2 What Is Science?

The wrong view of science betrays itself in the craving to be right; for it is not his possession of knowledge, of irrefutable truth, that makes the man of science, but his persistent and recklessly critical quest for truth.

Sir Karl Popper, *The Logic of Scientific Discovery*

So I left him, saying to myself, as I went away: Well, although I do not suppose that either of us knows anything really beautiful and good, I am better off than he is—for he knows nothing, and thinks that he knows. I neither know nor think that I know. In this latter particular, then, I seem to have slightly the advantage of him.

Socrates, in Plato's *Apology*

Test everything. Keep what is good.

Saint Paul, *First Letter to the Thessalonians*

OVERVIEW

- Comparative politics is the subfield of political science that focuses primarily on politics within countries. In Chapter 3 we define and examine the nature of politics. In this chapter we define and examine the nature of science.
- Science is a strategy for understanding and explaining the social and natural world that emphasizes the use of statements that can be examined to see whether they are wrong.
- Scientific explanations should explain previously puzzling facts, be logically consistent, and produce (many) potentially falsifiable predictions.
- All scientific explanations are tentative. We accept some explanations as provisionally true when they have withstood vigorous attempts at refutation more successfully than competing explanations.

Consider the following five statements. What do they all have in common?

1. Science is a collection of facts that tell us what we know about the world.
2. A scientific theory is one that has been proven.
3. "The sun revolves around the earth" is not a scientific statement.
4. If my theory is correct, then I should observe that rich countries are more likely to be democracies. I do observe that rich countries are more likely to be democracies. Therefore, my theory is correct.
5. Politics cannot be studied in a scientific manner.

The common element in these statements is that they are all, in some sense, wrong. Science is not a collection of facts that tell us what we know about the world. Scientific theories cannot be proven; thus, a scientific theory is not one that has been proven. The statement that the sun revolves around the earth is a scientific statement (even though it is false). The argument outlined in statement 4 is logically invalid; therefore, I cannot conclude that my theory is correct. And finally, politics can be studied in a scientific manner. We suspect that many of you will have thought that at least some of these statements were correct. To know why all of these statements about science are wrong, you will need to continue reading this chapter.

Science certainly has its detractors, largely because of what was experienced in the twentieth century. Some horrendous things were either done in the name of science or "justified" on scientific grounds or, at a minimum, made possible by science. Although we should never close our eyes to the harm that is sometimes done with science, we believe that it is as much a mistake to blame science for what some scientists have done in its name as it is to blame religion for what some believers have done in its name.

But what is science? First and foremost, science is a method; however, it is also a culture. The epigraphs at the start of this chapter are meant to capture what we might call the "culture" of science. Some of the negative views of science come from what people perceive the culture of science to be—cold, calculating, self-assured, arrogant, and, perhaps, even offensive. We believe, however, that at its best, the culture of science displays the characteristics encouraged by the otherwise very different thinkers who are quoted. The scientific method is, at its very core, a critical method, and those reflective individuals who use it are much more likely to be humbled than emboldened. Sir Karl Popper ([1959] 2003) reminds us that science is not a static set of beliefs to be conserved and that all knowledge is tentative. Socrates reminds us that an acute awareness of our own ignorance is always the first step toward knowledge. Saint Paul offers hope that our willingness to test all of our ideas will leave us something good to hang on to. As we'll demonstrate in this chapter, science isn't about certainty, it isn't merely about the orderly collection of facts, and it isn't about invoking authority to protect our ideas from uncomfortable evidence. Instead, science is about asking tough questions and providing answers that invite criticism. Science is about recognizing the limits of our knowledge without lapsing into irresponsible cynicism. And science is about using the best logic, methods, and evidence available to provide answers today, even though we recognize that they may be overturned tomorrow.

Comparative politics is a subfield of political science. But what exactly is political science? Well, it is the study of politics in a scientific way. How's that for a tautology? It is easy to see that, as it stands, this definition is not particularly informative. For example, what is politics? And what is science? In the next chapter we answer the first of these questions and seek to demarcate politics from other forms of social phenomena. In this chapter, though, we focus on the second question—what is science? Our goal is to provide an answer that resembles the way most practicing scientists would answer this question.

WHAT IS SCIENCE?

Is science simply a body of knowledge or a collection of facts, as many of us learn in high school? While there was a time when many scientists may have defined science in this way, this definition is fundamentally unsatisfactory. If this definition of science were accurate, then many of the claims about how the universe worked, such as those developed through Newtonian physics, would now have to be called unscientific, because they have been replaced by claims based on more recent theories, such as Einstein's theory of relativity. Moreover, if science were simply a collection of statements about how the world works, then we would not be able to appeal to science to justify our knowledge of the world without falling into the following circular reasoning:

"Science is a collection of statements about how the world works."
"How do we know if these statements are accurate?"
"Well, of course they're accurate! They're scientific!"

The body of knowledge that we call "scientific" may well be a product of science, but it is not science itself. Rather, science is a method for provisionally understanding the world. The reason for saying "provisionally" will become clear shortly. Science is one answer to the central question in epistemology (the study of knowledge): "How do we know what we know?" The scientist's answer to that question is, "Because we have subjected our ideas to the scientific method." Science, as Karl Popper indicates in one of the epigraphs at the start of this chapter, is the quest for knowledge. At this point, you might say that there are many ways to seek knowledge. Does this mean that meditation, reading scripture, and gazing at sunsets are all scientific activities? Although we agree that these are all ways of seeking knowledge, none of them is scientific. Science is a particular quest for knowledge. To use Popper's phrase, it is the "recklessly critical" pursuit of knowledge, in which the scientist continually subjects her ideas to the cold light of logic and evidence.

Although science is not the only route to knowledge, it may be unique in its emphasis on self-criticism. Scientists, like other scholars, can derive their propositions from an infinite number of sources. For example, Gregory Derry (1999) tells the story of how August Kekulé made an extremely important scientific breakthrough while hallucinating—half asleep—in front of the fireplace in his laboratory one night. He had spent days struggling to understand the spatial arrangement of atoms in a benzene molecule. In a state of mental and physical exhaustion, his answer appeared to him as he "saw" swirls of atoms joined in a particular formation dancing among the embers of his fireplace. In a flash of inspiration, he saw how the pieces of the puzzle with which he had been struggling fit together. This inspired understanding of the physical properties of organic compounds did not become a part of science that night, though. It did so only after the implications of his vision had withstood the critical and sober onslaught that came with the light of day. Thus, although flashes of insight can come from a variety of sources, science begins only when one asks, "If that is true, what else ought to be true?" And it ends—if ever—when researchers are satisfied that they have taken every reasonable pain to show that the implications of the insight are false and have failed to do so. Even then, however, the best answer is not the final answer—it is just the best "so far."

So, science is the quest for knowledge that relies on criticism. The thing that allows for criticism is the possibility that our claims, theories, hypotheses, ideas, and the like could be wrong. Thus, what distinguishes science from "non-science" is that scientific statements must be **falsifiable**—there must be some imaginable observation or set of observations that could falsify or refute them. This does not mean that a scientific statement will ever be falsified, just that there must be a possibility that it could be falsified if the "right" observation came along. Only if a statement is potentially testable is it scientific. We deliberately say "potentially testable" because a statement does not have to have been tested to be scientific; all that is required is that we can conceive of a way to test it.[1]

Scientific statements must be **falsifiable**. This means that they are potentially testable—there must be some imaginable observation that could falsify or refute them.

