. When adapting the idea of underdetermination to the moral realm, we should keep these distinctions in mind. The plausibility of moral underdetermination will depend strongly on the version defended and the strategy chosen to argue for it. Our survey of the scientific literature has hopefully prepared us for that task. But how exactly does all of this connect to our overall project of investigating underdetermination in the moral realm? In the next chapter, I take the next step toward this goal by transferring the idea of underdetermination from the scientific to the moral realm.

https://doi.org/10.1017/9781009492454.004 Published online by Cambridge University Press