# Chapter 8: Scientific Reasoning[[1]](#footnote-1)

Because the contributions of modern science are our culture's best examples of advances in knowledge, it is important for everyone to have some appreciation of how scientists reason. This chapter deeply examines the nature of scientific reasoning, showing how to assess the scientific claims we encounter in our daily lives, how to do good scientific reasoning, and how to distinguish science from mere pseudoscience. We begin with a description of science and a review of some of the methods of doing science.

## I. What is Science?

Science creates machinery of destruction. It spawns mutants. It spews radiation. It murders in order to dissect. Its apparently objective pose is a cover for callous indifference. Its consequence will be the annihilation of us all. Edward Rothstein said this….

Oops. Wait a minute, a message just arrived, and I need to make an announcement: “Dr. Frankenstein, please call your office.”

OK, where was I? Oh, yes, well, enough of this glowing praise of science. "Science" is the Latin term for knowledge. Latin philosophers used the term to refer to an organized body of knowledge, tracing this understanding to the fourth century BCE philosopher Aristotle, who sought to understand the nature of scientific inquiry.[[2]](#footnote-2) By "science" we will mean pure empirical science, the kind of science that makes observations and runs experiments trying to make predictions, create explanations and produce theoretical understanding of the physical world. This contemporary understanding of science does not match perfectly with Aristotle’s, for instance, Aristotle does not incorporate experimentation into his work. Nevertheless, the practice of formulating theories based upon careful observation and rigorous thinking has a long history predating even Aristotle and we see scientific developments across ancient cultures, e.g. Babylon, Africa and China. Our use of the term “science” in this text will rule out mathematics and formal logic, and it rules in physics, chemistry, and biology. At any particular time in history, science has what it claims is a body of knowledge, but actually science as we use the term in this text is more a way of getting knowledge than it is a body of knowledge.

Creating science is not what doctors, engineers and inventors do. These people apply science, but usually they do not do science in the sense of create science.

Consider engineering. Unlike scientists, the engineers primarily want to improve existing things that have been made by humans, such as tractors and X-ray machines, or they want to improve human beings’ abilities to move faster and to communicate more easily with people who are far away.

Scientists often make use of advances in engineering, but they have different primary concerns. Pure science is concerned primarily with understanding, explaining, and predicting. Engineering (or applied science) focuses rather on creating technology and controlling it, on getting machines to function as we want them to in a particular situation. That is how pure (or theoretical) scientists are different from engineers. Inventors and doctors are more like the engineers than like the scientists.

Proposing precise questions and seeking precise answers is one of the keys to successful science. With precision comes sophistication.

Although the scientist's vocabulary is often so technical that the rest of us cannot read a scientific research paper, science is not as distant from common sense as many people imagine. Science isn't the only way to know the world around us. They don't have a "lock" on knowledge. But scientists, like the rest of us, do look around at the world, try to explain what they observe, and are careful to back up what they say. Science is a slowed-down and more open and accountable image of what we normally do in coming to know about the world around us. Nevertheless, science isn't just common sense. Science is more cautious about what it claims to know, and it often overthrows traditional common sense in favor of new beliefs that can better stand up to testing.

Everybody agrees that science is important, even Edward Rothstein whose sarcastic remarks inspired the paragraph above about science spawning mutants and spewing radiation. But some people think science is much more important and valuable than others do. According to the distinguished historian of science Herbert Butterfield, the rise of European science in the seventeenth and eighteenth centuries is a turning point in our history; however, this view downplays the significance of developments across the centuries in a variety of civilizations.

The European scientific revolution was noteworthy for promoting the notion that scientific knowledge should be produced by the process that we now call the scientific method, which emerges out of the work of medieval thinkers, whose attention to observation, theory and experimentation led to developments in fields such as medicine, agriculture, optics, harmonics, metallurgy, and physics.[[3]](#footnote-3) At its heart, the contemporary scientific method is the method of testing hypotheses.[[4]](#footnote-4) The idea is that the true hypotheses will stand up to repeated testing while the false hypotheses eventually will get refuted.

In addition to biology, chemistry, and physics, which are the more commonly known sciences, another lesser-known science is stamp collecting. Here is why. Stamp collectors are careful; they use tools; they explain; they predict; and they make generalizations. These are marks of good science.

Stamp collectors are careful, like scientists. They measure and use tools such as rulers. They can explain why stamps have perforations and why they aren’t cubical. They can predict that most collections will have more three-cent stamps from 1944 than seventy-four cent stamps from 1944. They make generalizations, such as “There are more European stamps than Egyptian stamps.” So that's why stamp collecting is a science.



No, think again. Don’t believe everything you read. Stamp collecting is definitely not a science. It’s a hobby. All that reasoning I just performed was making the same kind of error as if I’d argued like this:

A woman has two legs, one nose, and breathes air. Mr. Dowden has two legs, one nose, and breathes air. Therefore, Mr. Dowden is a woman.

More is involved in being a woman, right? Similarly, more is involved in being a science. The difficulty is in being more specific about just what else is involved. Here is an attempt to specify what else.

Many philosophers of science would say that in addition to being precise, careful, using tools, explaining phenomena, predicting observations, and making generalizations, science also: (1) requires using the scientific method to justify its claims. More on this later. (2) Science assumes a background of no miracles and no supernatural causes. It is unscientific to say there was a hurricane in the Philippine Islands because God was angry with the people there. (3) Science has theories that are held tentatively and are falsifiable. That means science is opposed to dogma, and it requires science’s claims to be true or false depending on what the evidence is. If you have a theory that couldn’t be shown to be incorrect no matter what happens, then you aren’t doing science. Freud's theory of psychoanalysis has that defect.

## II. Reviewing the Principles of Scientific Reasoning

One fairly significant aspect of scientific reasoning distinguishes it from other reasoning: Its justification process can be more intricate. For example, you and I might look back over our experience of gorillas, seeing them in zoos and seeing pictures of them in books, and draw the conclusion that all gorillas are black. A biological scientist interested in making a statement about gorilla color would not be so quick to draw this conclusion; he or she would contact gorilla experts and would systematically search through information from all the scientific reports about gorillas to check whether the general claim about gorilla color has even one counterexample. Only if none were found would the scientist then say, "Given all the evidence so far, all gorillas are black." The scientific community as a whole is even more cautious. It would wait to see whether any other biologists disputed the first biologist's claim. If not, only then would the community agree that all gorillas are black. This difference between scientific reasoning and ordinary reasoning can be summed up by saying that scientific reasoning has higher standards of proof.

Scientists don't rummage around the world for facts just so they can accumulate more facts. They gather specific facts to reach general conclusions, the "laws of science." Why? Because a general conclusion encompasses a great variety of specific facts, and because a general claim is more useful for prediction, understanding and explanation, which are the three primary goals of science. Scientists aren't uninterested in specifics, but they usually view specific data as a steppingstone to a broader or more general overview of how the world works. This point can be expressed by saying that scientists prefer laws to facts. Although there is no sharp line between laws and facts, facts tend to be more specific; laws, more general.

The power that generality provides is often underestimated. At the zoo, suppose you spot a cage marked "Margay" although the margay is out of sight at the moment. You have never heard of a margay, yet you can effortlessly acquire a considerable amount of knowledge about the margay, just by noticing that the cage is part of your zoo's new rare-feline center. It’s cat-like. If so, then you know it cannot survive in an atmosphere of pure nitrogen, that it doesn't have gills, and that it was not hatched from an egg. You know this about the unseen margay because you know on scientific authority that no cat-like beings can survive in nitrogen, that no cats have gills, and that no cats are hatched from eggs. You don’t know all this first-hand, but you’ve heard it indirectly from scientists, and you’ve never heard of any serious disagreement. Of course, scientific generalizations can be wrong. And maybe no experiment has ever been performed to test whether margays can live on pure nitrogen. But you are confident that if there were serious suspicions, the scientists would act quickly to run the tests. Knowing this about how scientists act, you rest comfortably with the generalizations and with your newly acquired knowledge about margays.