What sorts of statements are not falsifiable? Tautologies are not falsifiable because they are true by definition. For example, the statement "Triangles have three sides" is a **tautology**. It is simply not possible ever to observe a triangle that does not have three sides because *by definition* if an object does not have three sides, it is not a triangle. It is easy to see that this statement is not testable and hence unscientific. Tautologies, though, are not always so easy to spot. Consider the following statement: "Strong states are able to overcome special interests in order to implement policies that are best for the nation." Is this a tautology? This statement may be true, but unless we can think of a way to identify a strong state without referring to its ability to overcome special interests, then it is just a definition

A **tautology** is a statement that is true by definition.

1. Indeed, a statement can be scientific even if we do not currently have the data or the technical equipment to test it. Our upcoming discussion of Einstein's special theory of relativity illustrates this point quite clearly.

and is, therefore, unscientific. In other words, whether this particular statement is scientific depends on how strong states are defined.

Other statements or hypotheses are not falsifiable, not because they are tautological, but because they refer to inherently unobservable phenomena. For example, the claims "God exists" and "God created the world" are not falsifiable because they cannot be tested; as a result, they are unscientific. Note that these claims may well be true, but it is important to recognize that science has nothing to do with the truth or falsity of statements. All that is required for a statement to be scientific is that it be falsifiable. It should be clear from this that we are not claiming that "nonscience" is nonsense or that it lacks meaning—this would clearly be a mistake. Nonfalsifiable statements like "God exists" may very well be true and have important and meaningful consequences—our claim is simply that they do not form a part of science. Having defined science as a critical method for learning about the world, we can now evaluate the basic elements of the scientific method in more detail.

THE SCIENTIFIC METHOD

Although there is no **scientific method** clearly written down that is followed by all scientists, it is possible to characterize the basic features of the scientific method in the following manner.

The **scientific method** describes the process by which scientists learn about the world.

Step 1: Question

The first step in the scientific process is to observe the world and come up with a question or puzzle. The very need for a theory or explanation begins when we observe something that is so unexpected or surprising that we ask, "Why did that occur?" Note that the surprise that greets such an observation, and that makes the observation a puzzle worth exploring, implies that the observation does not match some prior expectation or theory that we held about how the world works. Thus, we always have a preexisting theory or expectation when we observe the world; if we did not have one, we could never be surprised, and there would be no puzzles.

Step 2: Theory or Model

Once we have observed something puzzling, the next step is to come up with a theory or model to explain it. In what follows, we will talk of theories, models, and explanations interchangeably. Scientists use the word **theory** to describe a set of logically consistent statements that tell us why the things that we observe occur. It is important that these statements be logically consistent; otherwise we have no way of determining what their empirical predictions will be and, hence, no way to test them. Put differently, theories that are logically

A **theory** is a set of logically consistent statements that tell us why the things that we observe occur. A theory is sometimes referred to as a model or an explanation.

inconsistent should not, indeed cannot, be tested, because we have no way of knowing what observations would truly falsify them.

The principle of the **uniformity of nature** asserts that nature's operating mechanisms are unchanging in the sense that if *X* causes *Y* today, then it will also cause *Y* tomorrow and the next day and so on.

Most philosophers of science assume that all phenomena occur as a result of some recurring process. The principle of the **uniformity of nature** asserts that nature's operating mechanisms are unchanging in the sense that if *X* causes *Y* today, then it will also cause *Y* tomorrow and the next day and so on. If it does not, then we should not consider *X* a cause. Be careful to note that the principle of uniformity is a statement not that nature is unchanging, only that the laws of nature do not change (although our understanding of those laws will likely change over time). This is an important principle, because if this principle is rejected, we must accept the possibility that things "just happen." That is, we must accept that things happen for no reason. Casual observation of the sometimes maddening world around us suggests that this may, indeed, be true, but it is the job of scientists to attempt to impose order on the apparent chaos around them. In the social world, this process often begins by dividing the behavior we observe into systematic and unsystematic components. The social scientist then focuses her attention on explaining only the systematic components.[2]

So what should theories or models look like? It is useful to think of our starting puzzle or observation as the end result of some previously unknown process (Lave and March 1975). We can then speculate about what (hidden) processes might have produced such a result. In effect, we try to imagine a prior world that, if it had existed, would have produced the otherwise puzzling observation before us. This prior world then becomes our model explaining the observation.

Notice that this process of imagining prior worlds is one place—but surely not the only one—where imagination and creativity enter the scientific process. What scientists do to stimulate this creative process is itself not part of the scientific method. Essentially, anything goes. Nobel Prize–winning physicist Richard Feynman, who himself spent a lot of time hanging out in bars and playing Brazilian hand drums, describes science as "imagination in a straightjacket"—it is imagination constrained by what we already know about the world (Feynman 1967). Consequently, he suggests that there is no point engaging in flights of fancy about things that we know cannot exist (like antigravity machines). Whatever means we use to stimulate speculation about a prior world, if we can show through logical deduction that *if* that prior world existed, it would have produced the puzzling observation we started with, then we have a theory, or model. Note that we have only *a* theory; we do not necessarily have *the* theory. This is why we continually test the implications of our theory.

The model that we end up with will necessarily be a simplified picture of the world. It is impossible to have a descriptively accurate model of the world as an infinite number of

2. This suggests that you should be wary of anyone who tells you that you need to know everything before you can know anything.

details would have to be captured in such a model. Pure description is impossible—models are always going to leave many things out. As with all arts, much of the skill of modeling is in deciding what to leave out and what to keep in. A good model contains only what is needed to explain the phenomenon that puzzles us and nothing else. If we made our models too complex, we would have no way of knowing which elements were crucial for explaining the puzzling observation that we started with and which were superfluous. The purpose of a model is not to describe the world but to explain it, so descriptive accuracy is not a core value in model building. Details are important only to the extent that they are crucial to what we are trying to explain. For example, if we are interested in explaining an aircraft's response to turbulence, it is not important whether our model of the aircraft includes LCD screens on the back of the passengers' seats. In fact, such inconsequential details can easily distract our attention from the question at hand. Another benefit of simple models is that they invite falsification because they make it very clear what we should not observe. The more amendments and conditions placed on an explanation, the easier it is for scholars to dismiss apparently contradictory evidence.

It is important to remember that models are always developed with a specific goal in mind. This means that we should evaluate models in terms of how useful they are for achieving that goal. As the late Dutch economist Henri Theil (1971) once said, "models should be used, not believed." To emphasize this point, it can be helpful to think of models as being similar to maps. Like models, maps are simplified pictures of the world designed for a specific purpose. Consider the subway map of any city. The subway map is always a simplification of the city and, indeed, an inaccurate simplification in the sense that it provides inaccurate information about the relative distances between, and geographic positions of, particular locations. Despite this, the map is incredibly useful if one's goal is to move efficiently around the city using the subway system—the purpose for which the map was designed. Of course, this map would be less useful if one's goal was to walk above ground from one location to another. As with a map, one must not judge the value of a model in some abstract sense but in terms of how well it helps us understand some particular aspect of the world and explain it to others.

