Chapter 3

Evidence and Induction

3.1 *The Mother of All Problems*

In this chapter we begin looking at a very important and difficult problem, the problem of understanding how observations can provide evidence for a scientific theory. In some ways, this has been *the* fundamental problem over the last hundred years of philosophy of science. This problem was central to the project of logical empiricism, and it was a source of constant frustration. Abandoning logical empiricism does not make the problem go away, though. In some form or other, it arises for nearly everyone.

The aim of the logical empiricists was to develop a logical theory of evidence in science. In this chapter I'll often use the terminology they favored, in which the *confirmation* of theories is what has to be understood. Confirmation is not the same as proof; a theory can be confirmed by evidence but turn out to be wrong. Confirmation is a kind of support that theories can get from evidence, where this support is usually partial rather than decisive. The logical empiricists wanted to treat confirmation as an abstract relation between sentences. It has become fairly clear that their approach to the problem is doomed. The way to analyze evidence in science is to develop a different kind of theory. But it will take a lot of discussion, in this and later chapters, before the differences between approaches that will and will not work can emerge. The present chapter will look mostly at how the problem of evidence was tackled in the middle of the twentieth century. And that is a tale of woe.

Before looking at twentieth-century work on these issues, we must again spend some time further back in the past. The confirmation of theories is closely connected to another classic topic in philosophy: the *problem of induction*. What reason do we have for expecting patterns observed in our past experience to hold in the future? What justification do we have for using past observations as a basis for generalization about things we have not yet observed? The most famous discussions of induction come from the eighteenth-century Scottish empiricist David Hume. Hume asked, what reason do we have for thinking that the future will resemble the past? There is no contradiction in supposing that the future could be totally unlike the past. It is possible that the world could change radically at any point, rendering previous experience useless. How do we know this will not happen? We might say to Hume that when we have relied on past experience before, it has turned out well for us. The policy has a good track record. But Hume will reply that this is begging the question—presupposing what has to be shown. Induction has worked in the past, yes, but that's the *past*. We have successfully used "past pasts" to tell us about "past futures." But our problem is whether anything about the past gives us good information about what will happen tomorrow.

Hume concluded that we have no reason to expect the past to resemble the future. Hume was an "inductive skeptic." He accepted that we all use induction to make our way around the world. And he was not suggesting that we stop doing so (even if we could). Induction is psychologically natural to us. Despite this, Hume thought it had no rational basis; it is just a habit we all have. Hume's inductive skepticism has haunted empiricism ever since. The problem of confirmation is not the same as the classical problem of induction, but the two are closely related.

3.2 *Induction, Deduction, Confirmation, and Explanatory Inference*

The logical empiricists tried to show how observational evidence could provide support for a scientific theory. Again, there was no attempt to show that scientific theories can be proved. Error is always possible, but evidence can support one theory over another. The cases that were to be covered by this analysis included the simplest and most traditional cases of induction: if we see a multitude of cases of white swans, and no other colors, why does that give us reason to believe that all swans are white? But obviously not all cases of evidence in science are like this.

The observational support for Copernicus's theory that the Earth goes around the sun, and for Darwin's theory of evolution, seem to work very differently from a traditional induction. Darwin did not observe a set of individual cases of evolution and then generalize.

The logical empiricists wanted a theory of evidence that would cover all these cases. They were not trying to develop a recipe for confirming theories. Rather, the aim was to give an account of the relationships between the statements that make up a scientific theory and statements describing observations, which make the observations support the theory.

Next, I should say more about the distinction between deductive and inductive logic (a distinction introduced in chapter 2). Deductive logic is the well-understood and less controversial kind of logic. It is a theory of patterns of argument that transmit truth with certainty. These arguments have the feature that if the premises of the argument are true, the conclusion is guaranteed to be true. An argument of this kind is *deductively valid*. The most famous example of a logical argument is a deductively valid argument:

premises	All men are mortal.
	Socrates is a man.
conclusion	Socrates is mortal.

A deductively valid argument might have false premises. In that case, the conclusion might be false as well (although it also might not be). What you get out of a deductive argument depends on what you put into it.

The logical empiricists thought that deductive logic could not serve as a complete analysis of evidence and argument in science. Scientific theories do have to be logically consistent, but this is not the whole story. Many inferences in science are not deductively valid, but they still can be good inferences; they can provide support for their conclusions.

For the logical empiricists, there is a reason why so much inference in science is not deductive. As empiricists, they believed that all our evidence derives from observation. Observations are always of *particular* objects and occurrences. But the logical empiricists thought that the great aim of science is to discover and establish *generalizations*. Sometimes the aim was seen as describing "laws of nature," though this concept was also regarded with some suspicion (see chapter 11). The central idea was that science aims at formulating and testing generalizations, and these generalizations were seen as having, at least potentially, an infinite range of application. No finite number of observations can conclusively establish a generalization of this kind, so inferences from observations in support of generalizations are always nondeductive. In contrast, all it takes is one case of the right kind to prove a generalization to be *false*; this fact will loom large in the next chapter.