Definitions:

A **test** is an observation or an experiment intended to provide evidence about a claim.

A **law** of science is a sufficiently well-tested general claim.

A **theory** is a proposed explanation or a comprehensive, integrated system of laws that can be used in explaining a wide variety of phenomena.

### Testability, Accuracy, and Precision

If a proposed hypothesis (a claim) cannot be tested even indirectly, it is not scientific. This point is expressed by saying that scientists highly value testability. For example, suppose someone suggests, “The current laws of chemistry will hold true only as long as the Devil continues to support them. After all, the Devil made the laws, and he can change them on a whim. Luckily he doesn't change his mind too often.” Now, what is a chemist to make of this extraordinary suggestion? Even if a chemist were interested in pursuing the suggestion further, there would be nothing to do. There is no way to test whether the Devil is or isn't the author of the laws of chemistry. Does the Devil show up on any scientific instrument, even indirectly? Therefore, the Devil theory is unscientific.

Testability is a key ingredient of any truly scientific claim.

Scientists value accuracy and precision. An accurate measurement is one that agrees with the true state of things. A precise measurement is one of a group of measurements that agree with each other and cluster tightly together near their average. However, precision is valuable to science more than in the area of measurement. Precise terminology has helped propel science forward. Words can give a helpful push. How? There are two main ways. A bird may go by one name in the Southeastern United States but by a different name in Central America and by still a different name in Africa. Yet scientists the world over have a common Latin name for it. Thus, the use of precise terminology reduces miscommunication among scientists. Second, a precise claim is easier to test than an imprecise one. How do you test the imprecise claim that "Vitamin C is good for you"? It would be easier to run an experiment to check the more precise claim "Taking 400 milligrams of vitamin C per day will reduce the probability of getting a respiratory infection by fifty percent." If you can test a claim, you can do more with it scientifically. Testability is a scientific virtue, and precision is one path to testability.

Because the claims of social science are generally vaguer than the claims of physical science, social scientists have a tougher time establishing results. When a newspaper reports on biology by saying, "Vitamin C was shown not to help prevent respiratory infections," and when the paper reports on social science by saying, "Central America is more politically unstable than South America," we have a better understanding of the former, as "help prevent" can readily be given an unproblematic operational definition, whereas "politically unstable" is more difficult to define operationally. That is, the operation the biologist performs to decide whether something helps prevent respiratory infections can be defined more precisely, and easily, and accurately than the operation to be performed to decide whether one country is more politically stable than another.

### Reliability of Scientific Reporting

Almost every piece of scientific knowledge we have, we justify on the authority of what some scientist has said or is reported to have said. Because scientists are authorities on science, we usually take their word for things scientific. But chemists are not authorities on geology, and chemists who are experts in inorganic chemistry usually are not authorities on organic chemistry. Thus, when we are told that something is so because scientists believe it to be so, we should try to determine whether the proper authorities are being appealed to. Also, we know that scientists disagree on some issues but not on others, and we know that sometimes only the experts know which issues the experts disagree about. Is the reporter reporting the view of just one scientist, unaware that other scientists disagree? Scientists have the same moral failings as the rest of us, so we should also worry about whether a scientist might be biased on some issue or other. If a newspaper reporter tells us that the scientist's research on cloth diapers versus disposable diapers was not financed by the manufacturer of either diaper, we can place more confidence in the report.

Scientific journals are under greater pressure than daily newspapers to report the truth. A scientific journal will lose its reputation and its readers faster when there is a slipup than will the daily newspaper. So the stakes in reporting the truth are higher for journals. That is one reason the editors of scientific journals demand that authors provide such good evidence in their articles. If we read a report of a scientific result in a mainstream scientific journal, we can assume that the journal editor and the reviewers demanded good evidence. But if we read the report in a less reputable source, we have to worry that sloppy operational definitions, careless data collection, inaccurate instruments, or misunderstandings by the reporter may have colored the result.

When the stakes are high and we are asked to take an authority's word for something, we want independent verification. That means doing something more than merely buying a second copy of the newspaper to check whether what our first copy says is true. In medicine, it means asking for a second opinion from a different doctor. When the doctor says he wants to cut off your leg, you want some other doctor who is independent of the first doctor to verify that your leg really needs to be amputated. The term independent rules out your going to a partner in the first doctor's practice.

Ordinarily, though, we can't be bothered to take such pains to find good evidence. When we nonscientists read in the newspaper that some scientist has discovered something or other, we don't have enough time to check out the details for ourselves; we barely have enough time to read the reporter's account, let alone read his or her sources. So, we have to absorb what we can. In doing so, though, we who are critical thinkers are not blank slates willing to accept anything told to us. We are sensitive enough to ask ourselves: Does the report sound silly? Are any scientists protesting the result? What is the source of the report? We know that a reputable scientific journal article about some topic is more reliable than a reporter's firsthand interview with the author; we trust the science reporters for the national news magazines over those for a small, daily newspaper; and we know that daily newspapers are more reliable than independent bloggers and grocery store tabloids. But except for this, we nonscientists have severe difficulties in discriminating among the sources of information.

Suppose you were to read the following passage in a magazine: "To ensure the safety of raw fish, it should be frozen for at least five days at minus 4 degrees Fahrenheit (-20°C). That temperature kills all relevant parasitic worms so far tested." Should you believe what you read? It depends. First, ask yourself, "Where was it published and who said it?" In fact, the passage appeared in Science News, a well-respected, popular scientific publication. The magazine in turn was reporting on an article in an authoritative scientific publication, the New England Journal of Medicine. The journal in turn attributed the comment to Peter M. Schantz of the Centers for Disease Control in Atlanta, Georgia, a well-respected U.S. federal research laboratory. The magazine merely reported that Schantz said this. If you learned all this about the source of the passage in Science News, then you should probably accept what is said and add it to your knowledge.

You should accept it, but to what degree? You should still have some doubts based on the following concerns. The magazine did not say whether any other scientists disagreed with what Schantz said or even whether Schantz made this comment speculatively rather than as the result of a systematic study of the question. The occurrence of the word tested in the quote would suggest the latter, but you can't be sure. Nevertheless, you can reasonably suppose that the comment by Schantz was backed up by good science or the magazine wouldn't have published it the way it did— that is, with no warning that the claims by Schantz were not well supported. So, you can give Schantz's claims a high degree of belief, but you could be surer of what Schantz said if you had gotten direct answers to your concerns. Hearing from another scientific expert that Schantz's claims about fish are correct should considerably increase your degree of belief in his claims.

### Causal Explanations vs. Causal Arguments

Scientists and reporters of science present us with descriptions, explanations, and arguments. Scientists describe, for example, how ballistic missiles fall through the sky. In addition to description, scientists might also explain the phenomenon, saying why it occurs the way it does. The explanation will give the causes, and in doing so it will satisfy the following principle: **Explanations should be consistent with well-established results** (except in extraordinary cases when the well-established results are being overthrown with extraordinarily good evidence).

