Chapter 9
Explanation and Understanding

Abstract This chapter begins the transition from talking about knowledge in general to talking about scientific knowledge in particular. Although scientific knowledge is itself something that falls under the general account of knowledge, it has special features that are worth exploring. In this chapter the importance of explanation to scientific knowledge is brought to the forefront of the discussion. The nature of scientific explanation itself as well as its relation to understanding is explored in this chapter. It is made clear that good explanations are those which provide understanding of particular phenomena. In addition to examining the relationship between explanation and understanding, this chapter also examines what can make one explanation superior to another. This examination of explanatory virtues is very important because it is common in scientific practice to adopt a particular theory as a result of its being more virtuous than its competitors. Not only is such a practice common in science, it is something that we do routinely in our everyday life.

The two chapters in this part of the book shift from focusing on the nature of knowledge in general to exploring the nature of scientific knowledge in particular. In this first chapter we will take a close look at the primary aim of science and at how we go about achieving that aim. Afterward, in Chap. 10, we will explore how the vehicle for achieving the aim of science helps shed light on how we gain knowledge in science as well as how we gain knowledge in general. Let us get to work on the tasks of this chapter.

It is generally accepted that science has three primary aims: the prediction, control, and explanation of phenomena. Various branches of science emphasize some of these aims more than others. Theoretical physicists exploring the merits of String Theory do not spend a lot of time focusing on controlling phenomena whereas biochemists researching new medicines do focus a lot on controlling various infections and diseases. Despite these differences, it is commonly held that *explanation* is the most important aim of science (Strevens 2006).[1] As the

[1]Although this is widely held, it is not universally so. Theorists sympathetic to constructive empiricism, such as van Fraassen (1980, 1989), are apt to maintain that the primary aim of science is to construct theories that simply fit the observable phenomena—explanations that go beyond

K. McCain, *The Nature of Scientific Knowledge*, Springer Undergraduate Texts in Philosophy, DOI 10.1007/978-3-319-33405-9_9

National Research Council (2012, p. 52) puts the point, “the goal of science is the construction of theories that can provide explanatory accounts of features of the world.” Why is this commonly held to be true? That is, what makes explanation more important than prediction or control?

Before answering the question of the source of explanation’s importance it is worth noting the reason we are using the term “explanation” rather than “scientific explanation” here. Despite the fact that most philosophical theories of explanation are described as theories of “scientific explanation”, it is widely held that explanations in everyday life are roughly continuous with scientific explanations. The latter tend to be more precise and rigorous than our explanations in ordinary non-scientific contexts, but the differences between everyday explanations and scientific explanations are a matter of degree rather than differences in kind (Wilkenfeld 2014; Woodward 2003, 2014).

The close connection between everyday explanations and scientific explanations is not all that surprising when we consider a couple of facts. First, the practice of giving explanations is ubiquitous in our everyday lives. We often explain others’ behaviors, we explain our own behaviors, we explain why various things happened, and so on. For example, we explain why our friend went to the store by pointing out that she wanted some milk and believed the store to be the best place to get milk. We explain why we read a particular article by citing the facts that we found the title interesting and that we had the time to read it. And so on. Second, commonsense reasoning and the methods of scientists are quite similar. Of course, scientific methods tend to be much more exact and sometimes lead to rather surprising conclusions from the standpoint of commonsense, yet the two share many of the same commitments to views of “rationality, truth, objectivity, and realism” (Gauch 2012, p. 33). “Science inherits indispensible presuppositions about the world” from our ordinary, commonsense view of the world (Gauch 2012, p. 33). In fact, there are a number of clinical studies which suggest that “the very methodology that a mature science uses to identify dependencies in the world . . . is in some sense the built-in method that we use in our attempts to understand the world” (Grimm 2009, p. 88).[2] Science is in the business of giving explanations, and it seems that in our ordinary lives we are too.

---

what is observable are superfluous at best; epistemically unacceptable at worst. Two points are worth keeping in mind here though. First, the primacy of explanation in science and science education is widely endorsed in science education reform documents. Take the National Research Council’s (2012) framework for K-12 science education, for example. This framework consists of three dimensions: scientific and engineering practices, crosscutting concepts, and disciplinary core ideas. Explanation figures prominently in the first two of these three dimensions. It is similarly emphasized in other education reform documents such as AAAS (1993), National Research Council (1996, 2007), and NGSS Lead States (2013). Second, constructive empiricists themselves recognize that explanation is important—they simply claim that the goal of such explanations is only to adequately describe the observable world, not to give us knowledge of theoretical entities. So, even constructive empiricists, who deny that explanation is the primary aim of science, can agree that explanation is important.

[2]For the results of some of these clinical studies see Gopnik (1998), Gopnik and Glymour (2002), Gopnik, et al (2004), Gopnik and Sobel (2000), and Steyvers et al. (2003).

Not only is science and ordinary life in the same business when it comes to giving explanations, but it also seems that they conduct their business in a similar fashion. The similarity between explanations in scientific contexts and ordinary life is such that "the tendency in much of the recent philosophical literature has been to assume that there is a substantial continuity between the sorts of explanations found in science and at least some forms of explanation found in more ordinary non-scientific contexts, with the latter embodying in a more or less inchoate way features that are present in a more detailed, precise, rigorous etc. form in the former" (Woodward 2014, p. 1). What is more, in the philosophy of science it is assumed that a successful theory of explanation will be one that expresses what scientific and ordinary explanations have in common. "These assumptions help to explain ... why, ... discussions of scientific explanation so often move back and forth between examples drawn from bona-fide science (e.g., explanations of the trajectories of the planets that appeal to Newtonian mechanics) and more homey examples involving the tipping over of inkwells" (Woodward 2014, p. 1). So, a proper account of explanation should capture both the everyday explanations that we employ as well as the more precise explanations we find in science and advocated for in science education reform documents, (e.g. National Research Council 1996, 2007, 2012). In light of the similarities between scientific and ordinary explanations and the assumptions commonly made by those theorizing about the nature of scientific explanation, the present discussion should be understood to be a discussion of explanation in general—so that both everyday and scientific explanations are included.

Now to the question of why explanation is so important to science. Explanation's purpose, i.e. the goal of explanation, is what makes it the most important aim of science. Successful explanations increase our understanding of the world around us.[3] It is in virtue of this understanding that we are able to predict and control phenomena. So, it is not that prediction and control are not important goals of science. They are! In fact, some are apt to doubt whether a theory is really scientific if it does not yield new predictions.[4] The reason that explanation is primary is that it allows for prediction and control.[5] Consider the phenomenon of a piece

[3]Philosophers with very diverse views on the nature of explanation and understanding agree that the two are closely linked with explanation providing a means to achieving understanding. See for example, Achinstein (1983), de Regt (2009, 2013), de Regt and Dieks (2005), Friedman (1974), Harman (1986), Khalifa (2012), Khalifa and Gadomski (2013), Kim (1994), Kitcher (1981, 2002), Kvanvig (2003), Lewis (1986), Lipton (2004), Moser (1989), Railton (1993), Salmon (1984, 1998), Sober (1983), Strevens (2006), (2013), Trout (2002), van Fraassen (1980), von Wright (1971), Wilkenfeld (2013, 2014), and Woodward (2003). Even Hempel (1965) and Hempel and Oppenheim (1948), who thought that understanding was too subject dependent to be a proper focus of philosophical study, agree that explanations provide understanding.

[4]This is one of the major criticisms that some press against String Theory. For discussion of this criticism see Dawid (2013).