In many discussions of these topics, the logical empiricists (and some later writers) used a simple terminology in which all arguments are either deductive or inductive. Inductive logic was thought of as a theory of all good arguments that are not deductive. But this terminology can be misleading, and I will set things up differently.

I will use the term "induction" only for inferences that go from particular observations to generalizations. To use the most traditional example, the observation of a large number of white swans (and no swans of any other color) might be used to support the hypothesis that all swans are white. We could express the premises with a list of particular cases: "Swan 1 observed at time t_1 was white; swan 2 observed at time t_2 was white. . . ." Or we might simply say: "All the many swans observed so far have been white." The conclusion will be the claim that all swans are white—a conclusion that could well be false but is supported, to some extent, by the evidence. Sometimes the terms "enumerative induction" or "simple induction" are used for inductive arguments of this most traditional kind. Not all inferences from observations to generalizations have this form, though. (And a note to mathematicians: mathematical induction is really a kind of deduction, even though it has the superficial form of induction.)

A form of inference closely related to induction is *projection*. In a projection, we infer from a number of observed cases to arrive at a prediction about the next case, not to a generalization about all cases. We see a number of white swans and infer that the next swan will be white. Obviously there is a close relationship between induction and projection, though there are a variety of ways of understanding this relationship.

Other kinds of nondeductive inference are seen often in science and everyday life. For example, during the 1980s Luis and Walter Alvarez began claiming that a huge meteor had hit the Earth about 65 million years ago, causing a massive explosion and dramatic weather changes that coincided with the extinction of the dinosaurs (Alvarez et al. 1980). The Alvarez team claimed that the meteor caused the extinctions, but let's leave that aside here. Consider just the hypothesis that a huge meteor hit the Earth 65 million years ago. A key piece of evidence for this hypothesis is the presence of unusually high levels of some rare chemical elements, such as iridium, in layers in the Earth's crust that are 65 million years old. These elements tend to be found in meteors in much higher concentrations than they are near the surface of the Earth. This observation was taken to be good evidence supporting the theory that a meteor hit the Earth around that time.

If we set this case up as an argument, with premises and a conclusion, it clearly is not an induction or a projection. We are not inferring to a generalization, but to a hypothesis about a structure or event that would explain our data. Several terms are used in philosophy for inferences of this kind. C. S. Peirce called these "abductive" inferences as opposed to inductive ones. Others have called them "explanatory inductions" or "theoretical inductions." The most common term for them is "inference to the best explanation," but I will use a slightly different one—"explanatory inference."

So I will recognize two main kinds of nondeductive inference, induction and explanatory inference (along with projection, which is closely linked to induction). The problem of analyzing evidence includes all of these.

How are these kinds of inference related to each other? For many logical empiricists and others, induction is the most fundamental kind of nondeductive inference. Reichenbach claimed that all nondeductive inference in science can be reconstructed in a way that depends only on a form of inference that is close to traditional induction. What looks like an explanatory inference can be somehow broken down and reconstructed as a complicated network of inductions and deductions. Carnap did not make this claim, but he did seem to view induction as a model for all other kinds of nondeductive inference. Understanding induction was in some sense the key to the whole problem.

So one way to view the situation is to see induction as fundamental.

But it is also possible to do the opposite, to claim that explanatory inference is fundamental. Gilbert Harman argued in 1965 that inductions are justified only when they are explanatory inferences in disguise, and others have followed up this idea in various ways.

Explanatory inference seems much more common in science than induction. In fact, you might be wondering whether science contains any inductions of the simple, traditional kind. That suspicion is reasonable, but it turns out that science does contain inferences, including important ones, that at least look like traditional inductions. Here is one example. During the work that led to the discovery of the structure of DNA by James Watson and Francis Crick, a key piece of evidence was provided by "Chargaff's rules." These rules, described by Erwin Chargaff in the late 1940s, have to do with the relation between the amounts of the four bases—C, A, T, and G—that help make up DNA. Chargaff found that in the DNA samples he analyzed, the amounts of C and G were always roughly the same, and the amounts of T and A were always roughly the same (Chargaff 1951). This fact about DNA became important in the discussions of how DNA molecules are put together. I called it a "fact" just above, but of course Chargaff had not observed *all* the molecules of DNA that exist, and neither have we. Back then, Chargaff's claim rested on an induction from a small number of cases (though he did cover a diverse range of organisms). Today we can give an argument for why Chargaff's rules hold that is not just a simple induction; the structure of DNA explains why Chargaff's rules must hold. But it might seem that, back when the rules were originally discovered, the only reason to take the rules to describe all DNA was inductive.