Step 3: Implications (Hypotheses)

Once we have a model, the third step in the scientific process is to deduce implications from the model other than those that we initially set out to explain. Why do we say "other than those that we initially set out to explain"? Well, presumably the model that we construct will provide a logical explanation for the puzzling observation that we started with; after all, that is what it was designed to do! In other words, there is no way that a model can ever be falsified if only the observations that were employed to develop the model in the first place are used to test it. To actually test the model and allow for the possibility that it will be falsified, we will have to find other implications that can be deduced from it. We must ask ourselves, "If the prior world that we created to explain the phenomena that we originally found puzzling really did exist, what else ought to exist? What else should we be able to observe?"

As before, there is often room for incredible imagination here, because the complete list of logical implications of a model is seldom self-evident.

Good models are those that produce many different implications. This is so because each prediction represents another opportunity for the model to fail and, therefore, makes the model easier to falsify. This is good because if the model fails to be falsified, we gain more confidence in its usefulness. Fertile models—models with many implications—are also desirable because they encourage the synthesis of knowledge by encouraging us to see connections between ostensibly disparate events. Good models also produce surprising implications—they tell us something we would not know in the absence of the model. Models are not particularly useful if they tell us only what we already know. Surprise, however, is best appreciated in small doses. If every implication of a model is surprising, then either everything we thought about the world is wrong, or the model is.

Step 4: Observe the World (Test Hypotheses)

The fourth step is to examine whether the implications of the model are consistent with observation. Remember that the goal is not to dogmatically uphold the implications of our model or defend them in order to prove how right they are. On the contrary, we should try our best to falsify them, because it is only after a theory has withstood these attempts to overthrow it that we can reasonably start to have confidence in it. Although as many of the model's implications as possible should be tested, testing those that are most likely to be falsified is particularly important. Always submit a model to the harshest test that you can devise.

It is standard practice to stop and ask if other models—models that describe altogether different processes—might also explain the phenomena of interest. When this is the case (and it almost always is), it is incumbent upon the scientist to compare the implications of those other models with the implications of her own model. Although it is always the case that competing models have some of the same implications (otherwise they could not explain the same observations to begin with), it is typically the case that they will differ in some of their implications (otherwise they are not different models). The trick for a researcher is to identify these points of conflict between the different models and identify the relevant observations in the real world that would help her decide between them. This is what scientists refer to as a **critical test**. Ultimately, if a critical test is possible, observation will prove decisive in choosing between the models. This is because we know that there is only one world and the creative scientist has managed to get competing theories to say contradictory things about it—only one of the models can be consistent with the real world.

A **critical test** allows the analyst to use observation to distinguish between two or more competing explanations of the same phenomenon.

Step 5: Evaluation

If we observe the implications deduced from our theory, we say that our theory has been corroborated. Note that we cannot say that our theory has been verified or proven. This

Box 2.1

AN EXAMPLE OF THE SCIENTIFIC PROCESS

The Case of Smart Female Athletes

Because student athletes often miss classes to compete out of state, they frequently submit a letter from the athletic director asking for cooperation from their professors. Over the years, a certain professor has noticed through casual observation that women engaged in athletic competition frequently perform better academically than the average student. It is puzzling why female athletes would perform better in spite of missing classes. Can you think of a model—a process—that might produce such a puzzling observation?

You might start with the following conjecture:

- Female athletes are smart.

This is an explanation, but it is not a particularly good one. For example, it comes very close to simply restating the observation to be explained. One thing that could improve the explanation is to make it more general. This might lead you to a new explanation:

- Athletes are smart.

This model is certainly more general (but not necessarily more correct). Still, there are at least two problems with this model as things stand. First, it has no sense of process; it basically says that athletes share some inherent quality of smartness that leads them to perform better academically. In effect, this only pushes the phenomenon to be explained back one step; that is, we now need to know why athletes are smart. Second, the model comes close to being a tautology. It essentially says that athletes perform better academically because they are defined as being smart. This is problematic, as we saw earlier, because tautologies are not falsifiable—they cannot be tested; hence, they are not part of the scientific endeavor.

This might lead you to look for a new explanation or model that includes some sort of process that makes female athletes appear smart. You might come up with the following model:

- Being a good athlete requires a lot of hard work; performing well academically in college requires a lot of work. Students who develop a strong work ethic in athletics are able to translate this to their studies.

This is a much more satisfying model because it provides a process or mechanism explaining why female athletes might be more academically successful than other students. An appealing feature of the model is that the logic of the argument applies not only to female athletes but to any athlete. Indeed, it applies to any person involved in an activity that rewards hard work. Thus, we might generalize this model by removing the specific reference to athletes:

- **Work Ethic Theory:** Some activities provide a clear, immediate, and tangible reward for hard work—in fact, they may provide an external stimulus to work hard (coaches shouting through bullhorns, manipulating rewards and punishments based on effort, and so on). Individuals who engage in these activities develop a habit of working hard and so will be successful in other areas of life as well.

At this point, you should stop and ask yourself whether there are any alternative explanations for why female athletes are successful. Can you think of any? One alternative explanation is the following:

- **Excellence Theory:** Everyone wants to feel successful, but some people go long periods without success and become discouraged. Those individuals who experience success in one area of their life (perhaps based on talent, rather than hard work) develop a "taste" for it and devise strategies to be successful in other parts of their life. Anyone who achieves success in nonacademic areas, such as athletics, will be more motivated to succeed in class.

Another alternative explanation is the following:

- **Gender Theory:** In many social and academic settings, women are treated differently from men. This differential treatment often leads women to draw inferences that certain activities are "not for them." Because many athletic endeavors are gender specific, they provide an environment for women to develop their potential free from the stultifying effects of gender bias. The resulting sense of efficacy and autonomy encourages success when these women return to gendered environments like the classroom.

We now have three different or competing models, all of which explain the puzzling observation that we started with. But how can one evaluate which model is best? One way is to test some of the implications that can be derived from these theories. In particular, we would like to find some new question(s) to which the three models give different answers. In other words, we would like to conduct a critical test that would allow us to choose among the alternative reasonable models.

We might start by wondering whether being an athlete helps the academic performance of women more than men. Whereas the Work Ethic Theory and the Excellence Theory both predict that being an athlete will help men and women equally, the Gender Theory predicts that female athletes will perform better than nonathletic women but that male athletes will have no advantage over nonathletic men. Thus, collecting information on how well male and female athletes perform in class relative to male and female nonathletes, respectively, would allow us to distinguish between the Gender Theory and the other theories.