[5]In some ways this view may be a bit simplistic. As we will see later, one might think that explanation and prediction are very similar (perhaps even the same—one is simply backward looking in time and the other forward looking). Also, it is reasonable to think that making predictions is a factor that can make one explanation better than another. Hence, the relationship

of iron rusting. An explanation of this phenomenon is that iron and oxygen have a redox reaction when brought together in the presence of H_2O (either in liquid form or as moisture in the air). This explanation facilitates understanding of why the particular piece of iron rusted as well as why iron rusts in general. This understanding allows us to predict various phenomena. For instance, we can predict that another piece of iron will begin to rust after a time of being left outside in moist air. The understanding gained via this explanation can also help us to control the rusting of pieces of iron. Since we understand that in order for iron to rust there must be moisture, we could control whether a piece of iron rusts by keeping it in a completely dehydrated environment. The understanding provided by the explanation of the piece of iron's rusting allows us to both make predictions and to control phenomena. Without such understanding it is difficult to see how we could manage either of these feats—particularly controlling whether a piece of iron rusts. In light of the role that understanding plays in both prediction and control, it is maybe a bit misleading to even characterize science as having *three* primary aims. As Erwin Schrödinger explains, our entire scientific worldview rests on the "hypothesis that *the display of Nature can be understood*" (1954, p. 90). Perhaps it is more accurate to say that the primary aim of science is to produce understanding via explanations.[6] Using the understanding gained via explanations to yield accurate predictions and to allow for increased control of phenomena are important secondary aims of science.[7]

Recognizing the centrality of understanding to science is an important insight as is the fact that the vehicle by which we typically gain understanding is explanation.[8] However, these insights raise questions. What exactly is understanding? How do explanations, when they are successful, provide us with understanding? Answers to both of these questions will be explored in this chapter. We will begin with the latter question because by coming to better understand the nature of explanation and the sort of information that explanations convey we will be better equipped to explore the nature of understanding.

between explanation and prediction is not completely clear-cut. However, for the present purpose we can treat them as separate in which case explanation is primary.

[6]Given the central importance of explanations in science, it is unsurprising that constructing and understanding explanations is a major goal of science education reform (Braaten and Windschitl 2011).

[7]According to de Regt et al. (2009, p. 1) "In the eyes of most scientists, and of educated laypeople, understanding is the central goal of science . . . it seems commonplace to state that the desire for understanding is a chief motivation for doing science."

[8]Strevens (2013) argues that when it comes to science, understanding can only be achieved via explanations, i.e. scientific understanding without explanation is impossible. de Regt et al. (2009), Gijsbers (2013), Hindriks (2013), and Lipton (2009) disagree. They each argue that there can be cases in which one comes to have scientific understanding without possessing an explanation. It is worth noting that even though they argue that it is possible to gain understanding without having an explanation, none of these philosophers contest the claim that we typically come to have understanding via explanations or the claim that providing understanding is the goal of producing explanations.

9.1 Explanation

As we have noted, explanation is the method by which we typically gain our understanding of the world around us, whether this is in a controlled scientific context or in our everyday lives. Given the importance of explanation to science it is not surprising that there has been much philosophical discussion of the nature of explanation. As a result of this philosophical discourse, there are many competing philosophical theories of the nature of explanation. There are far too many theories of explanation for us to survey here, in fact. So, we will proceed by exploring what Wesley Salmon (1990, p. 12) refers to as the "fountainhead" of contemporary discussions on the nature of explanation—Carl Hempel's Deductive-Nomological Model (D-N Model) of explanation—and some of its permutations and problems.[9] The importance of this account of explanation can hardly be overstated. It is commonly accepted that it shaped the contemporary discussion of explanation since its arrival on the scene in the mid-twentieth century. Further, it is fair to say that all contemporary theories of explanation are in some way or other reactions to the D-N Model and the problems that were exposed for this theory. We will also very briefly discuss some of the major theories of explanation that have arisen after the D-N Model's fall from dominance before settling on a general picture of explanation that will be useful as a working model for our purposes. However, before we begin our examination of various theories of explanation we need to first consider a very important distinction and some general features of explanation. Let us begin with the distinction.

While it might seem obvious that the two are different, it is important to explicitly distinguish between *explanation* and *explaining*. On the one hand, explaining is a particular action that we sometimes perform. Often this is construed as a speech act in which we verbally or non-verbally communicate an explanation to someone else (Harman 1986). On the other hand, an explanation is "something one grasps or understands that makes things more intelligible" (Harman 1986, p. 67). Typically, the thing that we grasp is a set of propositions—in other words, an explanation is a collection of propositions (de Regt 2009; Kim 1994; Moser 1989; Strevens 2006; Woodward 2003). One very rough way to understand this distinction is simply that explanations are the propositions we communicate to one another when we explain things. In other words, explanations are sets of propositions that provide answers to why-questions—the sort of questions we seek to answer when we are explaining things (Lehrer 2000; Lipton 2004, 2009; Moser 1989; Woodward 2014). Another

[9]There were other important supporters of this view of explanation, including Braithwaite (1953), Gardiner (1959), Nagel (1961), and Popper (1959). Nonetheless, as Woodward (2014) notes, "unquestionably the most detailed and influential statement is due to Carl Hempel". In fact, Hempel's work on this theory (1942, 1965) and Hempel and Oppenheim (1948) is so influential that it is not uncommon to refer to the D-N Model as "Hempel's theory of explanation". We will actually follow this tradition and refer to the D-N Model and the Inductive-Statistical Model (I-S Model) jointly as "Hempel's theories".

way to understand the distinction between explanations and explaining is in terms of cognitive outcomes versus cognitive processes. Explaining is a cognitive process which, when successful, yields a particular cognitive outcome. This cognitive outcome, explanation, is one that promotes understanding. Despite this connection between explanations and explaining, the vast majority of theorists hold that explanations and the explanatory facts which compose them are independent of our acts (or even intentions) of explaining (Strevens 2006).[10] It is the nature of explanation that we are concerned with here, not our acts of explaining (conveying explanations to others). The nature of explanation, the exact features a collection of propositions has to have in order to be an explanation, is the question that theories of explanation seek to answer. We will briefly explore some of the answers that have been offered in the subsections which follow.

9.1.1 Hempel's Theories of Explanation

In his vastly influential work on the nature of explanation, Carl Hempel proposes two theories of explanation, the D-N Model and the Inductive-Statistical (I-S) Model, though the D-N Model is the most widely discussed of the two (perhaps because Hempel considered the explanations it offered far superior to those of the I-S Model).[11] These two models share some features. First, both adhere to the majority view that explanations are sets of propositions. Second, both construe explanations in the form of an argument where a set of propositions (the *explanans*) provide support for a proposition that describes the phenomenon to be explained (the *explanandum*). Third, both require that the explanans include at least one proposition which describes a general law of nature as well as propositions that accurately describe the empirical conditions relevant to the phenomenon's occurrence (empirical conditions which fall under the domain of the law(s) given in the explanans).

Although the D-N Model and the I-S Model are similar in ways, they are importantly different. The general laws in a D-N explanation must be deterministic whereas those in the I-S Model are statistical (it is because some of our fundamental laws seem to be indeterministic that the I-S Model was proposed in the first place). Since the D-N Model involves general deterministic laws and conditions satisfying the conditions of those laws, the explanandum is shown to follow deductively from the explanans in this sort of explanation. In other words, a successful D-N Model explanation is a logically valid argument (an argument where the truth of

[10]Two notable exceptions to this majority opinion are Achinstein (1983) and van Fraassen (1980). Both of these philosophers hold that explanation cannot exist without acts of communication.