It probably seems a good idea, then, to refuse to treat one of these kinds of inference as more fundamental than the other. Perhaps there is more than one kind of good nondeductive inference (and maybe there are others besides the ones I have mentioned). Philosophers often find it attractive to think that there is ultimately just one kind of nondeductive inference, because that seems to be a simpler situation. But this argument from simplicity is not very convincing.

Let us return to our discussion of how the problem was handled by the logical empiricists. They used two main approaches. One was to formulate an inductive logic that looked as much as possible like deductive logic. That was Hempel's approach. The other approach, used by Carnap, was to apply the mathematical theory of probability. In the next two sections of this chapter, I will discuss some famous problems for logical empiricist theories of confirmation. The problems are especially easy to discuss in the context of Hempel's approach. An examination of Carnap's view is beyond the scope of this book. Over the course of his career, Carnap developed very sophisticated models of confirmation using probability theory applied to artificial languages. Problems kept arising. More and more special assumptions were needed to make the results come out in a way that looked reasonable. There was never a knockdown argument against him, but the project came to seem less and less relevant to science, and it eventually ran out of steam.

Although Carnap's approach to analyzing confirmation did not work out, the idea of using probability theory to understand confirmation remains popular and has been developed in new ways. Certainly this looks like a good approach; it does seem that observing the raised iridium level in the Earth's crust made the Alvarez meteor hypothesis *more probable* than before. In chapter 12 I will describe new ways of using probability theory to understand the confirmation of theories.

Before moving on to some famous puzzles, I will discuss a simple proposal that may have occurred to you.

The term "hypothetico-deductivism" is used in several ways by people writing about science. Sometimes it is used to describe a view about testing and confirmation. According to this view, hypotheses in science are confirmed when their logical consequences turn out to be true. This idea covers a variety of cases; the confirmation of a white-swan generalization by observing white swans is one case, and another is the confirmation of a hypothesis about an asteroid impact by observations of the true consequences of this hypothesis.

As Clark Glymour has emphasized (1980), an interesting thing about this idea is that it is hopeless when expressed in a simple way, but something like it seems to fit well with many episodes in the history of science. One problem is that a scientific hypothesis will only have consequences of a testable kind when it is combined with other assumptions, as we have seen. But set that problem aside for a moment. The suggestion above is that a theory is confirmed when a true statement about observables can be derived from it.

This claim is vulnerable to many objections. For example, any theory *T* deductively implies *T-or-S*, where *S* is any sentence at all. But *T-or-S* can be conclusively established by observing the truth of *S*. Suppose *S* is observational. Then we can establish the truth of *T-or-S* by observation, and that confirms *T*. This is obviously absurd. Similarly, if theory *T* implies observation *E*, then the theory *T&S* implies *E* as well. So *T&S* is confirmed by *E*, and *S* here could be anything at all. (Note the similarity here to a problem discussed at the beginning of section 2.4.) There are many more cases like this.

The situation is indeed strange. People often regard a scientific hypothesis as supported when its consequences turn out to be true; this is taken to be a routine and reasonable part of science. But when we try to summarize this idea using simple logic, it seems to fall apart. Does the fault lie with the original idea, with our summary of the idea using basic logic, or with basic logic itself?

The logical empiricists' response, here and elsewhere, was to hang steadfastly onto the logic, and often to hang onto their preferred ways of translating scientific theories into a logical framework as well. This led them to question some quite reasonable-looking ideas about evidence and testing. It also led to a situation in which philosophy of science seemed to turn into an exercise in "logic-chopping" for its own sake. And as we will see in a moment, even the logic-chopping did not go well. Despite all this, there is a lot to learn from the problems faced by logical empiricism. Confirmation really is a puzzling thing. Let us look at some famous puzzles.

3.3 *The Ravens Problem*

The logical empiricists put a great deal of work into analyzing the confirmation of generalizations by observations of their instances. At this point we will switch birds, in accordance with tradition. How is it that repeated observations of black ravens can confirm the generalization that all ravens are black?

First I will deal with a simple suggestion that will not work. Some readers might be thinking that if we observe a large number of black ravens and no nonblack ones, then at least we are cutting down the number of ways in which the hypothesis that all ravens are black might be wrong. With each raven we see, there is one fewer raven that might fail to fit the theory. In some sense, the chance that the hypothesis is true should be slowly increasing. But this does not help much. First, the logical empiricists were concerned to deal with the case where generalizations cover an infinite number of instances. In that case, as we see each raven we are not reducing the number of ways in which the hypothesis might fail. Also, note that even if we forget this problem and consider a generalization covering just a finite number of cases, the kind of support that is analyzed here is a weak one. That is clear from the fact that we get no help with the problem of projection. As we see each raven, we know there are fewer and fewer ways for the generalization to be false, but this does not tell us anything about what to expect with the next raven we see.