Scientists who publicly claim to have the correct explanation for some phenomenon have accepted a certain burden of proof. It is their obligation to back up their explanation with an argument that shows why their explanation is correct. We readers of scientific news usually are more interested in the description and the explanation than in the argument behind it, and we often assume that other scientists have adequately investigated the first scientist's claim. This is usually a good assumption. Thus, reporters rarely include the scientific proof in their report, instead sticking to describing the phenomenon, explaining it, and saying that a certain scientist has proved that the phenomenon should be explained that way.

Scientific proofs normally do not establish their conclusions as firmly as mathematical proofs do. Scientific proofs are inductive; mathematical proofs are deductive. So, one scientific proof can be stronger than another scientific proof even though both are proofs. In any inductive scientific proof, there is never a point at which the conclusion has been proved beyond a shadow of all possible doubt. Nevertheless, things do get settled in science. Scientists proved that the Earth is round, not flat; and even though this result is not established beyond all possible doubt, it is established well enough that the scientific community can move on to examine other issues confident that new data will not require any future revision. In fact, you haven't a prayer of getting a research grant to double-check whether the Earth is flat.

One scientific proof can be stronger than another scientific proof even though both are proofs.

### Good Evidence

Many persons view science as some vast storehouse of knowledge. That is an accurate view, but we also should view science as a way of getting to that knowledge. This latter way of looking at science is our primary concern in this chapter. In acquiring knowledge, a good scientist adopts a skeptical attitude that says, "I won't believe you unless you show me some good evidence." Why do scientists have this attitude? Because it is so successful. Scientists who are so trusting that they adopt beliefs without demanding good evidence quickly get led astray; they soon find themselves believing what is false, which is exactly what science is trying to avoid.

What constitutes good evidence? How do you distinguish good from bad evidence? It’s not like the evidence appears with little attached labels of “good” and “bad.” Well, if a scientist reports that tigers won't eat vegetables, the report is about a phenomenon that is repeatable—namely, tiger meals. If the evidence is any good, and the phenomenon is repeatable, the evidence should be, too. That is, if other scientists rerun the first scientist's tests, they should obtain the same results. If not, the evidence was not any good. The moral here is that reproducible evidence is better than evidence that can't be reproduced. The truth is able to stand up to repeated tests, but falsehood can eventually be exposed. That is one of the major metaphysical assumptions of contemporary science.

A scientist who appreciates good evidence knows that having anecdotal evidence isn't as good as having a wide variety of evidence. For example, suppose a scientist reads an article in an engineering journal saying that tests of 300 randomly selected plastic ball bearings showed the bearings to be capable of doing the job of steel ball bearings in the electric windows of Honda cars.[[5]](#footnote-5)

A close-up of a ball bearing


The journal article reports on a wide variety of evidence, 300 different ball bearings. If a scientist were to hear from one auto mechanic that plastic bearings didn't hold up on the car he repaired last week, the scientist won't be a good logical reasoner if he (or she) immediately discounted the wide variety of evidence and adopted the belief of the one auto mechanic. We logical reasoners should trust the journal article over the single anecdote from the mechanic, although the mechanic's report might alert us to be on the lookout for more evidence that would undermine the findings of the journal article. One lemon does not mean that Honda’s electric windows need redesigning. If you discount evidence arrived at by systematic search, or by testing, in favor of a few firsthand stories, you’ve committed the fallacy of overemphasizing anecdotal evidence.

### A Cautious Approach with an Open Mind

The scientific attitude is also a cautious one. If you are a good scientist, you will worry initially that perhaps your surprising new evidence shows only that something is wrong somewhere. You won't claim to have revolutionized science until you’ve made sure that the error isn't in the faulty operation of your own measuring apparatus. If a change of beliefs is needed, you will try to find a change with minimal repercussions; you won't recommend throwing out a cherished fundamental law when you can just as easily revise it by changing that constant from 23 to 24 so that it is consistent with all data, given the margin of error in the experiments that produced the data. The cautious scientific attitude recognizes these principles: **Don't make a broader claim than the evidence warrants**, and don't reject strongly held beliefs unless the evidence is very strong. In short, don't be wildly speculative.

Scientists are supposed to think up reasonable explanations, but what counts as a reasonable explanation? An explanation that conflicts with other fundamental beliefs that science has established is initially presumed to be unreasonable, and any scientist who proposes such an explanation accepts a heavier than usual burden of proof. A related principle of good explanation is to not offer supernatural explanations until it is clear that more ordinary, natural explanations won't work.

In assessing potential new beliefs—candidates for new knowledge—scientists actively use what they already believe. They don't come into a new situation with a mental blank. When scientists hear a report of a ghost sighting in Amityville, they will say that the report is unlikely to be true. The basis for this probability assessment is that everything else in the scientists' experience points to there being no ghosts anywhere, and so not in Amityville, either. Because of this background of prior beliefs, a scientist will say it is more probable that the reporter of the Amityville ghost story is confused or lying than that the report is correct. Better evidence, such as multiple reports or a photograph, may prompt a scientist to actually check out the report, if Amityville isn't too far away or if someone provides travel expenses.

Good scientists don't approach new data with the self-assurance that nothing will upset their current beliefs. Scientists are cautious, but they are also open to new information, and they don't suppress counterevidence, relevant evidence that weighs against their accepted beliefs. They do search for what is new; finding it is how they get to be famous. So the scientific attitude requires a delicate balance.

Keep an open mind, but don't be so open that you spend most of your valuable time on wild goose chases.

### Discovering Causes, Creating Explanations, and Solving Problems

Contrary to what Francis Bacon recommended in 1600, clearing your head of the B.S. and viewing nature with an open mind is not a reliable way to discover the causes behind what you see. Unfortunately, there is no error-free way. Nevertheless, the discovery process is not completely chaotic. There are rules of thumb. For example, to discover a solution to a problem, scientists can often use a simple principle: Divide the problem into manageable components. This principle was used by the space program in solving the problem of how to travel to the moon. The manager of the moon program parceled out the work. Some scientists and engineers concentrated on creating a more powerful rocket engine; others worked on how to jettison the heavy, empty lower stages of the rocket; others designed the communication link between the Earth and the spaceship's computer; and still others created the robot mechanisms that could carry out the computer's commands during flight and after landing on the moon. In short: Divide and conquer.

Another principle of scientific discovery says to assume that similar effects are likely to have similar causes. The history of medicine contains many examples of using this principle effectively. Several times before 1847, Doctor Ignaz Semmelweis of the General Hospital in Vienna, Austria had tried but failed to explain the alarming death rate of so many women who gave birth in his maternity ward. They were dying of puerperal fever, a disease with gruesome symptoms: pus discharges, inflammation throughout the body, chills, fever, delirious ravings. One day, a Dr. Kolletschka, who worked with Semmelweis, was performing an autopsy on a puerperal fever victim when a clumsy medical student nicked Kolletschka's arm with a scalpel. A few days later Kolletschka died with the same symptoms as the women who died of puerperal fever. Semmelweis suspected a connection. Perhaps these were similar effects due to a similar cause. And perhaps whatever entered Kolletschka via the student's scalpel was also being accidentally introduced into the women during delivery. Then Semmelweis suddenly remembered that the doctors who delivered the babies often came straight from autopsies of women who had died of puerperal fever. Maybe they were bringing infectious material with them and it somehow entered the bodies of women during delivery of their babies. Semmelweis's suggestion of blaming the doctors was politically radical for his day, but he was in fact correct that this disease, which we now call blood poisoning, was caused by doctors transferring infectious matter from the dead mothers on the dissecting tables to the living mothers in the delivery rooms. Semmelweis's solution was straightforward. Doctors must be required to wash their hands in disinfectant before delivering babies. That is one reason that today doctors wash their hands between visits to patients.