[11]Technically, Hempel (1965) discusses a third form of explanation, what he calls "Deductive-Statistical" explanations. These are explanations that derive a statistical uniformity from a more general statistical law. However, deductive-statistical explanations conform to the same pattern as the D-N Model so we will not treat them separately (Woodward 2014).

the premises guarantees the truth of the conclusion) where true explanans guarantee that the phenomenon described in the explanandum had to occur. For example, a D-N explanation of the position of Mars at some particular time, *t*, would include “Newton’s laws of motion, the Newtonian inverse square law governing gravity, and information about the mass of the sun, the mass of Mars and the present position and velocity of each” in the explanans (Woodward 2014). These explanans (assuming that they are all true) would deductively entail that Mars has the particular position it does at *t*.

Successful I-S Model explanations are different. Since these explanations involve statistical (indeterministic) laws, the explanandum is not shown to deductively follow from the explanans in this sort of explanation. Instead, a successful I-S explanation shows that the explanans confer a high probability on the explanandum. For instance, if it is a statistical law that it is highly probable someone with disease X will recover after taking a particular drug, and Blake has taken the drug, then this information can be used (along with the law) to provide an I-S explanation of Blake’s recovery from X (Woodward 2014).

There is much more that can be said about both of Hempel’s theories of explanation including specific details about the I-S Model, which is very complex.[12] Nevertheless, we have a sufficient grasp of these theories for our purposes. Let us turn our attention to some of the objections to Hempel’s theories.

9.1.2 Objections to Hempel’s Theories

In spite of the fact that many believe Hempel’s theories are largely correct about the nature of explanation and that these theories dominated the thinking on the topic of explanation for 20 years or more, numerous objections have been raised against Hempel’s theories. Here we will consider two of the most prominent, and most decisive, objections—one for each model.[13]

Let us begin with perhaps the most famous objection to the D-N Model—Bromberger’s flagpole counterexample.[14] Consider a situation where there is a flagpole of a particular height on a large patch of clear, level ground. The sun is shining brightly, and there is nothing to interfere with the sun’s rays reaching the

[12] In addition to Hempel’s own work see Salmon (1990) for an excellent discussion of Hempel’s theories as well as the ensuing philosophical debate for the 40 years following the publication of Hempel and Oppenheim (1948).

[13] For discussion of these and several other objections to Hempel’s theories see Salmon (1990), Strevens (2006), and Woodward (2014).

[14] Although this counterexample is attributed to Bromberger and often called “Bromberger’s flagpole” (Salmon 1990), Bromberger’s (1966) actual example involved a tower and its shadow. Nonetheless, the idea behind the counterexample is the same, and it was Bromberger who introduced this sort of counterexample. So, we will follow the literature and refer to this as “Bromberger’s flagpole”.

flagpole or with the flagpole's casting a shadow on the ground. Given the empirical facts about the flagpole's height, the position of the sun in the sky, and the law of rectilinear propagation of light, it is possible to deduce the length of the shadow that the flagpole casts on the ground. Hence, using the empirical information along with the law about the rectilinear propagation of light we can construct a successful D-N explanation of the length of the flagpole's shadow. This is all well and good for the D-N Model. There is a problem lurking here though. Using the empirical information about the sun's position in the sky and the length of the flagpole's shadow along with the law of rectilinear propagation of light we can construct an equally successful D-N explanation of the height of the flagpole. But, this cannot be right. While it seems we can explain the length of the shadow in this way it does not seem that the height of the flagpole can be explained like this. The shadow does not explain why the flagpole is the height that it is. Rather, a plausible explanation of the height of the flagpole will refer to the intentions of the person who erected the flagpole, the materials they had to use, and so on. The general problem is that for any explanandum for which we can give a successful D-N explanation we can turn around and use that explanandum to explain some of the explanans. The D-N Model does not properly capture explanatory asymmetries—the heights of flagpoles can explain the lengths of their shadows, but the length of flagpoles' shadows do not explain their heights.

Now, for a much discussed objection to the I-S Model.[15] This sort of objection raises a problem for the fact that in a successful I-S explanation the explanans *must* make the occurrence of the explanandum very probable. This is problematic in cases where the only cause of a particular kind of phenomena is *x*, but *x* does not *often* cause phenomena of that sort. The classic example of this sort of case involves syphilis and paresis (Scriven 1959). Only syphilis causes paresis. Yet, in most cases where someone has syphilis he does not contract paresis. So, if we were to find out that Joe has syphilis, we might think his odds of getting paresis are better than if he did not have syphilis, but his odds of contracting paresis are still very low. When we discover that Joe has paresis it seems like citing the fact that he has syphilis is an explanation of his having paresis though. After all, only people with syphilis contract paresis. Nevertheless, according to the I-S Model citing Joe's syphilis is not a good explanation of his having paresis because the fact that someone has syphilis does not make it very probable that he will contract paresis. It seems the restriction that the explanans must make the explanandum very probable is problematic.

Of course, there is much more that can be said about these objections and Hempel's theories in general. In fact, a lot has been said—the literature on explanation is vast and discussions of Hempel's theories dominated that literature for several years. However, for our purposes it is enough to have a handle on Hempel's theories and some of the objections which led the majority of theorists to look elsewhere for an account of explanation.

[15] See Hempel (1965), Lipton (2004), Scriven (1959), and van Fraassen (1980) for discussion of this objection.

9.1.3 Alternatives to Hempel's Theories and a Working Model of Explanation

In the mid-twentieth century there was largely a consensus about the nature of explanation—Hempel's theories, particular the D-N Model, were taken to be correct (Woodward 2003). In the wake of many purported counterexamples to Hempel's theories there is no longer a consensus concerning explanation. Rather, a number of competing accounts of explanation have been proposed. A complete, or even moderately complete, survey of these various alternatives to Hempel's theories is not possible here; nor is it necessary for our purposes. Consequently, we will simply very briefly consider some of the major alternatives that have been proposed and some of the challenges they face before settling on a working model of explanation. This working model of explanation will be fairly broad and ecumenical, which will work nicely for our purposes—we do not need to settle the intimate details of the nature of explanation in order to better understand NOS.

An early alternative to Hempel's theories is Wesley Salmon's (1971) Statistical Relevance (SR) account of explanation. According to the SR account, explanation requires describing the factors that are statistically relevant to the occurrence of the explanandum. A particular factor is statistically relevant to *E*'s occurring when the probability of *E*'s occurrence is greater when the factor is present than when it is not. In the ideal case an SR explanation would describe all of the factors that are statistically relevant to a particular explanandum's occurrence. For example, in the case of Joe's syphilis and paresis mentioned above Joe's syphilis is statistically relevant to his contracting paresis because the probability of contracting paresis when one has syphilis, while low, is higher than the probability of contracting paresis without having syphilis. In light of this, it seems that the SR account does not share the I-S Model's problem mentioned above. It also avoids the flagpole problem that plagues the D-N Model. While the height of the flagpole is statistically relevant to the length of the shadow, it does not seem that the length of the shadow is statistically relevant to the height of the flagpole. So far, so good.