So let us look at the problem differently. Hempel suggested that, as a matter of logic, all observations of black ravens confirm the generalization that all ravens are black. More generally, any observation of an *F* that is also *G* supports the generalization "All *F*s are *G*." He saw this as a basic fact about the logic of support.

This looks like a reasonable place to start. And here is another obvious-looking point: any evidence that confirms a hypothesis *H* also confirms any hypothesis that is logically equivalent to *H*.

What is logical equivalence? Think of it as what we have when two sentences say the same thing in different terms. More precisely, if *H* is logically equivalent to *H**, then it is impossible for *H* to be true but *H** false, or vice versa.

But these two innocent-looking claims generate a problem. In basic logic, the hypothesis "All ravens are black" is logically equivalent to "All nonblack things are not ravens." Let us look at this new generalization. "All nonblack things are not ravens" seems to be confirmed by the observation of a white shoe. The shoe is not black, and it's not a raven, so it fits the hypothesis. But given the logical equivalence of the two hypotheses, anything that confirms one confirms the other. So the observation of a white shoe confirms the hypothesis that all ravens are black! That sounds ridiculous. As Nelson Goodman (1955) put it, we seem to have the chance to do a lot of "indoor ornithology"; we can investigate the color of ravens without ever going outside to look at one.

This simple-looking problem is hard to solve. Debate about it continues to this day. Hempel himself was well aware of this problem—he was the one who originally thought of it. But there has not been a solution proposed that everyone, or even a majority of people, has agreed upon.

One possible reaction is to accept the conclusion. This was Hempel's response. Observing a white shoe does confirm the hypothesis that all ravens are black, though presumably only by a tiny amount. Then we can keep our simple rule that whenever we have an "All *F*s are *G*" hypothesis, any observation of an *F* that is *G* confirms it and also confirms everything logically equivalent to "All *F*s are *G*." Hempel stressed that, logically speaking, an "All *F*s are *G*" statement is not a statement about *F*s but a statement about everything in the universe—the statement that if something is *F*, then it is *G*. We should note that according to this reply, the observation of the white shoe also confirms the hypothesis that all ravens are green, that all aardvarks are blue, and so on. Hempel was comfortable with this situation, but most others have not been.

A multitude of other solutions have been proposed. I will discuss just two ideas, which I regard as being on the right track.

Here is the first idea. Perhaps observing a white shoe or a black raven may or may not confirm "All ravens are black." It depends on other factors. Suppose we know, for some reason, that either (1) all ravens are black and ravens are extremely rare, or else (2) most ravens are black, a few are white, and ravens are common. Then a casual observation of a black raven will support (2), a hypothesis that says that not all ravens are black. If all ravens were black, we should not be seeing them at all. Similarly, observing a white shoe may or may not confirm a hypothesis, depending on what else we know. This reply was first suggested by I. J. Good (1967).

Good's move is very reasonable. We see here a connection to the issue of holism about testing, discussed in chapter 2. The relevance of an observation to a hypothesis is not a simple matter of the content of the two statements; it depends on other assumptions as well. This is so even in the simple case of a hypothesis like "All *F*s are *G*" and an observation like "Object *A* is both *F* and *G*." Good's point also reminds us how artificially simplified the standard logical empiricist examples are. No biologist would seriously wonder whether seeing thousands of black ravens makes it likely that all ravens are black. Our knowledge of genetics and bird coloration leads us to expect some variation, such as cases of albinism, even when we have seen thousands of black ravens and no other colors.

Here is a second suggestion about the ravens, which is consistent with Good's idea but goes further. Whether or not a black raven or a white shoe confirms "All ravens are black" might depend on the order in which you learn of the two properties of the object.

Suppose you hypothesize that all ravens are black, and someone comes up to you and says, "I have a raven behind my back; want to see what color it is?" You should say yes, because if the person pulls out a white raven, your theory is refuted. You need to find out what is behind his back. But suppose the person comes up and says, "I have a black object behind my back; want to see whether it's a raven?" Then it does not matter to you what is behind his back. You think that all ravens are black, but you don't have to think that all black things are ravens. In both cases, suppose the object behind his back is a black raven and he does show it to you. In the first situation, your observation of the raven seems relevant to your investigation of raven color, but in the other case it's irrelevant.

So perhaps the "All ravens are black" hypothesis is only confirmed by a black raven when this observation had the potential to refute the hypothesis, only when the observation was part of a genuine test.

Now we can see what to do with the white shoe. You believe that all ravens are black, and someone comes up and says, "I have a white object behind my back; want to see what it is?" You should say yes, because if he has a raven behind his back your hypothesis is refuted. He pulls out a shoe, however, so your hypothesis is OK. Then someone comes up and says, "I have a shoe behind my back; want to see what color it is?" In this case you need not care. It seems that in the first of these two cases, you have gained some support for the hypothesis that all ravens are black. In the second case, you have not.