A good method to use when trying to find an explanation of some phenomenon is to look for the key, relevant difference between situations in which the phenomenon occurs and situations in which it doesn't. Semmelweis used this method of discovery. You can use the same method to make discoveries about yourself. Suppose you were nauseous, then you vomited. You want to know why. The first thing to do is to check whether prior to your symptoms you ate something you'd never eaten before. If you discover there was something, it is likely to be the cause. Did you get those symptoms after eating raw tuna, but not after eating other foods? If so, you have a potentially correct cause of your problem.

To find a cause, look for a relevant difference between situations where the effect occurs and situations where it does not.

The rules of thumb we have just discussed can help guide scientific guessing about what causes what. There are a few other rules, some of which are specific to the kind of problem being worked on. Guessing is only the first stage of the discovery process. Before the guess can properly be called a discovery, it needs to be confirmed. This is the second stage, and one that is more systematic than the first, as we shall see.

### Confirming by Testing

To prove your hypothesis about tuna scientifically, you would need to run some tests. One test would be to eat the tuna again and see whether it causes the symptoms again. That sort of test might be dangerous to your health. Here is a better test: acquire a sample of the tuna and examine it under a microscope for bacteria known to cause the symptoms you had.

Suppose you do not have access to the tuna. What can you do? You might ask other people who ate the tuna: "Did you get sick, too?" Yes answers would make the correlation more significant. Suppose, however, you do not know anybody to ask. Then what? The difficulty now is that even if you did eat tuna before you got your symptoms, was that the only relevant difference? You probably also ate something else, such as french fries with catsup. Could this have been the problem instead? You would be jumping to conclusions to blame the tuna merely on the basis of the tuna eating being followed by the symptoms; that sort of jump commits the post hoc (or “after this”) fallacy. At this point you simply do not have enough evidence to determine the cause of your illness.

Let's reexamine this search for the cause, but at a more general level, one that will provide an overview of how science works in general. When scientists think about the world in order to understand some phenomenon, they try to discover some pattern or some causal mechanism that might be behind it. They try out ideas the way the rest of us try on clothes in a department store. They don't adopt the first idea they have, but instead are willing to try a variety of ideas and to compare them.

Suppose you, a scientist, have uncovered what appears to be a suspicious, unexplained correlation between two familiar phenomena, such as vomiting and tuna eating. Given this observed correlation, how do you go about explaining it? You have to think of all the reasonable explanations consistent with the evidence and then rule out as many as you can until the truth remains. One way an explanation is ruled out is when you collect reliable data inconsistent with it. Another way is if you notice that the explanation is inconsistent with accepted scientific laws. If you are unable to refute the serious alternative explanations, you will be unable to find the truth; knowledge of the true cause will elude you. This entire cumbersome process of searching out explanations and trying to refute them is called the scientific method of justifying a claim. There is no easier way to get to the truth. People have tried to take shortcuts by gazing into crystal balls, taking drugs, or contemplating how the world ought to be, but those methods have turned out to be unreliable.

Observation is passive; experimentation is active. Experimentation is a poke at nature. It is an active attempt to create the data needed to rule out a hypothesis. Unfortunately, scientists often cannot test the objects they are most interested in. For example, experimenters interested in whether some potential drug might be harmful to humans would like to test humans but must settle for other species of animal. Scientists get into serious disputes with each other about whether the results of testing on rats, rabbits, and dogs carry over to humans. This dispute is really a dispute about analogy; is the animal's reaction analogous to the human's reaction?

Scientists often collect data from a population in order to produce a general claim about that population. The goal is to get a representative sample, and this goal is more likely to be achieved if the sample size is large, random, diverse, and stratified. Nevertheless, nothing you do with your sampling procedure will guarantee that your sample will be representative. If you are interested in making some claim about the nature of polar bears, even capturing every living polar bear and sampling it will not guarantee that you know the characteristics of polar bears that roamed the Earth 2,000 years ago.

Relying on background knowledge about the population's lack of diversity can reduce the sample size needed for the generalization, and it can reduce the need for a random sampling procedure. If you have well-established background knowledge that electrons are all alike, you can run your experiment with any old electron; don't bother getting Egyptian electrons as well as Japanese electrons.

### Aiming to Disconfirm

In the initial stages of a scientific investigation, when a scientist has an idea or two to try out, it is more important to find evidence in favor of the idea than to spend time looking for disconfirming evidence. However, in the later stages, when a scientist is ready to seriously test the idea, the focus will turn to ways to shoot it down. Confirming evidence—that is, positive evidence or supporting evidence—is simply too easy to find. That is why the scientist designs an experiment to find evidence that would refute the idea if it were false. Scientists want to find the truth, but the good scientist knows that the proper way to determine the truth of some idea is to try to find negative, not positive, evidence. A scientific generalization, at least a universal one of the form "All X are Y," will have all sorts of confirming instances (things that are both X and Y), but it takes just one X that is not Y to refute the whole generalization. So disconfirming evidence is more valuable than confirming evidence at this later stage of scientific investigation. Failure to find the disconfirming evidence is ultimately the confirming evidence.

When a hypothesis can stand up to many and varied attempts to rule it out, the hypothesis is tentatively accepted as true or proved.

Although scientific reasoning is not so different from other kinds of logical reasoning, it is special in that its claims tend to be more precise, and the evidence backing up the claims is gathered more systematically. This completes our review of what earlier chapters have said about scientific reasoning. Let's now probe deeper into the mysteries of science.

### Looking for Alternative Explanations

Suppose you receive a letter asking you to invest your money with Grover Hallford and Associates (GHA), a new stock brokerage firm. You do have a little extra cash,[[6]](#footnote-6) so you don't immediately shut the idea out of your mind. The new stockbrokers charge the same rates as other major brokers who offer investment advice. GHA is unusual, though, in that it promises to dramatically increase your investment because, according to the letter, it has discovered a special analytic technique for predicting the behavior of the stock market. Normally you would have to pay for any stock advice from a broker, but to show good faith, the GHA letter offers a free prediction for you. It predicts that the price of IBM stock will close lower next Tuesday from where it closed at the end of trading on the previous day, Monday. You place the letter in file 13, the circular file (the garbage can). However, the following week you happen to notice that IBM stock did perform as predicted. Hmmm. What is going on?

A few days later you receive a second letter from GHA. It says that GHA is sorry you have not yet become a client, but, to once again show its good faith, the company asks you to consider its prediction that Standard Oil of New Jersey stock will close up next Tuesday from where it was at the end of Monday. Again you decline to let GHA invest your savings, but you do keep an eye on the stock price of Standard Oil of New Jersey during the next week. Surprisingly, the prediction turns out to be correct. A few days later you receive a third letter suggesting that you invest with GHA, containing yet another free stock tip, but warning that there is a limit to how much free advice you will receive. Are you now ready to invest with GHA? If not, how many more letters would you have to receive before you became convinced that the brokers truly do understand the logic of the stock market? If you demand thirty letters, aren't you being foolish and passing up the chance of a lifetime? Surely GHA is on to something, isn't it? Other brokers cannot perform this well for you. How often do you get a chance to make money so easily? Isn't GHA's unknown technique causing them to be able to make correct predictions? And even if GHA is cheating and somehow manipulating the market, you can still take advantage of this and make money, too. Think about what you would do if you were faced with this decision about investing.

You may not have been able to find a reasonable alternative explanation to GHA's claim that it understands the causal forces shaping the stock market. Many people cannot.

That's why the swindle works so well. However, it is a swindle, and it is illegal. What GHA did is to get a long mailing list and divide it in half. For their first letter, half of the people get a letter with the prediction that IBM stock will close higher next Tuesday; the other half get a letter making the opposite prediction—that IBM will not close higher.