The SR account is problematic, however. It has problems with cases where there is a common cause of two effects. Consider a case where a specific weather pattern typically leads to a storm, and the weather pattern also leads to a particular barometer reading. In such a case the barometer reading will be statistically relevant to the storm, when the barometer reading occurs it is more probable that the storm occurs than when the barometer reading does not occur. The storm will be similarly statistically relevant to the barometer reading though. The probability of the barometer reading is higher when the storm occurs than when it does not. But, intuitively the barometer reading does not explain the occurrence of the storm despite its statistical relevance to the storm's occurrence (Strevens 2006).[16]

[16]This sort of case is also a problem for the D-N Model.

Another alternative to Hempel's theories is the Unificationist account of explanation. The idea here is that explanation consists in subsuming the explanandum phenomenon under a system of laws which best unifies the relevant phenomena (Friedman 1974; Kitcher 1981, 1989). Very roughly, showing how phenomena or laws themselves can be taken to be instances of more fundamental laws creates a more unified picture of the world. The best explanations unify the most phenomena with the least number of fundamental principles. Unfortunately, the Unificationist account seems susceptible to the flagpole problem just like the D-N Model (Strevens 2006). After all, it seems that explaining the flagpole's height by appealing to the length of the shadow is just as unifying of a pattern as explaining the length of the shadow by appealing to the height of the flagpole. The reason is that, as we saw above, both of these "explanations" are apparently relying on the very same law and patterns to subsume the explanandum.[17]

The problems raised by the flagpole example and examples involving common causes, like the storm and the barometer, have led several theorists to think that the best way to approach explanation is by reference to causation. An early causal approach to explanation is Wesley Salmon's (1984) Causal Mechanical (CM) approach (adopted after his abandonment of the SR). On the CM account an explanation consists of a description of the causal influences of the explanandum. The sort of causal influences that are relevant to the CM approach are fundamental causes—close to how we would describe things in terms of the actions of fundamental particles in physics (Strevens 2006). There is a serious problem with this sort of account, however. As Strevens (2006) notes, in a normal case where a baseball shatters a window not only does the baseball count as part of the explanation of the window's shattering on the CM account, so do the shouts of the baseball players which cause very minute vibrations in the window. After all, these shouts do have a causal influence on the event of the window's shattering.[18] This seems implausible though because the fact that the baseball players are shouting seems irrelevant to the window's breaking.

The problem facing the CM account leads some causal theorists to focus on more nuanced accounts of causation. Some such accounts handle this sort of problem by allowing for multiple levels of causation. Often these accounts are cached out in terms of counterfactual views of causation (very roughly, accounts where *x* counts as a cause of *y* when it is the case that if *x* did not occur, *y* would not occur) (Lewis 1986; Woodward 2003). Other nuanced causal accounts agree with the CM account's focus on causal influence at a fundamental level while requiring the satisfaction of additional requirements in order for a particular fundamental causal influence to count as part of the explanation of the explanandum's occurrence (Salmon 1997; Strevens 2008).

[17] See Kitcher (1981) for attempts to address this issue.

[18] Strevens points out that Salmon appears to concede that this results from his theory.

Although such accounts of explanation hold promise, it is not clear that they can give us a completely general account of explanation.[19] Given their focus on causal relations it seems that some genuine explanations will not be properly accounted for under these theories. Assuming there can be genuine explanations in pure mathematics, it is very difficult to see how such explanations can be accounted for given any causal account of explanation.[20] Additionally, it seems like Philip can explain to his daughter why it is that Robert is her cousin by pointing out that Robert is the son of Philip's sister and that any children of one's parents' brothers or sisters are her cousins (Harman 1986). This seems like a perfectly good explanation, but it does not seem to be causal.

Obviously, our overview of alternatives to Hempel's theories is far from complete. There are many other accounts of the nature of explanation. There are also many other objections that have been raised for the theories we have discussed. Likewise, there are many responses on offer to the sorts of objections we have pointed out for these theories. There is no consensus on the exact nature of explanation because these objections and responses are still being raised and explored. Nevertheless, we now have a fairly good handle on the general issue of the nature of explanation and some of the surrounding debates. Before we turn our attention to the other major topic of this chapter, understanding, we will briefly discuss what we can take as a plausible working model of the nature of explanation for our purposes.

Our working model comes from Jaegwon Kim (1994). According to Kim, "*explanations track dependence relations*", that is to say, explanations are sets of propositions that provide information about dependence relations which hold between the explanans and the explanandum (1994, p. 68).[21] What makes Kim's proposal so useful as a working model of explanation is that there are a number of different kinds of dependence relations. Causal dependence is the sort of relation that holds when one thing causes another, so the sort of relations which causal accounts of explanation seek to capture fall within the purview of this model. The familial relations which constitute being a cousin are dependence relations too. The relation of constitution is itself a dependence relation. Mereological dependence, when "the properties of a whole, or the fact that a whole instantiates a certain property, may depend on the properties and relations had by its parts", is another sort of dependence relation (Kim 1994, p. 67). It is because of the fact that there are many kinds of dependence relations that our working model is ecumenical. It encompasses all of the sorts of relations that the prominent theories of explanation claim are of central importance. Of course, there are many details that would need to

[19]Of course, it might be a mistake to think that we can hope to find a theory of explanation that applies in all cases (Díez et al. 2013). Still, some argue that such a theory can be found (Nickel 2010). Until we have decisive reasons for thinking a fully general account of explanation is impossible, it seems to be a mark against a particular theory if it does not apply generally.

[20]For an overview of reasons for and against thinking there are genuine explanations in pure mathematics see Mancosu (2011).

[21]Strevens (2008) endorses this general conception of explanation as well.

be worked out in order for this dependence relation account to be a fully satisfactory final account of explanation. Fortunately, we do not need such an account for our purposes, and so, we do not need to trouble about these details. It is enough for us to have a working model of explanation, which we do.

Putting things together, we can understand an explanation to be a set of propositions where the explanans provides information about dependence relations which hold between phenomena described in the explanans and the phenomenon to be explained (the explanandum).[22] With this, undoubtedly abstract, working model in hand we are ready to turn our attention toward the aim of explanation: understanding.

9.2 Understanding

It is important to clarify the general sense of "understanding" that is our focus here before delving into more specific concerns. Since our primary focus throughout this book is scientific knowledge, we are especially interested in the sort of understanding which arises in science. For this reason, we will not concern ourselves with discussions of the myriad ways in which the word "understanding" might be used in English.[23] Also, the sense of "understanding" that concerns us is not merely the phenomenological sense in which an "explanation 'feels right'" or what is sometimes referred to as the "aha" experience of understanding (Trout 2002, p. 212). J.D. Trout (2002, 2005, 2007) argues, by appealing to psychological research concerning hindsight bias and overconfidence bias, that understanding construed as a phenomenological feeling of "rightness" is not a good indicator of the quality of an explanation. Although Trout tends to speak as if his arguments demonstrate that it is a mistake to think understanding is important for explanations or for science, they do not. Most philosophers who think providing understanding places constraints on what counts as an explanation are not thinking of understanding as a phenomenal feel, but as a kind of epistemic state. Although this epistemic state may be accompanied by the phenomenal feel that Trout targets, the two should not be conflated. It is the epistemic state we are concerned with, not merely the phenomenal "aha" feeling.[24]

[22]Admittedly, one might think that in order to have a "real" explanation one needs to provide information about causal dependence relations or information about natural laws governing the dependence relations in question. Two points about this concern are worth keeping in mind. First, insofar as it is plausible to think there are genuine explanations in pure mathematics it is unclear that such restrictions are necessary. Second, the addition of such restrictions would not affect the points made in this chapter. So, the reader is welcome to understand the account of explanation here as having such restrictions, if she believes they are necessary for genuine explanations.