But how can we distinguish between the Excellence Theory and the Work Ethic Theory? One difficulty frequently encountered when trying to devise critical tests is that alternative theories do not always produce clearly differentiated predictions. For example, we just saw that the Excellence Theory and the Work Ethic Theory both predict that athletics will help men and women academically. It turns out that these two theories have other predictions in common as well. The Excellence Theory clearly suggests that success in any nonacademic area of life is likely to encourage academic success. In other words, the Excellence Theory predicts that academic success will be associated with success in other areas of life. The problem is that success in many of these nonacademic areas may require hard work. As a result, if we observe, for instance, accomplished musicians performing well in our political science classes, it will be difficult to discern whether this is because they learned the value of hard work in music and transferred it to political science (Work Ethic Theory) or because they developed a “taste” for success as musicians that then inspired success in political science (Excellence Theory). In effect, the Excellence Theory and the Work Ethic Theory both predict that academic success will be associated with success in other areas of life.

If we want to distinguish between the Work Ethic Theory and the Excellence Theory, we need to imagine observations in which they produce different expectations. Sometimes, this requires further development of a theory. For example, we might expand the Excellence Theory to say that those people who develop a taste for excellence also develop a more competitive spirit. If this is true, then the Excellence Theory would predict that student athletes are likely to be more competitive and will perform better than other students even when playing relatively frivolous board games. Since even the most driven athletes are not likely to devote time to training for board games, the Work Ethic Theory predicts that athletes will perform the same as nonathletes in such trivial pursuits. Thus, we could look at the performance of athletes and nonathletes at board games to distinguish between the Excellence Theory and the Work Ethic Theory.

The three critical tests that we have come up with and their predictions are listed in Table 2.1. All that is now required is to collect the appropriate data and decide which model, if any, is best.

It is worth noting that there is considerable overlap between the predictions of our three theories. This is often the case in political science settings as well. The crucial point is not that each theory should yield a complete set of unique predictions, but that our theories should have sufficiently many distinct predictions that we can use observation to help us make decisions about which theories to embrace, however tentatively. Table 2.1 lists just some of the predictions that might help us to distinguish between the three theories outlined above. Can you think of any more?

Table 2.1 Three Critical Tests

	Theory		
Question	**Gender**	**Excellence**	**Work ethic**
Will athletics help women more than men?	Yes	No	No
Is academic success associated with success in other areas of life?	No	Yes	Yes
Are female athletes more successful at board games than women who are not athletes?	Yes	Yes	No

important point is one that we will return to in more detail in the next section of this chapter.[3] The fact that we can never prove a scientific explanation is why we earlier called science a method for "provisionally" understanding the world. Our theory may or may not be true. All we can conclude, if observations are consistent with our theoretical implications, is that our theory has not yet been falsified; we cannot rule out that it will not be falsified the next time it is tested. As you can see, the scientific method is an inherently critical method when it is "successful" (when a theory's predictions seem to be borne out), because it is precisely under these circumstances that it is most cautious in the claims that it makes.

Although we cannot ever prove our theories, we can claim that some theories are better corroborated than others. As a result, we can have more confidence in their conclusions. One might think that a theory that has been subjected to multiple tests is better corroborated than one that has not been subjected to many tests at all. However, this is not always the case. If we keep testing the same implication over and over again, it is not clear how much an additional test actually adds to the degree to which the theory is corroborated. What really matters is not so much how many times a theory has been corroborated, but the severity and variety of the tests to which it has been subjected. This, in turn, will depend on the degree to which the theory is falsifiable. Again, this is why we like our models to be simple and have multiple implications. In general, we will have more confidence in a theory that has survived a few harsh tests than a theory that has survived many easy ones. This is why scientists often talk about the world as if it were black-and-white rather than gray. Bold statements should be interpreted not as scientific hubris but rather as attempts to invite criticism—they are easier to falsify.

What happens if we do not observe the implications deduced from our theory? Can we conclude that our theory is incorrect based on one observation? The answer is "probably not." It is entirely possible that we have not observed and measured the world without error. Moreover, if we believe that human behavior is inherently probabilistic, then we might not want to reject theories on the basis of a single observation. In a world in which our tests are potentially fallible, we should not relegate a theory to the dustbin of intellectual history the minute one of its implications is shown to be false. Instead, we must weigh the number, severity, and quality of the tests that the theory's implications are subjected to and make a judgment. And most important, this judgment should be made with an eye toward what would replace the theory should we decide to discard it. This is why some scientists say that it takes a theory to kill a theory. Further, if we do embrace a new theory and disregard an alternative, it should be because the new theory is more consistent with all of the implications of both theories. Developing a new theory that explains the facts that the old theory found

3. Many scientists, however, slip into the language of verification when reporting their results. Instead of simply saying that their test has failed to falsify their hypotheses or is consistent with their theory, they will claim that the test has shown that their theory is correct. For example, they might claim that their test shows that wealth causes democracies to live longer when, in fact, all they can conclude is that they were unable to falsify the claim that wealth causes democracies to live longer.

inconvenient without also explaining the many facts that the old theory accurately predicted is called *ad hoc* explanation. Because this practice does not expose the new theory to falsification as strenuously as it does the old theory, it is not consistent with sound scientific practice.

AN INTRODUCTION TO LOGIC

In the previous section, we talked in a rather casual way about constructing and testing scientific explanations. In order to better appreciate the important connection between theory construction and theory testing, it is useful to devote some time to the study of logic. The study of logic is, first and foremost, about learning to be careful about how we construct and evaluate arguments.

Throughout our lives, we are confronted by people trying to convince us of certain things through arguments. Politicians make arguments as to why we should vote for their party rather than the party of their opponents. National leaders provide arguments for why certain policies should be implemented or abandoned. Lawyers make arguments as to why certain individuals should be found guilty or innocent. Professors make arguments as to why students should spend more time in the library and in class rather than at parties. It is important for you to know when these arguments are logically valid and when they are not. If you cannot distinguish between a valid and an invalid argument, other people will be able to manipulate and exploit you. You will be one of life's suckers. In this section, we give you some tools to determine whether an argument is valid or not.

Valid and Invalid Arguments

What is an argument? An **argument** is a set of logically connected statements, typically in the form of a set of **premises** and a **conclusion**. An argument is **valid** when accepting its premises compels us to accept its conclusions. An argument is **invalid** if, when we accept the premises of an argument, we are free to accept or reject its conclusions. One way to represent an argument is in the form of a **categorical syllogism** that consists of a major premise, a minor premise, and a conclusion. The major premise is typically presented as a conditional statement, such as "If *P*, then *Q*." The "if" part of the conditional statement (in this case "If *P*") is called the *antecedent,* whereas the "then" part of it (in this case "then *Q*") is called the *consequent.* An example of a conditional statement is "If a country is wealthy [antecedent], then it will be a democracy [consequent]." The minor premise consists of a claim about either the antecedent or the consequent in the conditional statement (major premise). The conclusion is a claim that is thought to be supported by the premises.

An **argument** is a set of logically connected statements, typically in the form of a set of premises and a conclusion. A **premise** is a statement that is presumed to be true within the context of an argument leading to a conclusion. A **conclusion** in an argument is a claim that is thought to be supported by the premises. A **valid argument** is one in which, if you accept the premises, you are compelled to accept the conclusion. An **invalid argument** is one in which, if you accept the premises, you are free to accept or reject the conclusion. A **categorical syllogism** is a specific type of argument that consists of a major premise, a minor premise, and a conclusion.