Having no ability to predict the stock market, GHA merely waits until next Tuesday to find out who received a letter with the correct prediction. Only that half then gets a second letter. Half of the second letters say Standard Oil of New Jersey stock will go up; the other half say it won't. After two mailings, GHA will have been right two times in a row with one-fourth of the people it started with. The list of names in the lucky fourth is divided in half and GHA generates a new letter. Each new mailing cuts down by 50 percent the number of people GHA has given good advice to, but if the company starts with a long enough list, a few people will get many letters with correct predictions. You are among those few. This explains why you have received the letters. Along the way, many people will have sent their hard-earned money to GHA, money that will never be returned. This swindle is quite effective. Watch out for it. And don't use it yourself on anybody else.

Once again we draw a familiar moral. The degree of belief you should give to a claim that A causes B (that GHA's insight into the stock market causes its correct predictions) is improved or lessened depending on whether you can be more or less sure that reasonable alternative explanations can be ruled out. Thinking up these alternatives is crucial to logical reasoning. Without this creativity you can be more easily led away from the truth, that is, conned.

## III. Creating Scientific Explanations

The power to explain is a mark of your having discovered the truth. Those who can explain more know more. Since at least the fourth century BCE, we have known that the Earth is round, not flat.[[7]](#footnote-7) How did we draw this conclusion? Not by gathering many positive reports from people declaring that the Earth is round while failing to receive any negative reports declaring it to be flat or simply by photographing the Earth from space as we first did in 1946. The evidence was more indirect: the hypothesis that the Earth is round enabled so many things to be explained that otherwise were unexplainable.



By assuming that the Earth is round we can explain why Magellan's ship could keep sailing west from Spain yet return to Spain. By assuming that the Earth is round we can make sense of the shape of eclipses of the moon (they are round shadows of our round Earth). By assuming that the Earth is round we can explain why, when we look away from port with our telescope at a ship sailing toward the horizon, the top of the ship disappears after the bottom, not before. By assuming that the Earth is round we can explain why the sun can shine at midnight in the arctic. All these facts would be deep mysteries without the round-Earth hypothesis, and it would be nearly a miraculous coincidence if all these facts fit so well with an assumption that was false; therefore, the assumption is a fact. The moral is that **science is propelled forward by its power to explain**.

### Probabilistic and Deterministic Explanations

The best explanations of an event usually give us a good reason to have expected the event. Suppose you want to explain why apples fall off from the apple tree and hit the ground. One untestable explanation would be that it was the apple's "time" to leave the tree. That explanation appeals to a supernatural notion of fate or destiny. A scientific explanation is that the apple fell because it absorbed enough water through its stem that its weight increased above the maximum downward force that the brittle stem could resist.

Because explaining people's behavior is harder than explaining the behavior of apples, the current principles of psychology are less precise than the principles of physics. Psychologists depend on rules of thumb; physical scientists have deterministic laws that indicate what will happen rather than what might happen. For example, why did Sarah decide not to go out with Wayne when he mentioned he had an extra ticket to the concert? After talking with her, a psychologist might explain her action this way:

1. Wayne suggested that Sarah spend her time doing something she believed wouldn't be interesting to her.
2. People will not usually do what they have little interest in doing, nor what they perceive to be against their self-interest.

Sentence 1 states the relevant initial facts of the situation, and sentence 2 expresses the relevant law of psychology. This law is less precise than the law of gravity. It is only probabilistic, not deterministic, because it doesn't say what will happen but only what probably will happen. Using 1 and 2 in advance, we could predict only what Sarah probably would do, not what she will do. Psychology can't give a deterministic explanation. Such is the current state of that science.

Suppose you asked why you can see through glass but not through concrete, and you were told: “Because glass is transparent.” That answer is appropriate for an elementary school student, but not for a more sophisticated audience. After all, transparent merely means being able to be seen through. The explanation is trivial. Up until 1926, however, no one had a better explanation. Glass’s being transparent was just one of the brute facts of nature. It was accepted, but no deeper explanation could show why. Then, in 1926, the theory of quantum mechanics was discovered. From the principles of quantum mechanics, it was possible to deduce that anything made of glass should permit light to pass through. Similarly, quantum mechanics allowed us to find out why water is wet. These examples illustrate two main points: (1) General theories are more valuable than mere collections of specific facts, because with a general theory you can explain a large variety of individual facts. (2) If you can deduce a phenomenon from some well- accepted principles, you have a much deeper explanation of the phenomenon than if you can't carry out this deduction.

### Fruitful and Unfruitful Explanations

Untestable explanations are avoided by good scientists, but fruitful explanations are highly valued. To appreciate this virtue of fruitfulness, consider the scientists' favorite explanation of what caused the demise of the dinosaurs 65 million years ago. Four explanations or specific theories have been proposed in the scientific literature: the sex theory, the drugs theory, the violence theory, and the crime theory.

* According to the sex theory, 65 million years ago the world's temperature increased a few degrees. This increase warmed the male dinosaurs' testes to the point that they became infertile.
* According to the drug theory, 65 million years ago the world's first psychoactive (mind altering) plants evolved. Dinosaurs ate these plants, overdosed, and died.
* According to the violence theory, 65 million years ago some violent global event—perhaps caused by an asteroid or volcano—led to the dinosaur extinctions.
* According to the crime theory, 65 million years ago the first small mammals got braver and more clever. Some mammals learned to steal dinosaur eggs, which caused the dinosaur extinctions.

Of all four theories, current science favors the violence theory. Why? There are two reasons: it has been successfully tested, and it has been fruitful. The other three theories are testable in principle, but they are too hard to test in practice. The soft parts of male dinosaurs don't leave fossils, so the sex theory cannot be tested by looking for fossil remains. The drug theory is too hard to test because nothing much is known about which drugs were in which plants so long ago. The crime theory is too hard to test because there is no practical way to check whether little mammals did or didn't steal the dinosaur eggs. On the other hand, the violence theory can be. Suppose a violent global event threw dust into the air, darkening the Earth, leading to cold weather and the end of most plant photosynthesis. Digging down to the 65-million-year layer should reveal a thin layer of dust, no matter where in the world the scientists dig down. And indeed, scientists have discovered a layer of dust there containing a high concentration of a very rare element, iridium. Although naturally scarce on the Earth's surface, the element is relatively abundant both in asteroids and deep inside volcanoes.

In addition to its having stood up to this observational test, the violence theory is favored because it is so fruitful. That is, scientists can imagine many interesting and practical ways in which the theory can be tested. They can search satellite photos looking for 65-million-year-old asteroid craters. At suspected crater sites, they can analyze rocks for signs of impact—tiny fractures in shocked quartz. Digging might reveal pieces of an asteroid. A large speeding asteroid would ionize the surrounding air, making it as acidic as the acid in a car battery, so effects of this acidity might be discovered. Imagine what that rain would do to your car's paint. Scientists can also examine known asteroids and volcanoes for unusual concentrations of other chemical elements in addition to iridium. Ancient beaches can be unearthed to look for evidence of a huge tidal wave having hit them 65 million years ago. All these searches and examinations are under way today, and there has been much success in finding data consistent with the violence theory and little uncontested counterevidence.

Thus, the violence theory is the leading contender for explaining the dinosaur extinctions not because the alternative explanations have been refuted but because of its being successfully tested (so far) and its being so fruitful.