[23]For a discussion of these issues see Franklin (1981).

[24]It is worth noting that most philosophers agree with Trout that this sort of "aha" experience is neither necessary nor sufficient for possessing genuine understanding. However, several

"Understanding" is "a cognitive success term", that is to say "understanding" denotes a mental state with positive epistemic status; it is a sort of cognitive achievement (Elgin 2007, p. 31).[25] There are various accounts of the nature of understanding construed as a positive epistemic state. For instance, Jonathan Kvanvig (2003, p. 192) says that "understanding requires the grasping of explanatory and other coherence-making relationships in a large and comprehensive body of information . . . understanding is achieved only when informational items are pieced together by the subject in question." Similarly, Linda Zagzebski (2001, p. 241) claims, "*understanding is the state of comprehension of nonpropositional structures of reality*." She goes on to identify three conditions of understanding:

1. It is acquired through mastering a *techne* [practical art/skill]
2. Its object is not a discrete proposition but involves seeing part/whole relations
3. It involves representing some portion of the world non-propositionally

Catherine Elgin (2007, p. 35) maintains, "understanding is primarily a cognitive relation to a fairly comprehensive, coherent body of information." The accounts of understanding that Kvanvig, Zagzebski, and Elgin offer, though different in important ways, seem to capture the general idea of what we have in mind when we claim that understanding is the goal of explanation. Since for our purposes a rough account of the nature of understanding will suffice, we will assume that understanding corresponds to something generally along the lines of what Kvanvig, Zagzebski, and Elgin propose.

Now that we have briefly outlined the relevant notion of "understanding" in broad strokes it is time to explore some of the details.

philosophers have argued Trout's arguments concerning this phenomenal sense of understanding do not undermine the importance of understanding as an aim of science or a goal of everyday explanations (de Regt 2004, 2009; Grimm 2009; Lipton 2009). Some (Lipton 2009; Grimm 2009) go so far as to argue that the "aha" experience may not be as misleading as Trout suggests, and in fact, it may be a reliable guide to the presence of genuine understanding in some cases.

[25] Although we tend to have an intuitive grasp of the notion, clearly defining "positive epistemic status" is difficult. Perhaps the best way to understand this notion is by noting other familiar states which have positive epistemic status such as justified/rational beliefs and knowledge on the one hand and states which lack positive epistemic status such as opinions, beliefs arising from wishful thinking, and guesses on the other hand. In this sense "understanding" is a term for a kind of mental state which belongs in the former group of mental states as opposed to the latter. The idea that understanding is a cognitive achievement (a mental state with positive epistemic status) is widespread. It is shared by those who hold very different views concerning other aspects of understanding, see, for example, Elgin (1996, 2007), Grimm (2001, 2006, 2014a, b), Khalifa (2011, 2012, 2013), Khalifa and Gadomski (2013), Kvanvig (2003, 2009), Pritchard (2009, 2014), Wilkenfeld (2013) and Zagzebski (2001).

9.2.1 Understanding Phenomena Versus Understanding Theories

Before exploring one of the central contemporary debates concerning understanding it will be helpful to first draw a further distinction between senses of understanding. We have already seen that for our discussion of understanding we are not concerned with the "aha" feeling of understanding, but rather with a cognitive achievement, a mental state with positive epistemic status. It is important to distinguish two sorts of cognitive achievements which are both kinds of understanding, and both relevant to our discussion: *understanding phenomena* (*UP*) and *understanding theories* (*UT*) (de Regt 2009).[26]

UP is the sort of understanding one has when she understands why or how a particular phenomenon occurs. This is the sort of understanding that is the aim of explanations, and so it is UP that is the primary aim of science. As we noted above, several philosophers claim that one gains UP by grasping a correct explanation of a particular phenomenon's occurrence. Thus, it is UP that has been our focus when discussing the relation between explanation and understanding.

At this point, one might wonder why we are bothering to distinguish UP and UT, or even discuss UT at all since UP is the sort of understanding that is the aim of science. The reason for this is simple. It is plausible that when it comes to scientific matters one can only truly possess UP, if she has UT (de Regt 2009; Strevens 2013). In other words, it is plausible that UT is necessary to have UP in scientific contexts.[27] So, what exactly is UT?

UT is the sort of understanding one has when she has the ability to use a theory.[28] The sense of "theory" that is relevant here is not the colloquial sense that occurs in ordinary English. Often in ordinary English the term "theory" is used to signify a claim or hypothesis that is still the subject of significant doubt. However, "among scientists, ... the term is often used to describe an established sub-discipline in which there are widely accepted laws, methods, applications, and foundations ... for

[26]Strevens (2013) draws a similar distinction between what he calls "understanding why" and "understanding with". The former corresponds to de Regt's UP while the latter corresponds to his UT.

[27]Plausibly, one way of understanding the distinction between UP and UT is in terms of their objects. The object of UP is natural phenomena; the object of UT is abstract content (theories).

[28]There is at least one other important sense of understanding, which is often characterized as "understanding-that". For example, Belle *understands that* the theory of relativity says X. While this sort of understanding is important, we will not focus on it here. The primary reason for this is that ascriptions of this sort of understanding are widely held to simply be knowledge ascriptions (see Kvanvig 2009; Pritchard 2009). For instance, when we say, "Belle understands that the theory of relativity says X" all we are saying is that "Belle knows that the theory of relativity says X". While this sort of understanding is important, it is fairly clear that it is simply a precondition of UT. One cannot have the ability to use a theory to construct explanations without knowing what the theory says. For present purposes we can assume individuals have this sort of understanding of the theories in question.

which there is strong empirical support" (Rosenberg 2012, p. 115).[29] It is the latter, scientific sense, of "theory" that is important for UT.[30] According to de Regt (2009, p. 37), UT "amounts to the ability of scientists to use relevant theories to construct explanations."[31] Whether or not you possess this ability with respect to a particular theory will depend on a variety of things. Of course, the virtues and nature of the theory itself will make a difference to whether you understand it in this sense. However, things like your "capacities, background knowledge, and background beliefs" (de Regt 2009, p. 33) will all make a difference to whether you have UT. This is extremely important because it is plausible that UT is necessary for UP. In light of this connection between the two kinds of understanding and the various conditions which affect one's having UT, it is possible that when "two people who possess exactly the same theories and background knowledge, one may achieve understanding of a phenomenon while the other does not" because the former might have the requisite capacities for the necessary UT while the latter does not (de Regt 2009, p. 37). [32]

It is worth emphasizing that the distinction between UP and UT along with their connection to one another helps make it clear how UP is a cognitive *achievement*. On this approach, gaining UP requires the exercising of various capacities, such as one's UT, in order to either come up with or properly appreciate an explanation which yields UP. It is not sufficient for understanding that you simply be informed

[29] These two very different understandings of the nature of theories, the ordinary and the scientific, may be partly to blame for some of the misguided objections to evolution such as the "it is just a theory" objection. For more on this and other misguided objections to evolution see McCain and Weslake (2013). See Kampourakis (2014) for an in-depth discussion of some of the common misunderstandings which lead to resistance against accepting evolutionary theory.

[30] Also, see National Research Council (2012) for a similar account of scientific theories. Of course, there are important issues concerning how we should understand the fundamental nature of scientific theories. It is not clear whether theories are best understood as ultimately collections of axiomatized sentences or nonlinguistic models, or both. Fortunately, for present purposes it is not necessary to settle the debate concerning the fundamental ontology of scientific theories. For more on this see Winther (2015).