TABLE 2.2 Affirming the Antecedent: A Valid Argument

	General form	Specific example
Major premise	If *P*, then *Q*	If a country is wealthy, then it will be a democracy.
Minor premise	*P*	The country is wealthy.
Conclusion	Therefore, *Q*.	Therefore, the country will be a democracy.

Four types of conditional argument can be represented with a syllogism—arguments that affirm or deny the antecedent and those that affirm or deny the consequent. Which of these four types of argument are valid, and which are invalid? Recall that a valid argument is one such that if you accept that the premises are true, then you are compelled to accept the conclusion as true. Let's start by considering what happens when we affirm the antecedent. An example is shown in Table 2.2.

The major premise states, "If *P* is true, then *Q* must be true." The minor premise says that "*P* is true." Together, these premises compel us to accept that the conclusion is true. As a result, the argument is valid. In other words, the major premise states, "If a country is wealthy [antecedent], then it will be a democracy [consequent]." The minor premise says, "The observed country is wealthy." It logically follows from this that the observed country must be a democracy. To see why this type of argument is valid, consider the general form of this argument in set-theoretic form. This is shown in Figure 2.1. The major premise indicates that the set of cases where *P* occurs is a subset of the cases where *Q* occurs. The minor premise maintains that *P* does occur. Figure 2.1 clearly shows that if the case in question is in *P*, as the minor premise affirms, then the case must also be in *Q*. Thus, the argument is valid—we are compelled to conclude *Q*.

FIGURE 2.1 Major Premise: If *P*, Then *Q*

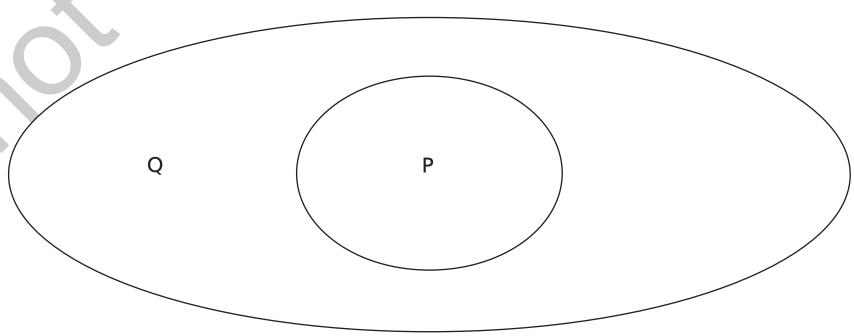


TABLE 2.3 Denying the Antecedent: An Invalid Argument

	General form	Specific example
Major premise	If *P*, then *Q*	If a country is wealthy, then it will be a democracy.
Minor premise	Not *P*	The country is not wealthy.
Conclusion	Therefore, not *Q*.	Therefore, the country will not be a democracy.

Now let's consider what happens when we deny the antecedent. An example is shown in Table 2.3. Once again, the major premise can be represented in set-theoretic terms by Figure 2.1. The difference from the previous example is that the minor premise now asserts that *P* is not the case; that is, it denies the antecedent. If we accept this, does it necessarily follow that *Q* is not the case, as the conclusion maintains? Figure 2.1 clearly illustrates that even if our case is not in *P*, it could still be in *Q*. As a result, it does not logically follow from observing "not *P*" that *Q* is not the case. Therefore, this is an invalid argument. This is because we can contradict the conclusion (not *Q*) without running into a contradiction with either the major premise or the minor premise. Since a valid argument compels us to accept its conclusion given that its premises are true, this is sufficient to demonstrate that arguments that deny the antecedent are invalid.

In the context of our running example, does it follow from the fact that the observed country is not wealthy that it will not be a democracy? Intuitively, we can imagine that there may be other reasons why a country is a democracy even though it is not wealthy. Indeed, one example of a nonwealthy democracy is India. An important point here, though, is that the argument is invalid, not because we can come up with an example of a real democracy that is not wealthy (India), but rather because we are not compelled to accept the conclusion based on the truthfulness of the major and minor premises. It may be confusing for readers that there is no direct connection between the factual accuracy of an argument's conclusion and the validity of the argument itself—a valid argument can have a conclusion that is factually false, and an invalid argument can have a conclusion that is factually true. If we restrict our attention only to whether the argument is valid as it applies to our democracy example, we must ask, "Does the major premise claim that wealth is the only reason why a country will be a democracy?" The answer is clearly no. The major premise states only what will happen if a country is wealthy. It makes no claim as to what might happen if a country is not wealthy. It is for this reason, and this reason alone, that the argument is invalid.

Now let's consider what happens when we affirm the consequent. An example is shown in Table 2.4. As before, the major premise can be represented in set-theoretic terms by Figure 2.1. The difference this time is that the minor premise now asserts that *Q* is the case; that is, it affirms the consequent. If we accept that the premises are true, are we compelled to accept the conclusion that *P* is the case? Figure 2.1 clearly illustrates that the fact that our

Table 2.4 Affirming the Consequent: An Invalid Argument I

	General form	Specific example
Major premise	If *P*, then *Q*	If a country is wealthy, then it will be a democracy.
Minor premise	*Q*	The country is a democracy.
Conclusion	Therefore, *P*	Therefore, the country is wealthy.

case is in *Q* does not necessarily mean that it is also in *P*. As a result, the argument is invalid—we are not compelled to accept the conclusion based on the premises.

In the context of our running example, an argument that affirms the consequent confuses necessity and sufficiency. Although the major premise states that wealth is sufficient for democracy—wealthy countries will be democracies—it does not assert that wealth is necessary for democracy. In other words, the major premise does not state that wealth is the only cause of a country's democracy. Consequently, we cannot make a valid inference from the fact that a country is a democracy to the claim that the country must be wealthy—it may be wealthy, or it may not be. Recall that to show that an argument is invalid, it is not necessary to show that its conclusion is false; we have to show only that it doesn't have to be true.

Finally, let's consider what happens when we deny the consequent. An example is shown in Table 2.5. As always, the major premise can be represented in set-theoretic terms by Figure 2.1. The difference this time is that the minor premise now denies that *Q* is the case; that is, it denies the consequent. If we accept that the premises are true, are we compelled to accept the conclusion that "not *P*" is the case? Figure 2.1 clearly shows that the fact that our case is not in *Q* necessarily means that it is not in *P*. As a result, the argument is valid—we are compelled to accept the conclusion based on the premises. In the context of our running example, the major premise indicates that all wealthy countries are democracies and the minor premise states that the country is not a democratic one. If these premises are both true, then it logically follows that our country cannot be wealthy.

Our brief foray into the study of logic indicates that if complex arguments can be broken down into categorical syllogisms, then it is possible to classify all arguments into one of four types according to whether they affirm or deny the consequent or antecedent. Two of these

Table 2.5 Denying the Consequent: A Valid Argument I

	General form	Specific example
Major premise	If *P*, then *Q*	If a country is wealthy, then it will be a democracy.
Minor premise	Not *Q*	The country is not a democracy.
Conclusion	Therefore, not *P*	Therefore, the country is not wealthy.

TABLE 2.6 What Types of Conditional Arguments Are Valid?