This brings us to the edge of a controversy about scientific methodology. The other alternative theories of dinosaur extinctions have not been refuted; they have not even been tested. But if they have not been refuted, and if proving the violence theory requires refuting all the alternative theories, doesn't it follow that the violence theory will never be proved, no matter how much new positive evidence is dug up by all those searches and examinations mentioned above? This question cannot be answered easily. We will end our discussion of this problem about scientific reasoning with the comment that not only is there much more to be learned about nature, but there are also unsolved problems about the nature of the science itself.

## IV. Testing Scientific Explanations

If you don’t test the claim, you don’t know it’s true.

### Designing a Scientific Test

It is easy to agree that scientific generalizations should be tested before they are proclaimed as true, and it is easy to agree that the explanations based on those generalizations also should be tested. However, how do you actually go about testing them? The answer is not as straightforward as one might imagine. The way to properly test a generalization differs dramatically depending on whether the generalization is universal (all A are B) or non-universal (some but not all A are B). When attempting to confirm a universal generalization, it is always better to focus on refuting the claim than on finding more examples consistent with it. That is, look for negative evidence, not positive evidence. For example, if you are interested in whether all cases of malaria can be cured by drinking quinine, it would be a waste of research money to seek confirming examples. Even 20,000 such examples would be immediately shot down by finding just one person who drank quinine but was not cured. On the other hand, suppose the generalization were non-universal instead of universal, that is, that most cases of malaria can be cured by drinking quinine. Then the one case in which someone drinks quinine and is not cured would not destroy the generalization. With a non-universal generalization the name of the game would be the ratio of cures to failures. In this case, 20,000 examples would go a long way toward improving the ratio.

There are other difficulties with testing. For example, today's astronomers say that all other galaxies on average are speeding away from our Milky Way galaxy because of the Big Bang explosion. This explosion occurred 13.7 billion years ago, when the universe was smaller than the size of a pea. Can this explanation be tested to see whether it is correct? You cannot test it by rerunning the birth of the universe. But you can test its predictions. One prediction that follows from the Big Bang hypothesis is that microwave radiation of a certain frequency will be bombarding Earth from all directions. This test has been run successfully, which is one important reason why today's astronomers generally accept the Big Bang as the explanation for their observations that all the galaxies on average are speeding away from us. There are several other reasons for the Big Bang theory having to do with other predictions it makes of phenomena that do not have good explanations by competing theories.

We say a hypothesis is confirmed or proved if several diverse predictions are tested and all are found to agree with the data while none disagree. Similarly, a hypothesis gets refuted if any of the actual test results do not agree with the prediction. However, this summary is superficial—let's see why.

### Retaining Hypotheses Despite Negative Test Results

If a scientist puts a hypothesis to the test, and if the test produces results inconsistent with the hypothesis, there is always some way or other for the researcher to hold onto the hypothesis and change something else. For example, if the meter shows “7” when your hypothesis would have predicted “5,” you might rescue your hypothesis by saying that your meter wasn't working properly. However, unless you have some good evidence of meter trouble, this move to rescue your hypothesis in the face of disconfirming evidence commits the fallacy of ad hoc rescue. If you are going to hold on to your hypothesis no matter what, you are in the business of propaganda and dogma, not science. Psychologically, it is understandable that you would try to rescue your cherished belief from trouble. When you are faced with conflicting data, you are likely to mention how the conflict will disappear if some new assumption is taken into account. However, if you have no good reason to accept this saving assumption other than that it works to save your cherished belief, your rescue is an ad hoc rescue.

In 1790 the French scientist Lavoisier devised a careful experiment in which he weighed mercury before and after it was heated in the presence of air. The remaining mercury, plus the red residue that was formed, weighed more than the original. Lavoisier had shown that heating a chemical in air can result in an increase in weight of the chemical. Today, this process is called oxidation. But back in Lavoisier’s day, the accepted theory on these matters was that a posited substance, “phlogiston,” was driven off during any heating of a chemical. If something is driven off, then you would expect the resulting substance to weigh less. Yet Lavoisier’s experiments clearly showed a case in which the resulting substance weighed more. To get around this inconsistency, the chemists who supported the established phlogiston theory suggested their theory be revised by assigning phlogiston negative weight. The negative-weight hypothesis was a creative suggestion that might have rescued the phlogiston theory. It wasn't as strange then as it may seem today because the notion of mass was not well understood. Although Isaac Newton had believed that all mass is positive, the negative-weight suggestion faced a more important obstacle. There was no way to verify it independently of the phlogiston theory. So, the suggestion appeared to commit the fallacy of ad hoc rescue.

A new hypothesis can avoid the charge of committing the fallacy of ad hoc rescue if it can meet two conditions: (1) The hypothesis must be shown to be fruitful in successfully explaining phenomena that previously did not have an adequate explanation. (2) The hypothesis's inconsistency with previously accepted beliefs must be resolved[[8]](#footnote-8) without reducing the explanatory power of science. Because the advocates of the negative-weight hypothesis were unable to do either, it is appropriate to charge them with committing the fallacy. As a result of Lavoisier's success, and the failure of the negative-weight hypothesis, today's chemists do not believe that phlogiston exists.

And Lavoisier’s picture gets a prominent place in history:[[9]](#footnote-9)



### Three Conditions for a Well-Designed Test

As a good rule of thumb, three definite conditions should hold in any well-designed test. First, if you use an experiment or observation to test some claim, you should be able to deduce the predicted test result from the combination of the claim plus a description of the relevant aspects of the test's initial conditions. That is, if the claim is really true, the predicted test result should follow. Second, the predicted test result should not be expected no matter what; instead, the predicted result should be unlikely if the claim is false. For example, a test that predicts water will flow downhill is a useless test because water is expected to do so no matter what. Third, it should be practical to check on whether the test did or did not come out as predicted, and this checking should not need to presume the truth of the claim being tested. It does no good to predict something that nobody can check.[[10]](#footnote-10)

To summarize, ideally a good test requires a prediction that meets these three conditions; it is

1. deducible or at least probable, given that the claim is true,
2. improbable, given that the claim is false, and
3. verifiable.

A good test of a claim will be able to produce independently verifiable data that should occur if the claim is true but shouldn't occur if the claim is false.

### Deducing Predictions for Testing

Condition 1, the deducibility condition, is somewhat more complicated than a first glance might indicate. Suppose you suspect that one of your co-workers named Philbrick has infiltrated your organization to spy on your company's chief scientist, Oppenheimer. To test this claim, you set a trap. Philbrick is in your private office late one afternoon when you walk out declaring that you are going home. You leave a file folder labeled "Confidential: Oppenheimer's Latest Research Proposal” on your desk. You predict that Philbrick will sneak a look at the file. Unknown to him, your office is continually monitored on closed-circuit TV, so you will be able to catch him in the act.

Let's review this reasoning. Is condition 1 satisfied for your test? It is, if the following reasoning is deductively valid:

Philbrick has the opportunity to be alone in your office with the Oppenheimer file folder (the test's initial conditions). Philbrick is a spy (the claim to be tested). So, Philbrick will read the Oppenheimer file while in your office, (the prediction)

This reasoning might or might not be valid depending on a missing premise. It would be valid if a missing premise were the following:

If Philbrick is a spy, then he will read the Oppenheimer file while in your office if he has the opportunity and believes he won’t be detected doing it (background assumption).

Is that premise acceptable? No. You cannot be that sure of how spies will act. The missing premise is more likely to be the following hedge:

If Philbrick is a spy, then he will probably read the Oppenheimer file while in your office if he has the opportunity and believes he won’t be detected doing it (new background assumption).