[31] Plausibly, using a theory to make predictions about phenomena is simply an aspect of constructing explanations. After all, if explanations amount to information about dependence relations, then they allow one to make predictions about what will happen as well as to explain what has happened—explanation and prediction are simply backward looking (explanations) and forward looking (predictions) approaches to the same dependence relations. For this reason, we will focus primarily on explanation and the role that it plays in UT and UP. If one is convinced that explanation and prediction are not this closely connected, then she can construe the current discussion of UT in terms of the ability to construct explanations *and* the ability to make predictions.

[32] Both UT and UP come in degrees. One might have a greater or lesser degree of UT, and so understand a particular phenomenon that is explained by the relevant theory to a greater or lesser degree. In the text we are considering the high-end of the scale of UT—being able to construct explanations of phenomena requires a fairly significant degree of UT. This level of UT is something to strive for as an educational goal, but it is good to keep in mind that one can exhibit UT and UP without reaching this highest level. For example, one might have UT of a theory and UP of phenomena by being able to appreciate explanations of phenomena that are provided by a particular theory without being able to come up with such explanations on her own.

of an explanation of a particular phenomenon; you must be able to appreciate how the explanation provides an account of why or how the phenomenon occurs as it does.

9.2.2 Understanding and Knowledge

Now that we have gotten clearer on some of the important distinctions and basic points about the nature of understanding as well as its connection to explanation, it is time to briefly explore what is currently a major debate concerning explanation: the relationship between understanding and knowledge.

Interestingly, most philosophers of science hold that understanding is a kind of knowledge, but most epistemologists hold that understanding is not knowledge.[33] This debate is ongoing, and it does not need to be settled here. Nonetheless, we will briefly examine some of the considerations that have been adduced for thinking understanding (of either the UP or UT variety) is not a kind of knowledge. By doing so we will gain a sense of this debate.

To begin, one reason for thinking understanding is not a kind of knowledge is due to the purported transparency of understanding. Zagzebski (2001) argues that understanding is distinct from knowledge because while the former is transparent the latter is not. Zagzebski rightly notes that it is often the case that one knows that *p* without being aware of the factors which make it the case that she knows *p*. The reason for this is that a large number of the beliefs we have which are instances of knowledge depend upon other background evidence we possess for their justification. However, it is not the case that we are consciously aware of all of this background evidence. Consider your belief that you are reading this book. Setting aside skeptical worries, we can agree that your belief constitutes knowledge for you, i.e. you know that you are reading this book. The justification which you possess for this belief is dependent upon both your current perceptions and a large body of background evidence like the following: your knowledge of what a book is, your knowledge of what it is like to read a book, your knowledge of what experiences constitute your reading a book as opposed to your noting that someone else is reading a book, and so on. It is very probable that you are not consciously aware of all of the pieces of background evidence on which your knowledge that you are currently reading a book depends. So, your knowledge that you are reading a book is not transparent in the sense of your being aware of the factors which make it the case that you know it. Hence, Zagzebski is correct in claiming that knowledge is not always transparent because S can know that *p* without being aware of all of the factors which make it the case that she knows *p*.

---

[33]For a sampling of those who hold that understanding is a kind of knowledge see Achinstein (1983), Grimm (2006, 2009, 2014b), Khalifa (2011, 2013), Khalifa and Gadomski (2013), Kitcher (2002), Salmon (1990), Strevens (2008, 2013), and Woodward (2003). For a sampling of those who hold that understanding is not a kind of knowledge see Elgin (1996, 2007), Grimm (2001), Hills (2009, 2010), Kvanvig (2003, 2009), and Pritchard (2009, 2014).

Although knowledge is not transparent in this way, one might think that understanding is. According to Zagzebski, it is not possible for someone to have understanding of something and yet fail to be aware of the factors which make it the case that she understands it. This is an intuitive way of thinking of understanding, especially since understanding involves being aware of various relations between phenomena. In order to be said to understand phenomena one must be aware of the phenomena in some way and she has to be aware of the relations between the phenomena. Someone cannot lack awareness of this sort and still understand. Since it is the awareness of these relations which constitutes understanding, if someone understands something, she is aware of the factors which make it the case that she understands it. These sorts of considerations lead Zagzebski to plausibly claim that understanding and knowledge are two distinct kinds of things because understanding is transparent in this fashion, and knowledge is not.[34]

A further reason one might have for thinking understanding is not a kind of knowledge concerns factivity. As we have noted in earlier chapters, knowledge is factive—it is not possible to know that *p*, if *p* is false. Yet, some, such as Catherine Elgin (1996, 2007), argue that understanding is not factive.[35] Elgin (1996, p. 123) notes that the idea that "Objects in a vacuum fall toward the Earth at a rate of 32 ft/s" provides us with understanding of gravity and the rate of falling objects. Although this claim about the rate of falling objects provides understanding, it is not strictly speaking true. As Elgin observes, this claim ignores several factors which affect the rate at which an object will fall toward the earth, factors such as the force of the gravitational attractions of all other objects. Elgin even goes so far as to claim that the more complicated truth of this matter would be less likely to contribute to understanding of the phenomena than the simplified, but false, claim mentioned above.[36] Thus, Elgin claims that understanding is not a kind of knowledge because understanding is not factive.[37]

Perhaps the most persuasive reason to think that understanding is not a kind of knowledge comes from consideration of Gettier style cases. Jonathan Kvanvig (2003, pp. 197–98) makes a strong case for thinking understanding is not a kind of knowledge by way of the following example:

> Consider, say, someone's historical understanding of the Comanche dominance of the southern plains of North America from the late seventeenth until the late nineteenth centuries. Suppose that if you asked this person any question about this matter, she would

[34]See Grimm (2006) for criticisms of Zagzebski's claim that understanding is transparent in this way.

[35]It is worth noting that other philosophers who think understanding is not a kind of knowledge still maintain that understanding is factive, e.g. Kvanvig (2003, 2009) and Pritchard (2009, 2014).

[36]Elgin may be correct in this respect regarding the general person and her interests. However, it is highly dubious that the more complicated true statement would be less likely to contribute to the understanding of someone who comprehends this complicated statement as easily as the ordinary person does the simpler truth.

[37]The issue of how best to understand the role of idealizations in science is too complex to get into here. For now, it is simply worth mentioning that it is not clear that one must accept Elgin's view of how idealizations figure into our understanding of science.

> answer correctly. Assume further that the person is answering from stored information; she is not guessing or making up answers, but is honestly averring what she confidently believes the truth to be. Such an ability is surely constitutive of understanding, and the experience of query and answer, if sustained for a long enough period of time, would generate convincing evidence that the person in question understood the phenomenon of Comanche dominance of the southern plains.

Kvanvig adds to this example that the person gathered this information from an accurate history book, but that most all history books on the subject are inaccurate. Further, the person in this example had no reason to prefer the accurate history book that she did consult to any of the many inaccurate ones. So, it is a matter of luck that she came to possess accurate information, and her true beliefs about this subject are only accidentally connected to the truth. Thus, it seems the person in this example does not have knowledge of this subject because she is in a situation which matches the cases in the Gettier literature (notice, in particular, the similarity of this case to the fake barn case discussed in Chap. 8).