	Antecedent	Consequent
Affirm	Valid	Invalid
Deny	Invalid	Valid

arguments are valid, but the other two are invalid. Specifically, affirming the antecedent and denying the consequent are valid arguments—if you accept the major and minor premises, you are compelled to accept the conclusion. In contrast, denying the antecedent and affirming the consequent are invalid arguments—if you accept the major and minor premises, you are not compelled to accept the conclusion. These results are summarized in Table 2.6.

Testing Theories

We obviously think that it is important for you to be able to distinguish between valid and invalid arguments so that you are not manipulated or exploited by others. However, this brief introduction to logic is also important because it tells us something about the way that scientists test their theories and explanations. Suppose we want to explain why rich countries are much more likely to be democracies than poor countries. One possible explanation for why this might be the case is given in the following statements:[4]

1. Living in a dictatorship is risky—if you are one of the dictator's friends, you will do extremely well; but if you are not, you will do extremely poorly.
2. Living in a democracy is less risky—democratic leaders have to spread the goodies (and the pain) around more evenly. This means that you are less likely to do extremely well or extremely poorly in a democracy.
3. Rich people are less likely to take risks than poor people because they have more to lose. This means that countries with lots of rich people are more likely to be democracies than dictatorships.

This short explanation provides reasons why rich countries might be more likely to be democracies than poor countries. How good is this explanation, though? Does this argument have any testable implications? One implication is that rich democracies should live longer than poor democracies. This is because people in rich democracies should be less likely to take the "risk" of becoming a dictatorship; in contrast, people in poor democracies might wonder what they have to lose.

How can we use observations of the real world to evaluate our proposed explanation? It is often the case that the implications of an explanation are more readily observable than the

4. This is a simplified version of an argument presented by Przeworski (2001). It will be discussed more fully in Chapter 6.

TABLE 2.7 Affirming the Consequent: An Invalid Argument II

General form	Example	Specific example
If *P*, then *Q*	If our theory *T* is correct, then we should observe some implication *I*.	If our theory is correct, then we should observe that rich democracies live longer than poor democracies.
Q	We observe implication *I*.	Rich democracies live longer than poor democracies.
Therefore, *P*	Therefore, our theory *T* is correct.	Therefore, our theory is correct.

elements of the explanation itself. Consider the example we are using. Although it may be possible to compare the distribution of good and bad outcomes in dictatorships and democracies, the claims that people differ in their propensity to take risks and that this propensity is related to their level of income are difficult to observe. This is because the propensity to take risks is an internal and psychological attribute of individuals. For similar reasons, scholars typically evaluate their explanations by observing the real world to see if the implications of their explanations appear to be true based on the assumption, "If my theory is true, then its implications will be true." If we take this to be our major premise and the truth or falsity of the theory's implications as the minor premise, then we might be able to use observations to draw inferences about our theory or explanation.

Suppose our theory's implications were borne out by our observation that rich democracies live longer than poor democracies. Can we conclude that our theory is true? Note that if we were to do so, we would be engaging in reasoning that affirmed the consequent. This fact is shown more clearly in Table 2.7. As you know by now, affirming the consequent is an invalid form of argument. The major premise says only that if the theory is correct, then the implications should be observed. It never says that the only way for these implications to be produced is if the theory is correct. In other words, processes other than those described in our theory may produce the observation that rich countries live longer than poor countries. Put differently, the mere fact of observing the predicted implication does not allow us to categorically accept or reject our theory.

Suppose now that our observations did not bear out our theory's implications; that is, we did not observe that rich democracies live longer than poor democracies. Can we conclude that our theory is incorrect? Note that if we were to do so, we would be engaging in reasoning that denies the consequent. This fact is shown more clearly in Table 2.8. As you know by now, denying the consequent is a valid form of argument. In other words, by accepting the premises, we are compelled to accept the conclusion that our theory is not correct.

If we compare the two previous examples, we can see an important asymmetry as regards the logical claims that can be made on the basis of "confirming" and "disconfirming" observations. When an implication of our theory is confirmed, the most we can say is that the

TABLE 2.8 Denying the Consequent: A Valid Argument II

General form	Example	Specific example
If *P*, then *Q*	If our theory *T* is correct, then we should observe some implication *I*.	If our theory is correct, then we should observe that rich democracies live longer than poor democracies.
Not *Q*	We do not observe implication *I*.	Rich democracies do not live longer than poor democracies.
Therefore, not *P*	Therefore, our theory *T* is incorrect.	Therefore, our theory is incorrect.

theory may be correct. This is because neither of the two possible conclusions—our theory is correct or our theory is not correct—contradicts our major and minor premises. In other words, we cannot say that our theory is correct or verified. In contrast, if we find that an implication of our theory is inconsistent with observation, then we are compelled by logic to accept that the theory is false—this is the only conclusion that is consistent with our observation. Thus, although we can know that a theory must be incorrect in light of a disconfirming case, all that we can say in light of a confirming case is that a theory may be correct (it may also be wrong). What does this mean? It means that we are logically justified in having more confidence when we reject a theory than when we do not. This, in turn, implies that the knowledge encapsulated in theories that have not been rejected remains tentative and can never be proven for sure—scientific theories can never be proven. Even if we are utterly convinced that our major and minor premises are true, all that we can logically conclude from a confirming instance is that the theory has not yet been falsified.

This asymmetry between confirming and disconfirming cases led the philosopher of science Sir Karl Popper ([1959] 2003, 280–81) to conclude:

> The old scientific ideal of *episteme*—of absolutely certain, demonstrable knowledge—has proved to be an idol. The demand for scientific objectivity makes it inevitable that every scientific statement must remain *tentative for ever.* . . . With the idol of certainty . . . there falls one of the defenses of obscurantism which bar the way to scientific advance. For the worship of this idol hampers not only the boldness of our questions, but also the rigor and integrity of our tests. The wrong view of science betrays itself in the craving to be right; for it is not his *possession* of knowledge, of irrefutable truth, that makes the man of science, but his persistent and recklessly critical *quest* for truth.

If confirming observations do not prove that our theory is correct, does this mean that they are of no use whatsoever? The answer is no. Imagine that we start with a set of implications derived from a theory and then observe some facts. In other words, let's start with the theory and then observe the world. If we do this, then it is possible that our observations will contradict our theory. If it turns out that our observations are consistent with our theory, then we can have a greater measure of confidence in our theory because it withstood the very real chance of being falsified. We cannot say that our theory is verified or confirmed, just that we have more confidence in it. If our observations are inconsistent with our theory, then we can draw valid inferences about the truthfulness of our theory—we can conclude that it is wrong. This approach to doing science, which forms the basis of the scientific method described earlier, is called falsificationism. **Falsificationism** is an approach to science in which scientists generate or "deduce" testable hypotheses from theories designed to explain phenomena of interest. It emphasizes that scientific theories are constantly called into question and that their merit lies only in how well they stand up to rigorous testing. Falsificationism forms the basis for the view of science employed in this book.

> **Falsificationism** is an approach to science in which scientists generate testable hypotheses from theories designed to explain phenomena of interest. It emphasizes that scientific theories are constantly called into question and that their merit lies only in how well they stand up to rigorous testing.