Although it is more plausible that this new background assumption is the missing premise used in the argument for the original prediction, now the argument isn't deductively valid. That is, the prediction doesn't follow with certainty, and condition 1 fails. Because the prediction follows inductively, it would be fair to say that condition 1 is "almost" satisfied. Nevertheless, it is not satisfied. Practically, though, you cannot expect any better test than this; there is nothing that a spy must do that would decisively reveal the spying. Practically, you can have less than ideal tests about spies or else no tests at all.

In response to this difficulty with condition 1, should we alter the definition of the condition to say that the prediction should follow either with certainty or probability? No. The reason why we cannot relax condition 1 can be appreciated by supposing that the closed-circuit TV does reveal Philbrick opening the file folder and reading its contents. Caught in the act, right? Your conclusion: Philbrick is a spy. This would be a conclusion many of us would be likely to draw, but it is not one that the test justifies completely. Concluding with total confidence that he is a spy would be drawing a hasty conclusion because there are alternative explanations of the same data. For example, if Philbrick were especially curious, he might read the file contents yet not be a spy. In other words, no matter whether the prediction comes out to be true or false, you cannot be sure the claim is true or false. So, the test is not decisive because its result doesn't settle which of the two alternatives is correct.

Yet being decisive is the mark of an ideally good test. We would not want to alter condition 1 so that this indecisive test can be called decisive. Doing so would encourage hasty conclusions. So the definition of condition 1 must stay as it is. However, we can say that if condition 1 is almost satisfied, then when the other two conditions for an ideal test are also satisfied, the test results will tend to show whether the claim is correct. In short, if Philbrick snoops, this tends to show he is a spy. More testing is needed if you want to be surer.

This problem about how to satisfy condition 1 in the spy situation is analogous to the problem of finding a good test for a non-universal generalization. If you suspect that most cases of malaria can be cured with quinine, then no single malaria case will ensure that you are right or that you are wrong. Finding one case of a person whose malaria wasn't cured by taking quinine doesn't prove your suspicion wrong. You need many cases to adequately test your suspicion.

The bigger issue here in the philosophy of science is the problem of designing a test for a theory that is probabilistic rather than deterministic. To appreciate this, let’s try another scenario. Suppose your theory of inheritance says that, given the genes of a certain type of blue-eyed father and a certain type of brown-eyed mother, their children will have a 25 percent chance of being blue-eyed. Let's try to create a good test of this probabilistic theory by using it to make a specific prediction about one couple's next child. Predicting that the child will be 25 percent blue-eyed is ridiculous. On the other hand, predicting that the child has a 25 percent chance of being blue-eyed is no specific prediction at all about the next child. Specific predictions about a single event can't contain probabilities. What eye color do you predict the child will have? You should predict it will not be blue-eyed. Suppose you make this prediction, and you are mistaken. Has your theory of inheritance been refuted? No. Why not? Because the test was not decisive. The child's being born blue-eyed is consistent with your theory's being true and also with its being false. The problem is that with a probabilistic theory you cannot make specific predictions about just one child. You can predict only that, if there are many children, then 25 percent of them will have blue eyes and 75 percent won't. A probabilistic theory can be used to make predictions only about groups, not about individuals.

The analogous problem for the spy in your office is that when you tested your claim that Philbrick is a spy you were actually testing a probabilistic theory because you were testing the combination of that specific claim about Philbrick with the general probabilistic claim that spies probably snoop. They don’t always snoop. Your test with the video camera had the same problem with condition 1 as your test with the eye color. Condition 1 was almost satisfied in both tests, but strictly speaking it wasn't satisfied in either.

Our previous discussion should now have clarified why condition 1 is somewhat more complicated than a first glance might indicate. Ideally, we would like decisive tests or, as they are also called, crucial tests. Practically, we usually have to settle for tests that only tend to show whether one claim or another is true. The stronger the tendency, the better the test. If we arrive at a belief on the basis of these less than ideal tests, we are always in the mental state of not being absolutely sure. We are in the state of desiring data from more tests of the claim so that we can be surer of our belief, and we always have to worry that someday new data might appear that will require us to change our minds. Such is the human condition. Science cannot do better than this.

## IV. Detecting Pseudoscience

The word science has positive connotations, the word pseudoscience has negative connotations. Science gets the grant money; pseudoscience doesn't. Calling some statement, theory, or research program “pseudoscientific” suggests that it is silly or a waste of time. It is pseudoscientific to claim that the position of the planets at the time a person is born determines the person's personality and major life experiences. It is also pseudoscientific to claim that spirits of the dead can be contacted by mediums at seances. Astrology and spiritualism may be useful social lubricants, but they aren't scientific.

Despite a few easily agreed-upon examples such as these two, defining pseudoscience is difficult. One could try to define science and then use that to say pseudoscience is not science, or one could try to define pseudoscience directly. A better approach is to try to find many of the key features of pseudosciences. A great many of the scientific experts will agree that pseudoscience can be detected by getting a “no” answer to the first two questions or a “yes” answer to any of the remaining three:

1. Do the "scientists" have a theory to test?
2. Do the "scientists" have reproducible data that their theory explains better than the alternatives?
3. Do the "scientists" seem content to search around for phenomena that are hard to explain by means of current science; that is, do the scientists engage in mystery mongering?
4. Are the "scientists" quick to recommend supernatural explanations rather than natural explanations?
5. Do the "scientists" use the method of ad hoc rescue while treating their own views as unfalsifiable?

The research program that investigates paranormal phenomena is called parapsychology. What are the paranormal phenomena we are talking about here? They include astral travel, auras, psychokinesis (moving something without touching it physically), plant consciousness, psychic healing, speaking with the spirits, witchcraft, and ESP—that is telepathy (mind reading), clairvoyance (viewing things at a distance), and precognition (knowing the future).

None of the parapsychologists' claims to have found cases of cancer cures, mind reading, or foretelling the future by psychic powers have ever stood up to a good test. Parapsychologists cannot convincingly reproduce any of these phenomena on demand; they can only produce isolated instances in which something surprising happened.

Parapsychologists definitely haven't produced repeatable phenomena that they can show need to be explained in some revolutionary way.

Rarely do parapsychologists engage in building up their own theories of parapsychology and testing them. Instead, nearly all are engaged in attempts to tear down current science by searching for mysterious phenomena that appear to defy explanation by current science. Perhaps this data gathering is the proper practice for the prescientific stage of some enterprise that hopes to revolutionize science, but the practice does show that the enterprise of parapsychology is not yet a science.

Regarding point 1, scientists attack parapsychologists for not having a theory-guided research program. Even if there were repeatable paranormal phenomena, and even if parapsychologists were to quit engaging in mystery mongering, they have no even moderately detailed theory of how the paranormal phenomena occur. They have only simplistic theories such as that a mysterious mind power caused the phenomenon or that the subject tapped a reserve of demonic forces or that the mind is like a radio that can send and receive signals over an undiscovered channel. Parapsychologists have no more-detailed theory that permits testable predictions. Yet, if there is no theory specific enough to make a testable prediction, there is no science.

## V. Paradigms and Possible Causes

Your car's engine is gummed up today. This has never happened before. Could it be because at breakfast this morning you drank grapefruit juice rather than your usual orange juice? No, it couldn't be. Current science says this sort of explanation is silly. OK, forget the grapefruit juice. Maybe the engine is gummed up because today is Friday the 13th. No, that is silly, too. A scientist wouldn't even bother to check these explanations. Let's explore this intriguing notion of what science considers "silly" versus what it takes seriously. What causes the pain relief after swallowing a pain pill? Could it be the favorite music of the inventor? No, that explanation violates medical science's basic beliefs about what can count as a legitimate cause of what. Nor could the pain relief be caused by the point in time when the pill is swallowed. Time alone causes nothing, says modern science. The pain relief could be caused by the chemical composition of the pill, however, or perhaps by a combination of that with the mental state of the person who swallowed the pill. The general restrictions that a science places on what can be a cause and what can't are part of what is called the paradigm of the science. Every science has its paradigm.