Although it seems the subject in this case does not have knowledge, Kvanvig argues that she does have understanding.[38] He points out that not only does she answer questions about this phenomenon correctly, but her "capacity for answering is counterfactual supporting . . . Ask anything about the phenomenon, and one would get a correct answer" (2003, p. 198). The person's awareness of the relations between the information she possesses concerning the Comanche dominance of the southern plains is not affected by how she came to be aware of those relations nor is her awareness of how the information she has about this phenomenon relates to her larger body of general historical information. Hence, it appears that the person in this example does not know because of the Gettier style situation she finds herself in, but she does understand. Therefore, it seems Kvanvig's example shows that understanding and knowledge are different in a significant way; understanding is not adversely affected by certain kinds of luck which keep one from having knowledge.[39]

Admittedly, the reasons we have examined in this section do not provide conclusive evidence that understanding is not a kind of knowledge, and the proposed

[38]See Hills (2009, 2010) and Pritchard (2009) for similar sorts of examples thought to show that understanding is distinct from knowledge. Grimm (2006, 2014b) and Khalifa (2013) contest whether such cases show that understanding is distinct from knowledge. It is worth noting, however, that Grimm does this by arguing that understanding is very similar to knowledge-how. As we noted in chapter two, knowledge-how seems to be importantly different than propositional knowledge (a difference which Grimm acknowledges). So, it may be that although Grimm argues that understanding is a kind of knowledge, his view is much closer to those who claim that understanding is not knowledge than it may seem at first glance.

[39]Pritchard (2009) helpfully distinguishes between two kinds of luck that arise in Gettier style cases. The first, which he terms "Gettier-style luck", is the sort of luck we find in Gettier's original cases and cases like the sheep in the field (all three of which were discussed in Chap. 8). The second, "environmental luck", is the sort of luck we find in fake barn cases. Pritchard argues that understanding is incompatible with Gettier-style luck, but it is compatible with environmental luck. Whereas, he claims that knowledge is incompatible with both kinds of luck.

distinction between understanding and knowledge is not without its detractors. Nonetheless, the fact that these issues are still up for debate should not trouble us. The purpose of this section was not to establish definitively that understanding is not a kind of knowledge, or that it is. Instead, the purpose of this section was merely to provide a sampling of some of the major points of contention in this debate. One final thought is worth mentioning. Even if it turns out that understanding is distinct from knowledge, this should not trouble us despite the fact that understanding is a primary aim of science. The reason for this is, as we have mentioned in Chap. 8 and earlier, the term "knowledge" with respect to scientific knowledge may perhaps be better understood as simply a general term for the sort of cognitive achievement we possess when we have understanding or simply a term for a strongly justified belief. At any rate, our conception of NOS may benefit from deeper appreciation of this debate, but it is not held hostage to its outcome.

9.3 Conclusion

In this chapter we have explored the nature of the fundamental aim of science, understanding. We have also examined the primary method by which we attain understanding in science, explanations. There are many ongoing debates about a number of issues concerning both understanding and explanation. However, we have enough here to help further our appreciation of the goals of science as well as how it is that we go about achieving those goals. One final point about understanding and explanation bears mentioning here. It seems plausible that since explanations aim at providing understanding, one very important criterion for determining whether one explanation is superior to another is the amount of understanding the explanation would provide if it were true (Lipton 2004). In the next chapter we will consider how we might compare explanations in order to gain the justification for accepting explanations and theories which is necessary for scientific knowledge. We will also see how the general explanatory method employed in science might be extended to provide an account of how our beliefs are justified more generally. Thus, we will explore the extent to which scientific knowledge is continuous with knowledge in general.

References

AAAS (American Association for the Advancement of Science). (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

Achinstein, P. (1983). *The nature of explanation*. New York: Oxford University Press.

Braaten, M., & Windschitl, M. (2011). Working toward a stronger conceptualization of scientific explanation for science education. *Science Education, 95*, 639–669.

Braithwaite, R. (1953). *Scientific explanation*. Cambridge, UK: Cambridge University Press.

Bromberger, S. (1966). Why-questions. In R. G. Colodny (Ed.), *Mind and cosmos* (pp. 86–111). Pittsburgh: University of Pittsburgh Press.

Dawid, R. (2013). *String theory and the scientific method*. Cambridge, UK: Cambridge University Press.
de Regt, H. W. (2004). Discussion note: Making sense of understanding. *Philosophy of Science, 71*, 98–109.
de Regt, H. W. (2009). Understanding and scientific explanation. In H. W. de Regt, S. Leonelli, & K. Eigner (Eds.), *Scientific understanding: Philosophical perspectives* (pp. 21–42). Pittsburgh: University of Pittsburgh Press.
de Regt, H. W. (2013). Understanding and explanation: Living apart together? *Studies in History and Philosophy of Science, 44*, 505–509.
de Regt, H. W., & Dieks, D. (2005). A contextual approach to scientific understanding. *Synthese, 144*, 137–170.
de Regt, H. W., Leonelli, S., & Eigner, K. (2009). Focusing on scientific understanding. In H. W. de Regt, S. Leonelli, & K. Eigner (Eds.), *Scientific understanding: Philosophical perspectives* (pp. 1–17). Pittsburgh: University of Pittsburgh Press.
Díez, J., Khalifa, K., & Leuridan, B. (2013). General theories of explanation: Buyer beware. *Synthese, 190*, 379–396.
Elgin, C. Z. (1996). *Considered judgment*. Princeton: Princeton University Press.
Elgin, C. Z. (2007). Understanding and the facts. *Philosophical Studies, 132*, 33–42.
Feldman, R. (1974). An alleged defect in Gettier counter-examples. *Australasian Journal of Philosophy, 52*, 68–69.
Franklin, R. L. (1981). Knowledge, belief and understanding. *Philosophical Quarterly, 31*, 193–208.
Friedman, M. (1974). Explanation and scientific understanding. *Journal of Philosophy, 71*, 5–19.
Gardiner, P. (1959). *The nature of historical explanation*. Oxford: Oxford University Press.
Gauch, H. G., Jr. (2012). *Scientific method in brief*. Cambridge, UK: Cambridge University Press.
Gijsbers, V. (2013). Understanding, explanation, and unification. *Studies in History and Philosophy of Science, 44*, 516–522.
Gopnik, A. (1998). Explanation as orgasm. *Minds and Machines, 8*, 101–118.
Gopnik, A., & Glymour, C. (2002). Causal maps and Bayes nets: A cognitive and computational account of theory formation. In P. Carruthers, S. Stich, & M. Siegal (Eds.), *The cognitive basis of science* (pp. 117–132). Cambridge, UK: Cambridge University Press.
Gopnik, A., & Sobel, D. (2000). Detecting blickets: How young children use information about novel causal powers in categorization and induction. *Child Development, 71*, 1205–1222.
Gopnik, A., Glymour, C., Sobel, D., Schulz, L., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychology Review, 111*, 1–31.
Grimm, S. (2001). Ernest Sosa, knowledge, and understanding. *Philosophical Studies, 106*, 171–191.
Grimm, S. (2006). Is understanding a species of knowledge? *British Journal for the Philosophy of Science, 57*, 515–535.
Grimm, S. (2009). Reliability and the sense of understanding. In H. W. de Regt, S. Leonelli, & K. Eigner (Eds.), *Scientific understanding: Philosophical perspectives* (pp. 83–99). Pittsburgh: University of Pittsburgh Press.
Grimm, S. (2014a). Understanding. In S. Bernecker & D. Pritchard (Eds.), *The Routledge companion to epistemology* (pp. 84–94). New York: Routledge.
Grimm, S. (2014b). Understanding as knowledge of causes. In A. Fairweather (Ed.), *Virtue epistemology naturalized: Bridges between virtue epistemology and philosophy of science* (pp. 329–346). Dordrecht: Springer.
Harman, G. (1986). *Change in view*. Cambridge, MA: MIT Press.
Hempel, C. G. (1942). The function of general laws in history. *Journal of Philosophy, 39*, 35–48.
Hempel, C. G. (1965). *Aspects of scientific explanation and other essays in the philosophy of science*. New York: The Free Press.
Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science, 15*, 135–175.
Hills, A. (2009). Moral testimony and moral epistemology. *Ethics, 120*, 94–127.