So perhaps some white-shoe observations do confirm "All ravens are black," and some black-raven observations don't. Perhaps there is only confirmation when the observations arise during a genuine test, a test that has the potential to disconfirm as well as confirm.

Hempel saw the possibility of a view like this. His responses to Good's argument and to the order-of-observation point were similar, in fact. He said he wanted to analyze a relation of confirmation that exists just between a hypothesis and an observation itself, regardless of extra information we might have, and regardless of the order in which observations were made. But perhaps Hempel was wrong; perhaps there is no such relation. We cannot answer the question of whether an observation of a black raven confirms the generalization unless we know something about the way the observation was made and unless we make assumptions about other matters as well.

Hempel thought that some observations are just "automatically" relevant to hypotheses, regardless of what else is going on. That is true in the case of the deductive refutation of generalizations; no matter how we come to see a nonblack raven, that is bad news for the "All ravens are black" hypothesis. But what is true for deductive disconfirmation is not true for confirmation.

Clearly this discussion of order-of-observation does not entirely solve the ravens problem. Why does order matter? And what if both properties are observed at once? I will return to this issue in chapter 12, using a more complex framework. Putting it briefly, perhaps we can only understand confirmation and evidence by taking into account the procedures involved in generating data.

I will make one more comment on the ravens problem. This one is a digression, but it does help illustrate what is going on. In psychology there is a famous experiment called the "selection task" (Wason and Johnson-Laird 1972). The experiment has been used to show that many people (including highly educated people) make bad logical errors in certain circumstances. The experimental subject is shown four cards with half of each card masked. The subject is asked to answer this question: "Which masks do you have to remove to know whether it is true that if

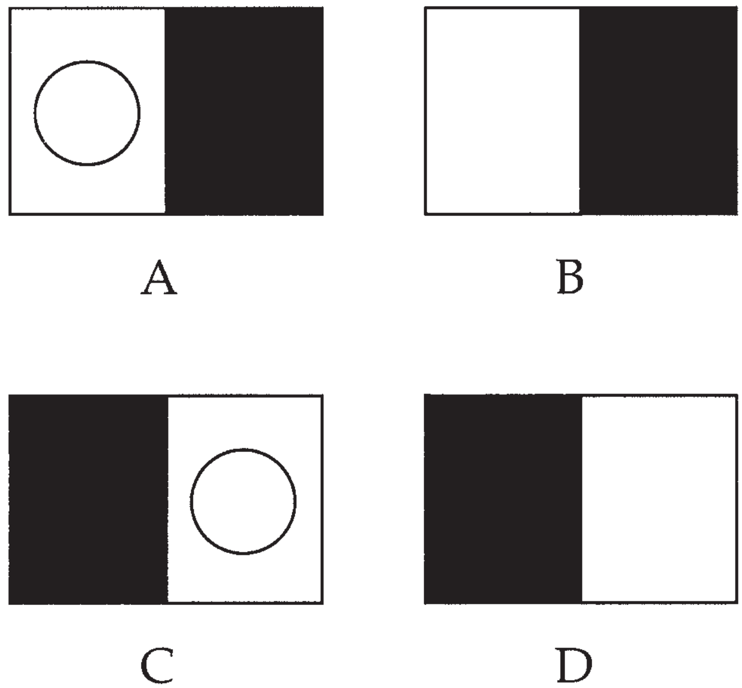

Figure 3.1. The Wason selection task

there is a circle on the left of a card, there is a circle on the right as well?" Have a look at figure 3.1 and try to answer the question yourself before reading the next paragraph.

Large majorities of people in many (though not all) versions of this experiment give the wrong answer. Many people tend to answer "only card *A*" or "card *A* and card *C*." The right answer is *A* and *D*. Compare this to the ravens problem; the problems have the same structure. I am sure Hempel would have given the right answer if he had been a subject in the four-card experiment, but the selection task might show something interesting about why confirmation has been hard to analyze. For some reason it is difficult for people to see the importance of "card *D*" tests in cases like this, and it is easy for people to think, wrongly, that "card *C*" tests are important. If you are investigating the hypothesis that all ravens are black, card *D* is analogous to the situation when someone says he has a white object behind his back. Card *C* is analogous to the situation where he says he has a black object behind his back. Card *D* is a real test of the hypothesis, but card *C* is not. Unmasking card *C* is evidentially useless, even though it may fit with what the hypothesis says. Not all observations of cases that fit a hypothesis are relevant as tests.

3.4 *Goodman's "New Riddle of Induction"*

In this section I will describe an even more famous problem, revealed by Nelson Goodman (1955). This argument looks strange, and it is easy to misinterpret. But the issues it raises are very deep.