The approach to science that we have described here takes a clear stance in the debate between "deductive" and "inductive" approaches to learning. The **deductive approach to learning** formulates an expectation about what we ought to observe in light of a particular theory about the world and then sets out to see if our observations are consistent with that theory. The **inductive approach to learning**, on the other hand, starts with a set of observations and then tries to ascertain a pattern in the observations that can be used to generate an explanation for the observations. Induction is problematic because in order to be successful it must rest at some point on the fallacy of affirming the consequent—the fact that observation precedes theory construction means that the theory is never exposed to potential falsification! Popper ([1959] 2003) suggests that, in fact, the biggest problem with induction is not so much that it is wrong but that it is impossible. Observational facts do not just present themselves to observers—we always decide which facts to pay attention to and which to ignore. As we noted earlier, the hunch that tells us what to observe and what to ignore, that is, what constitutes a puzzle worth explaining, constitutes a theory. In this respect, scholars who claim to be engaged in an inductive inquiry are actually engaged in an implicit deductive endeavor. If it is true that we are "all deductivists" as Popper claims, then the argument for deduction amounts to a claim that it is better to use theory explicitly than to use it implicitly.

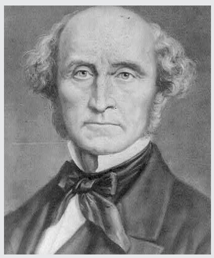
> The **deductive approach to learning** involves formulating an expectation about what we ought to observe in light of a particular theory about the world and then sets out to see if observation is consistent with that theory. With deduction, theory precedes observation. The **inductive approach to learning** starts with a set of observations and then tries to ascertain a pattern in the observations that can be used to generate an explanation for the observations. With induction, observation precedes theory.

Having described the scientific method, we would like to briefly dispel certain myths that have developed about science. Some of these myths have been promoted by opponents of the scientific project, but others, unfortunately, have been sustained by scientists themselves.

Box 2.2

THE COMPARATIVE METHOD

An Overview and Critique



An undated portrait of John Stuart Mill.

Hader, Ernst, Artist, Williams, Sophus, photographer. J. Stuart Mill/E. Hader, pinxit; phot. u. verl. v. Sophus Williams, Berlin W. Berlin: Sophus Williams, 1884. Image. Retrieved from the Library of Congress, https://www.loc.gov/item/2004678579/. (Accessed October 26, 2016.)

The **comparative method**, also known as Mill's methods, involves the systematic search for the necessary and sufficient causes of political phenomena. The comparative method comprises the Method of Agreement and the Method of Difference. The **Method of Agreement** compares cases that "agree" in regard to the phenomenon to be explained.

You will, undoubtedly, encounter excellent work by scholars who claim to be proceeding inductively. The most common method of inductive research in comparative politics is known as the **comparative method**. It is also known as Mill's methods because it is based on a formal set of rules outlined by John Stuart Mill in his 1872 book, *A System of Logic.* Mill actually outlined two different methods. One is called the Method of Agreement, and the other is called the Method of Difference. Political scientists who employ these methods collect observations of the world and then use these observations to develop general laws and theories about why certain political phenomena occur.[5] In employing these methods, the goal is to identify the causes of political events.

Mill's **Method of Agreement** compares cases that "agree" in regard to the political phenomenon to be explained. To see how this works, suppose that we want to explain the occurrence of democracy. Common sense might suggest that if we want to know what causes democracy, we should study democracies.[6] We could observe two or more contemporary democracies and take note of their features. For example, we might compare the United Kingdom, Belgium, and the United States, as we do in Table 2.9. All three countries "agree" in regard to the outcome to be explained—they are all democracies.

Table 2.9 Mill's Method of Agreement

Country	Democracy	Wealth	Ethnically homogeneous	Parliamentary system
UK	Yes	Yes	Yes	Yes
Belgium	Yes	Yes	No	Yes
US	Yes	Yes	Yes	No

5. For example, Weber ([1930] 1992) employs Mill's methods to explain the rise of capitalism; Moore ([1966] 1999) to determine why some countries are democracies but others are dictatorships; Skocpol (1979) to examine social revolutions; Katznelson (1985) to analyze the variation in the organizational patterns of the working class in the United States and the United Kingdom; and Kalyvas (1996) to explain the rise of Christian democracy in western Europe.

6. As we will see, the kind of sense needed to do good science often turns out to be very "uncommon."

What, if anything, can we infer from such a comparison? Well, we observe that the United Kingdom is a wealthy, relatively homogeneous parliamentary democracy. Belgium is a wealthy, heterogeneous parliamentary democracy.[7] And the United States is a wealthy, relatively homogeneous presidential democracy. Assuming that the classification of our observations is correct, we can conclude that ethnic homogeneity is not a **necessary condition** for democracy. This is because Belgium is a democracy despite being ethnically diverse. We can also conclude that having a parliamentary system is not a necessary condition for democracy. This is because the United States is a democracy despite having a presidential system. Wealth alone survives as a potential necessary condition for democracy in our three observations—all three democracies are wealthy. Based on the evidence in this simple example, then, a scholar using Mill's Method of Agreement would conclude that democracy is caused by wealth, or economic development.

A **necessary condition** is a circumstance in whose absence the phenomenon in question cannot occur. A **sufficient condition** is a circumstance in whose presence the phenomenon in question must occur.

Note that Mill's Method of Agreement does not allow us to determine whether wealth is a **sufficient condition** for democracy. To determine this, you would need to look for wealthy countries that are not democracies. If you found such a country, you would know that wealth is not sufficient for democracy. Thus, to evaluate whether wealth is sufficient for democracy, we need to examine nondemocracies as well as democracies. This obviously cannot be done with Mill's Method of Agreement because the outcome to be explained would not "agree" for all of the cases. It turns out, though, that we can evaluate claims about sufficient (and necessary) causes using Mill's Method of Difference.

Mill's **Method of Difference** compares cases that "differ" in regard to the outcome to be explained. To evaluate whether wealth is a sufficient condition for democracy, we must go back out into the real world to observe some nondemocracies.[8] Imagine that the first nondemocracy that we observe is Mexico prior to 1990. As Table 2.10 indicates, Mexico was a relatively wealthy, ethnically homogeneous presidential country in this period. The case of Mexico prior to 1990 tells us that wealth is not a sufficient condition for democracy. This is because Mexico is wealthy but not a democracy. Not only does Mill's Method of Difference allow us to determine whether certain features are sufficient to produce democracy but it also allows us to find out if those features are necessary for democracy. In this sense, it is "stronger" than the Method of Agreement.

The **Method of Difference** compares cases that "disagree" in regard to the outcome to be explained.

From the set of four observations in Table 2.10, we can make the following conclusions:

- Wealth is not sufficient for democracy in light of the Mexican case. It may, however, be a necessary condition.

---

7. Belgium's population is fairly evenly split between Dutch-speaking Flemish and French-speaking Walloons. There is also a sizable German-speaking population in the east of the country and a nontrivial number of non-European immigrants.

8. Here's an example of science depending on uncommon sense. Note the somewhat surprising implication that if you want to know what is sufficient to produce democracy, you must study nondemocracies.