Hills, A. (2010). *The beloved self: Morality and the challenge from egoism*. New York: Oxford University Press.
Hindriks, F. (2013). Explanation, understanding, and unrealistic models. *Studies in History and Philosophy of Science, 44*, 523–531.
Kampourakis, K. (2014). *Understanding evolution*. Cambridge, UK: Cambridge University Press.
Khalifa, K. (2011). Understanding, knowledge, and scientific antirealism. *Grazer Philosophische Studien, 83*, 93–112.
Khalifa, K. (2012). Inaugurating understanding or repackaging explanation? *Philosophy of Science, 79*, 15–37.
Khalifa, K. (2013). Understanding, grasping, and luck. *Episteme, 10*, 1–17.
Khalifa, K., & Gadomski, M. (2013). Understanding as explanatory knowledge: The case of Bjorken scaling. *Studies in History and Philosophy of Science, 44*, 384–392.
Kim, J. (1994). Explanatory knowledge and metaphysical dependence. *Philosophical Issues, 5*, 51–69.
Kitcher, P. (1981). Explanatory unification. *Philosophy of Science, 48*, 507–531.
Kitcher, P. (1989). Explanatory unification and the causal structure of the world. *Minnesota Studies in the Philosophy of Science, 13*, 410–505.
Kitcher, P. (2002). Scientific knowledge. In P. Moser (Ed.), *The Oxford handbook of epistemology* (pp. 385–407). New York: Oxford University Press.
Kvanvig, J. (2003). *The value of knowledge and the pursuit of understanding*. Cambridge, UK: Cambridge University Press.
Kvanvig, J. (2009). The value of understanding. In A. Haddock, A. Millar, & D. Pritchard (Eds.), *Epistemic value* (pp. 95–111). Oxford: Oxford University Press.
Lehrer, K. (2000). *Theory of knowledge* (2nd ed.). Boulder: Westview Press.
Lewis, D. (1986). Causal explanation. In *Philosophical papers, volume II* (pp. 214–240). Oxford: Oxford University Press.
Lipton, P. (2004). *Inference to the best explanation* (2nd ed.). New York: Routledge.
Lipton, P. (2009). Understanding without explanation. In H. W. de Regt, S. Leonelli, & K. Eigner (Eds.), *Scientific understanding: Philosophical perspectives* (pp. 43–63). Pittsburgh: University of Pittsburgh Press.
Mancosu, P. (2011). Explanation in mathematics. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2011 Edition). http://plato.stanford.edu/archives/sum2011/entries/mathematics-explanation/
McCain, K., & Weslake, B. (2013). Evolutionary theory and the epistemology of science. In K. Kampourakis (Ed.), *The philosophy of biology: A companion for educators* (pp. 101–119). Dordrecht: Springer.
Moser, P. K. (1989). *Knowledge and evidence*. Cambridge, UK: Cambridge University Press.
Nagel, E. (1961). *The structure of science: Problems in the logic of scientific explanation*. New York: Harcourt, Brace and World.
National Research Council. (1996). *National science education standards*. Washington, DC: The National Academies Press.
National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Committee on Science Learning, Kindergarten through Eighth Grade. R. A. Duschl, H. A. Schweingruber, & A. W. Shouse (Eds.), Board on Science Education, Center for Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: The National Academies Press.
Nickel, B. (2010). How general do theories of explanation need to be? *Nous, 44*, 305–328.
Popper, K. R. (1959). *The logic of scientific discovery*. London: Hutchinson.

Pritchard, D. (2009). Knowledge, understanding and epistemic value. In A. O'Hear (Ed.), *Epistemology: Royal Institute of Philosophy Supplement 64* (pp. 19–44). Cambridge, UK: Cambridge University Press.

Pritchard, D. (2014). Knowledge and understanding. In A. Fairweather (Ed.), *Virtue epistemology naturalized: Bridges between virtue epistemology and philosophy of science* (pp. 315–328). Dordrecht: Springer.

Railton, P. (1993). Probability, explanation, and information. In D.-H. Ruben (Ed.), *Explanation* (pp. 160–181). New York: Oxford University Press.

Rosenberg, A. (2012). *Philosophy of science: A contemporary introduction* (3rd ed.). New York: Routledge.

Salmon, W. (1971). Statistical explanation. In W. Salmon (Ed.), *Statistical explanation and statistical relevance* (pp. 29–87). Pittsburgh: University of Pittsburgh Press.

Salmon, W. (1984). *Explanation and the causal structure of the world*. Princeton: Princeton University Press.

Salmon, W. (1990). *Four decades of scientific explanation*. Minneapolis: University of Minnesota Press.

Salmon, W. (1997). Causality and explanation: A reply to two critiques. *Philosophy of Science, 64*, 461–477.

Salmon, W. (1998). *Causality and explanation*. New York: Oxford University Press.

Schrödinger, E. (1954). *Nature and the Greeks*. Cambridge, UK: Cambridge University Press.

Scriven, M. (1959). Explanation and prediction in evolutionary theory. *Science, 130*, 477–482.

Sober, E. (1983). Equilibrium explanation. *Philosophical Studies, 43*, 201–210.

Steyvers, M., Tenenbaum, J., Wagenmakers, E., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science, 27*, 453–489.

Strevens, M. (2006). Scientific explanation. In D. M. Borchert (Ed.), *Encyclopedia of philosophy* (2nd ed.). Detroit: Macmillan Reference USA.

Strevens, M. (2008). *Depth: An account of scientific explanation*. Cambridge, MA: Harvard University Press.

Strevens, M. (2013). No understanding without explanation. *Studies in History and Philosophy of Science, 44*, 510–515.

Trout, J. D. (2002). Scientific explanation and the sense of understanding. *Philosophy of Science, 69*, 212–233.

Trout, J. D. (2005). Paying the price for a theory of explanation: De Regt's discussion of Trout. *Philosophy of Science, 72*, 198–208.

Trout, J. D. (2007). The psychology of scientific explanation. *Philosophy Compass, 2*, 564–591.

van Fraassen, B. C. (1980). *The scientific image*. Oxford: Oxford University Press.

van Fraassen, B. C. (1989). *Laws and symmetry*. Oxford: Oxford University Press.

von Wright, G. H. (1971). *Explanation and understanding*. Ithaca: Cornell University Press.

Wilkenfeld, D. (2013). Understanding as representation manipulability. *Synthese, 190*, 997–1016.

Wilkenfeld, D. (2014). Functional explaining: A new approach to the philosophy of explanation. *Synthese, 191*, 3367–3391.

Winther, R. G. (2015). The structure of scientific theories. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2015 Edition). http://plato.stanford.edu/archives/spr2015/entries/structure-scientific-theories/

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. New York: Oxford University Press.

Woodward, J. (2014). Scientific explanation. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2014 Edition). http://plato.stanford.edu/archives/win2014/entries/scientific-explanation/

Zagzebski, L. (2001). Recovering understanding. In M. Steup (Ed.), *Knowledge, truth, and duty: Essays on epistemic justification, responsibility, and virtue* (pp. 235–252). New York: Oxford University Press.