Goodman's immediate goal was to show that there cannot be a purely "formal" theory of confirmation. He does not think that confirmation is impossible, or that induction is a myth. He just thinks they work differently from the way many philosophers have thought.

What would a formal theory of confirmation look like? The easiest way to explain this is to look at deductive arguments. Recall the most famous deductively valid argument:

Argument 1

premises	All men are mortal.
	Socrates is a man.
	———————————
conclusion	Socrates is mortal.

The premises, if they are true, guarantee the truth of the conclusion. But the fact that the argument is a good one does not have anything in particular to do with Socrates or manhood. Any argument that has the same form is just as good. That form is as follows:

All *F*s are *G*.
Object *a* is *F*.
———————————
Object *a* is *G*.

Any argument with this form is deductively valid, no matter what we substitute for "*F*," "*G*," and "*a*." As long as the terms we put in do pick out definite properties or classes of objects, and as long as the terms retain the same meaning all the way through the argument, the argument will be valid.

The deductive validity of arguments depends only on the form or pattern of the argument, not the content. This is something the logical empiricists wanted to build into their theory of induction and confirmation. Goodman aimed to show that this is impossible; there can never be a formal theory of induction and confirmation.

How did Goodman do it? Consider argument 2.

Argument 2

All the many emeralds observed, in diverse circumstances,
prior to the year 2050 have been green.

═══════════════════════════

All emeralds are green.

This looks like a good inductive argument. (Like some of the logical empiricists, I use a double line between premises and conclusion to indicate that the argument is not supposed to be deductively valid.) The argument does not give us a guarantee; inductions never do. And if you would prefer to express the conclusion as "Probably, all emeralds are green," or something similar, that will not make a difference to the rest of the discussion. (If you know something about minerals, you might object that emeralds are often regarded as green by definition: emeralds are beryl crystals made green by trace amounts of chromium. Please just regard this as another unfortunate choice of example by the literature.)

Now consider argument 3:

Argument 3

All the many emeralds observed, in diverse circumstances, prior to the year 2050 have been grue.

═══════════════════════════

All emeralds are grue.

Argument 3 uses a new word, "grue." We define "grue" as follows:

> **GRUE**: An object is *grue* if and only if it was first observed before the year 2050 and is green, *or* if it was not first observed before 2050 and is blue.

The world contains lots of grue things; there is nothing strange about grue objects, even though there is something strange about the word. The grass outside my door as I write this is grue. The sky outside on July 1, 2055, will be grue, if it is a clear day. An individual object does not have to change color in order to be grue—this is a common misinterpretation. Anything green that is at some point observed before 2050 passes the test for being grue. So, all the emeralds we have seen so far have been grue.[1]

Argument 3 does not look like a good inductive argument. Argument 3 leads us to believe that emeralds observed in the future will be blue, on the basis of previously observed emeralds being green. The argument also conflicts with argument 2, which looks like a good argument. But arguments 2 and 3 have exactly the same form. That form is as follows:

> All the many *E*s observed, in diverse circumstances, prior to 2050 have been *G*.
>
> ═══════════════════════
>
> All *E*s are *G*.

We could represent the form even more schematically than this, but that does not matter to the point. Goodman's point is that two inductive arguments can have the exact same form, but one argument can be good

---

1. You are probably reading this before 2050, and might then wonder how we know the premise to be true. Why not say an object is grue if and only if it has been observed before now and is green, or has never been observed and is blue? It is fine to present things that way, as long as we treat "now" in the argument as a fixed time, not something that keeps moving. I think it's easier to get a grip on the argument by choosing a particular date in the future and supposing that all the emeralds seen before that date are green.

while the other is bad. So what makes an inductive argument a good or bad one cannot be just its form. Consequently, there can be no purely formal theory of induction and confirmation. Note that the word "grue" works perfectly well in *de*ductive arguments. You can use it in the form of argument 1, and it will cause no problems. But induction is different.

Suppose Goodman is right so far. We then need to work out what is wrong with argument 3. This is the new riddle of induction. The obvious thing to say is that there is something about the word "grue" that makes it inappropriate for use in inductions. A good theory of induction should include a restriction of some kind on the terms that occur in inductive arguments. "Green" is OK and "grue" is not.

This has been the most common response to the problem. But as Goodman says, it is hard to spell out the details of such a restriction. Suppose we say that the problem with "grue" is that its definition includes a reference to a specific time. Goodman's reply is that whether or not a term is defined in this way depends on which language we take as our starting point. To see this, let us define a new term, "bleen."

> BLEEN: An object is *bleen* if and only if it was first observed before the year 2050 and is blue, *or* if it was not first observed before 2050 and is green.

We can use the English words "green" and "blue" to define "grue" and "bleen," and if we do so we must build a reference to time into the definitions. But suppose we spoke a language that was like English except that "grue" and "bleen" were basic, familiar terms and "green" and "blue" were not. Then if we wanted to define "green" and "blue," we would need a reference to time.