That is, at any particular time, each science has its own particular problems that it claims to have solved; and, more important, it has its own accepted ways of solving problems that then serve as a model for future scientists who will try to solve new problems. These ways of solving problems, including the methods, standards, and generalizations generally held in common by the community of those practicing the science, is, by definition, the paradigm of that science.

The paradigm in medical science is to investigate what is wrong with sick people, not what is right with well people. For a second example, biological science can explain what causes tigers to like meat rather than potatoes in terms of evolutionary factors that give rise to this preference, but biologists would turn to chemists to determine how the preference involves the chemical makeup of the meat, and not the history of zipper manufacturing or the price of rice in China. The paradigm for biological science limits what counts as a legitimate biological explanation. When we take a science course or read a science book, we are slowly being taught the paradigm of that science and, with it, the ability to distinguish silly explanations from plausible ones. Silly explanations do not meet the basic requirement for being a likely explanation, namely coherence with the paradigm. Sensitivity to this consistency requirement was the key to understanding the earlier story about Brother Bartholomew. Scientists today say that phenomena should not be explained by supposing that Bartholomew or anybody else could see into the future; this kind of "seeing" is inconsistent with the current paradigm. It is easy to test whether people can foresee the future if you can get them to make specific predictions rather than vague ones. Successfully testing a claim that someone can foresee the future would be a truly revolutionary result, upsetting the whole scientific world-view, which explains why many people are so intrigued by tabloid reports of people successfully foretelling the future.

Suppose a scientist wants to determine whether adding solid carbon dioxide to ice water will cool the water below 32 degrees Fahrenheit (0 degrees Celsius). The scientist begins with two glasses containing equal amounts of water at the same temperature.

The glasses touch each other. Solid carbon dioxide is added to the first glass, but not the second. The scientist expects the first glass to get colder but the second glass not to. This second glass of water is the control because it is just like the other glass except that the causal factor being tested—the solid carbon dioxide—is not present in it. After twenty minutes, the scientist takes the temperature of the water in both glasses. Both are found to have cooled, and both are at the same temperature. A careless scientist might draw the conclusion that the cooling is not caused by adding the carbon dioxide, because the water in the control glass also got cooler. A more observant scientist might draw another conclusion, that the experiment wasn’t any good because the touching is contaminating the control. The two glasses should be kept apart during the experiment to eliminate contamination. The paradigm of the science dictates that the glasses not touch because it implies that glasses in contact will reach a common temperature in much faster than glasses not in contact.

For a second example of the contamination of experimental controls, suppose a biologist injects some rats with a particular virus and injects control rats with a placebo—some obviously ineffective substance such as a small amount of salt water. The biologist observes the two groups of rats to determine whether the death rate of those receiving the virus is significantly higher than the death rate of those receiving the placebo. If the test is well run and the data show such a difference, there is a correlation between the virus injection and dying. Oh, by the way, the injected rats are kept in the same cages with the control rats. Oops. This contamination will invalidate the entire experiment, won't it?

Reputable scientists know how to eliminate contamination, and they actively try to do so. They know that temperature differences and disease transmission can be radically affected by physical closeness. This background knowledge that guides experimentation constitutes another part of the paradigm of the sciences of physics and biology. Without a paradigm helping to guide the experimenter, there would be no way of knowing whether the control group was contaminated. There would be no way to eliminate experimenter effects, that is, the unintentional influence of the experimenter on the outcome of the experiment. There would be no way of running a good test. That fact is one more reason that so much of a scientist's college education is spent learning the science's paradigm.

## VI. Review of Major Points

When scientists are trying to gain a deep understanding of how the world works, they seek general patterns rather than specific facts. The way scientists acquire these general principles about nature is usually neither by deducing them from observations nor by inductive generalization. Instead, they think about the observations, then guess at a general principle that might account for them, then check this guess by testing. When a guess or claim is being tested, it is called a hypothesis. Testing can refute a hypothesis. If a hypothesis does not get refuted by testing, scientists retain it as a prime candidate for being a general truth of nature, a law. Hypotheses that survive systematic testing are considered to be proved, although even the proved statements of science are susceptible to future revision, unlike the proved statements of mathematics.

Scientific reasoning is not discontinuous from everyday reasoning, but it does have higher standards of proof. This chapter reviewed several aspects of scientific reasoning from earlier chapters, including general versus specific claims, testing by observation, testing by experiment, accuracy, precision, operational definition, pseudoprecision, the role of scientific journals, independent verification, consistency with well-established results, reproducible evidence, anecdotal evidence, a scientist’s cautious attitude and open mind, attention to relevant evidence, the scientific method of justifying claims, disconfirming evidence, and the methods of gaining a representative sample.

Deterministic explanations are preferred to probabilistic ones, and ideal explanations enable prediction of the phenomenon being explained. Explanations are preferred if they are testable and fruitful. A good test requires a prediction that is (1) deducible, (2) improbable, and (3) verifiable.

Science provides the antidote to superstition. There are criteria that can be used to detect pseudoscience.

A reasonable scientific explanation is coherent with the paradigm for that science. Only by knowing the science's paradigm can a scientist design a controlled experiment that does not contaminate the controls and that eliminates effects unintentionally caused by the experimenter.

1. This chapter is based on *Logical Reasoning* by Bradley H. Dowden. [↑](#footnote-ref-1)
2. <https://plato.stanford.edu/entries/scientific-method/>. Note, Aristotle’s work went well past considerations of the nature of scientific inquiry. He likewise wrote on logic, biology, astronomy, physics, economics, political theory, metaphysics and ethics (amongst other things). Aristotle was a member of Plato’s Academy; founded in 387 BCE, the Academy broke ground in a diverse number of scientific fields, including geometry. In fact, Euclid’s geometry is a codification of the work done in Plato’s institute. [↑](#footnote-ref-2)
3. Theorizing regarding the nature of scientific inquiry was its own field of study in the middle ages, whose insights proved essential to contemporary science (L. Laudan, L., “Theories of scientific method from Plato to Mach”, History of Science (1968): 7(1), 1–63. [↑](#footnote-ref-3)
4. However, it is a mistake to suppose that the incorporation of testing into scientific inquiry didn’t occur prior to the work of European scientists in the seventeenth and eighteenth centuries. For instance, attention to the importance of experimentation enabled Ibn al-Haytham (born c. 965, Basra, Iraq—died c. 1040, Cairo, Egypt) to make significant contribution to the science of optics. [↑](#footnote-ref-4)
5. The photo below by Solaris 2006. [↑](#footnote-ref-5)
6. Some textbook authors make some fantastic assumptions, don't they? [↑](#footnote-ref-6)
7. Plato proposes the idea in his book *Phaedo*, dated around 380 BCE. [↑](#footnote-ref-7)
8. There is an exception to this, however. A new hypothesis might be inconsistent with previously accepted beliefs because it is correct and the old views were not. In this case, we wouldn’t want to resolve the inconsistency, we want to set about finding more evidence for the new hypothesis. [↑](#footnote-ref-8)
9. Portrait of Antoine-Laurent Lavoisier And His Wife (1788) by Jacques-Louis David [↑](#footnote-ref-9)
10. These criteria for a good test are well described by Ronald Giere in Understanding Scientific Reasoning (New York: Holt, Rinehart and Winston, 1979), pp. 101-105. [↑](#footnote-ref-10)