> GREEN: An object is *green* if and only if it was first observed before the year 2050 and is grue, *or* if it was not first observed before 2050 and is bleen.

(You can see how it will work for "blue.") Goodman claimed that whether or not a term "contains a reference to time" or "is defined in terms of time" is a *language-relative* matter. Terms that look OK from the standpoint of one language will look odd from another. So if we want to rule out "grue" from inductions because of its reference to time, then whether an induction is good or bad will depend on which language we treat as our starting point. Goodman thought this conclusion was fine. A good induction, for Goodman, must use terms that have a history of normal use in our community. That was his own solution to his problem. Most other philosophers did not like this at all. It seemed to say that the value of inductive arguments depended on irrelevant facts about which language we happen to use.

Consequently, many philosophers have tried to focus not on the words "green" and "grue" but on the *properties* that these words pick out, or the *classes* or *kinds* of objects that are grouped by these words. We might argue that greenness is a natural and objective feature of the world, and grueness is not. Putting it another way, the green objects make up a "natural kind," a kind unified by real similarity, while the grue objects are an artificial or arbitrary collection. Then we might say: a good induction has to use terms that we have reason to believe pick out natural kinds. Taking this approach plunges us into hard problems in other parts of philosophy. What is a property? What is a "natural kind"? These problems have been controversial since the time of Plato. (I'll have a look at them in chapter 10.)

Although Goodman's "riddle" is abstract and uses invented terms, it has interesting links to real problems in science. I think Goodman's problem encapsulates within it several distinct methodological issues. First, there is a connection between Goodman's problem and the "curve-fitting problem" in data analysis. Suppose you have a set of data points in the form of x and y values, and you want to discern a general relationship expressed by the points by fitting a function to them. The points in figure 3.2 fall almost exactly on a straight line, and that seems to give us a natural prediction for the y value we expect for $x = 4$. However, there is an infinite number of different mathematical functions that fit our three data points (as well or better) but which make different predictions for the case of $x = 4$. How do we know which function to use? Fitting a strange function to the points seems to be like preferring a grue induc-

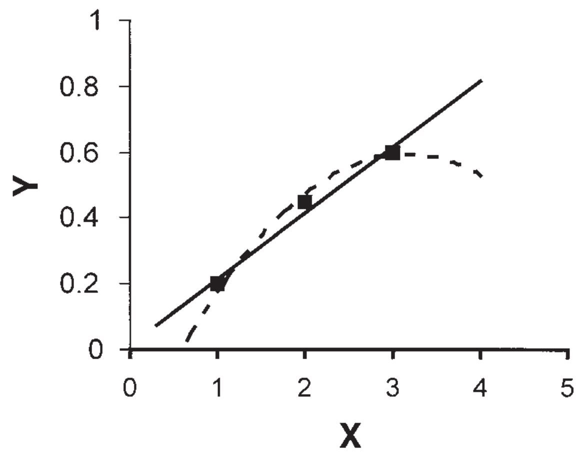

Figure 3.2. The curve-fitting problem

tion over a green induction when inferring from the emeralds we have seen.

Scientists dealing with a curve-fitting problem like this may have extra information telling them what sort of function is likely in a particular case, or they may prefer a straight line on the basis of *simplicity*. That suggests a way in which we might deal with Goodman's original problem. Perhaps the green induction is to be preferred on the basis of its simplicity?

That might work, but is it really so clear that the green-emerald induction is simpler? Goodman will argue that the simplicity of an inductive argument depends on which language we assume as our starting point, for the kinds of reasons given earlier in this section. For Goodman, what counts as a simple pattern depends on which language you speak or which categorization you assume. Hume, in the original problem of induction, wondered whether we had reason to believe that the future would resemble the past. Goodman, using his "new riddle," argued that whether or not some imagined future "resembles the past" depends on which language we habitually use.

A separate question is whether simpler theories really deserve some sort of preference in science. It is common to think so, but the issue is not so clear—this I will tackle in chapter 13.

That (almost) concludes our first foray into the problems of induction and confirmation. These problems are simple but very resistant to solution. For a good part of the twentieth century, it seemed that even the most innocent-looking principles about induction and confirmation led straight into trouble.

Some of these problems with evidence might seem to be consequences of a particular way of setting things up—an approach to evidence that puts induction in too central a place. That is roughly what I think. But here, near the end of the chapter, I want to emphasize another side to the role of induction in science, a side that involves its historical role, going back even before Hume's skeptical arguments.

In the history of science and philosophy, the idea of induction is associated with Francis Bacon, who wrote in the early seventeenth century and strongly advocated the use of induction in science. Bacon's sense of induction was not quite the same as the sense seen in logical empiricism, or Hume, but it is related. For Bacon, what was important was a contrast between different methods—different ways of approaching investigation. When Bacon wrote, he was trying first to urge people to move away from trusting ancient authorities. But once we are doing our own investigations, he wanted people to also move away from an approach in which the researcher looks for very general and basic principles first, and then seeks to understand particular things in terms of those principles. Instead, Bacon advocated a method in which you build up slowly, from lots of particular bits and pieces of data, toward bigger and more general ideas.

What Bacon had in mind in this building-up process was not just the tallying of cases (one black raven, another black raven . . .). His approach included looking for negative cases, and he also emphasized intervention in nature—experiment, not just observation. He said a scientist should "twist the lion's tail," not merely take in what's going on. All this was quite early in the Scientific Revolution, and what the idea of induction carried with it at that time was partly an ideal of *flexibility* in intellectual life. This embrace of flexibility is a very important idea, and a valuable part of the empiricist tradition. Philosophers, though, writing afterward and looking for a more systematic and exact treatment of induction, had a tendency to bring everything down to the relationship between particular instances and generalizations that cover them. They have thought that the relationship between individual cases and generalizations must be the heart of the matter. Or perhaps some of them thought that this relationship was at least the right place to start—the best route in to an understanding of evidence. Neither view turned out to be right.

3.5 *Optional Section: A Little More about Hypothetico-Deductivism*

Having explored induction, it would be good to have another look at the approach to evidence I moved over quickly in an earlier part of this chapter, the *hypothetico-deductive* (HD) approach. This short section goes down some twisting paths and can be skipped.

According to a simple HD view, hypotheses in science are confirmed when their observational consequences turn out to be true. As I said, this approach initially seems closer to most science than the approach which tries to reduce everything to inductive arguments from instances to generalizations. But it also seems to run into problems as soon as you fill the idea out.

I said earlier: "One problem is that a scientific hypothesis will only have consequences of a testable kind when it is combined with other assumptions, as we have seen. But set that problem aside for a moment." Let's look at it now. To move as quickly as possible, I'll use a very simple example again. It is not true that "All ravens are black" implies that you will see any black ravens. All it does on its own is forbid observations of other kinds of ravens. If you believe that some particular bird is a raven with its natural coloration intact, and also that all ravens are black, then you can infer a prediction about how things should look. But even in this case, the prediction comes from more than the generalization itself.

Suppose a prediction made by a theory together with background assumptions does work out. Perhaps Newton's laws of motion and gravity, together with other assumptions, are used to successfully predict an observation about what we see in the night sky—the return of a comet, for example. Then it seems that the whole network of claims operating here, the laws and the other assumptions, all get some credit from the prediction turning out well. However, we can add some apparently irrelevant claims to Newton's laws and still get the same prediction. Newton's laws plus Freud's psychology, along with the background assumptions we needed before, will give us the same successful prediction about the sky. Do we give some credit to Freud as well? One might reply that Freud's psychology played no role here. It should get no credit. We do just as well, in this case, with Freud in the picture and without him. However, quite a lot of the content of Newton's physics also played no role in the prediction. Newton's laws cover all physical bodies, not just comets and stars and planets. Should we give credit only to some parts of Newton's theory, when it does well with the comet? A reply: "But Newton's laws don't distinguish comets from other physical bodies, and that is part of what made his theory so good." Right, but you don't need this aspect of his view to derive predictions about the comet.

Perhaps we might just accept that only the part of Newton's theory that we need to make the prediction should get any credit. But the bit of the theory we need might turn out to be rather small. We have ended up in a surprising place yet again. Despite all this, it is hard to believe that there is nothing right in the HD approach.

Further Reading and Notes

Hume's classic discussions of induction are in his *Treatise of Human Nature* ([1739] 1978) and his shorter and more readable *Enquiry Concerning Human Understanding* (1748). Once again, Hempel's *Aspects of Scientific Explanation* (1965) is a central source for logical empiricist ideas on this topic. Skyrms's *Choice and Chance* (2000) is a good introductory book on these issues, and it introduces probability theory as well. For Bacon, see Klein and Giglioni (2016).

For a discussion of explanatory inference, see the Harman paper I cited (1965), Peter Lipton's *Inference to the Best Explanation* (1991), and Achinstein's *The Book of Evidence* (2001). For "abduction," see Peirce ([1898] 1993). There are several different ways of using order-of-observation to address the ravens problem. The version that influenced me is in Horwich's *Probability and Evidence* (1982). The Wason/Hempel connection has been noted independently by a number of people. The first may have been Yael Cohen (1987).

Goodman's classic presentation of his "new riddle of induction" is in *Fact, Fiction and Forecast* (1955). The problem is in chapter 3 (along with other interesting ideas), and his solution is in chapter 4. Douglas Stalker has edited a collection on Goodman's riddle, called *Grue!* (1994). The Quine and Jackson chapters are particularly good.