Table 2.10 Mill's Method of Difference

Country	Democracy	Wealth	Ethnically homogeneous	Parliamentary system
UK	Yes	Yes	Yes	Yes
Belgium	Yes	Yes	No	Yes
US	Yes	Yes	Yes	No
Mexico	No	Yes	Yes	No

- Ethnic homogeneity is neither necessary for democracy in light of the Belgian case nor sufficient for democracy in light of the Mexican case.
- A parliamentary system is not necessary for democracy in light of the United States case. It may, however, be a sufficient condition based on the Belgium and United Kingdom cases.

Mill's methods are widely employed in comparative political science, where they form the basis of the popular "most similar systems" and "most different systems" research designs (Collier 1993; Lijphart 1971, 1975; Przeworski and Teune 1970).[9]

It is easy to see why the comparative method is so appealing—it claims to be able to identify the necessary and sufficient conditions for political phenomena. The problem, as many scholars have noted, though, is that certain fairly restrictive assumptions must be met before analysts can draw valid inferences from Mill's methods (Lieberson 1991, 1994; Sekhon 2004). For example, one must assume that there is only one cause for a political phenomenon like democracy and that this cause is deterministic; that is, it *always* produces the political phenomenon (democracy). One must also assume that all potential causes have been identified and that all causal factors work independently of each other. These assumptions are particularly problematic given that the comparative method does not provide us with any help in determining when they will be met. In our view, at least one of these assumptions is likely to be violated in almost any social scientific application. Mill (1872) himself recognized this and warned scholars against using his methods to explain the political world. As he put it,

> Nothing can be more ludicrous than the sort of parodies on experimental reasoning which one is accustomed to meet with, not in popular discussion only, but in grave treatises, when the affairs of nations are the theme. "How," it is asked, "can an institution be bad, when the country has prospered under it?" "How can such or such causes have contributed to the prosperity of one country, when another has prospered without them?" Whoever

9. Somewhat confusingly, the most similar systems design is equivalent to Mill's Method of Difference. It requires that the analyst find cases that are identical to each other except in regard to the outcome to be explained and one key condition. The most different systems design is equivalent to Mill's Method of Agreement. It requires the analyst to choose cases that are as different as possible except in regard to the outcome to be explained and one key condition.

> makes use of an argument of this kind, not intending to deceive, should be sent back to learn the elements of some of the more easy physical sciences. (p. 324)

These reservations are sufficiently worrisome on their own that analysts should be reluctant to accept uncritically claims based on the application of Mill's methods. A more fundamental problem is at issue here, however. Even if the analyst could be satisfied that the assumptions underpinning the comparative method were met, she would have established only that certain phenomena occur together; she would not have provided an explanation of the outcome in question. That is, Mill's methods are empirical methods—they tell us what happens, not why the phenomena occur together. Put differently, all they say is that *Y* happened when *X* was present; this is roughly equivalent to saying that the sun came up because the rooster crowed. An essential missing ingredient is a sense of process, a story about why *Y* appears to happen when *X* happens. The story about the process that produces the outcomes we see is what scientists call a *theory,* and these stories cannot necessarily be reduced to a set of circumstances that covary with the outcome we wish to explain.

Finally, we should note that the asymmetry between confirmation and falsification that we noted previously has important implications for the methods we use to build knowledge. When scholars use the comparative method, they go out into the real world to collect observations and look for patterns in the data. Those factors that cannot be eliminated as potential causes by Mill's methods become our explanation. Each new case that exhibits the same pattern in the data confirms or verifies our conclusion. Note that because the comparative method starts with observations, it relies entirely on the process of affirming the consequent. If we identify causes only after we have observed the data, as the comparative method requires, we have no chance of ever coming across disconfirming observations. This is because our "theory" is essentially just a restatement of the patterns in our observations.[10] This is a real problem, whether the researcher is employing the comparative method on a small number of cases or analyzing large data sets looking for patterns. No matter how many cases these researchers observe that appear to exhibit the predicted pattern, they are never logically justified in claiming that their conclusions have been confirmed or verified.

You might wonder whether there is any way to avoid these problems. The answer is yes. Imagine that we start with a set of implications derived from a theory and then observe some facts. In other words, let's start with the theory and then observe the world rather than the other way around. It is now at least possible for our observations to contradict our theory. If it turns out that our observations are consistent with our theory, then we can have a greater measure of confidence in our theory because it withstood the very real chance of being falsified. If our observations are inconsistent with our theory, then we can draw valid inferences about the truthfulness of our theory—we can conclude that it is wrong. This approach to doing science, as we have seen, is called falsificationism, and it forms the basis for the view of science employed in this book.

---

10. This suggests that the comparative method is, at most, suitable only for developing theories and not for testing them.

MYTHS ABOUT SCIENCE

The first myth is that science proves things and leads to certain and verifiable truth. This is not the best way to think about science. It should be clear by now from our discussion that the best science can hope to offer are tentative statements about what seems reasonable in light of the best available logic and evidence. It may be frustrating for students to realize this, but science can speak with more confidence about what we do not know than what we do know. In this sense, the process of scientific accumulation can be thought of as the evolution of our ignorance. We use the scientific method because it is the best tool available to interrogate our beliefs about the (political) world. If we hold on to any beliefs about the (political) world, it is because, after we have subjected them to the most stringent tests we can come up with, they remain the most plausible explanations for the phenomena that concern us.

The second myth is that science can be done only when experimental manipulation is possible. This is clearly false. For theories to be scientific, they need only be falsifiable. There is no claim that the tests of these theories need to be carried out in an experimental setting. Many of the natural sciences engage in research that is not susceptible to manipulation. For example, all research on extinct animals, such as dinosaurs, must be conducted without the aid of experimental manipulation because the subjects are long dead. In fact, there is also no claim that a theory must be tested before it can be called scientific. Einstein presented a special theory of relativity in 1905 that stated, among other things, that space had to be curved, or warped. It took fourteen years before his theory was tested with the help of a solar eclipse. No scientist would claim that Einstein's theory was unscientific until it was tested. Put simply, scientific theories must be potentially testable, but this does not mean that they stop being scientific if they are yet to be actually tested.

The third myth is that scientists are value neutral. It is necessary here to distinguish between the method of science and the individuals—the scientists—who engage in science. The scientific method itself is value neutral. As we have indicated in this chapter, science is simply a method that involves generating and evaluating logically consistent sets of falsifiable statements about the world. Scientists, though, may not be value neutral (Longino 1987; Haraway 1988). It is important to remember that the pursuit of knowledge about the world is closely entangled with attempts by people to change the world. As a result, the types of research questions that are asked and the interpretation of scientific results are likely to be infused with the specific values and biases held by individual scientists and those who use their research. The lack of diversity in most scientific disciplines, whether in terms of gender, race, income, class, sexuality, religion, ethnicity, and so on, along with the power structure that exists in many societies, means that some research areas are less studied than others and that certain viewpoints are excluded or less privileged than others when it comes to interpreting scientific evidence (Smith 1974; Collins 1986, 1989; Carroll and Zerilli 1993). In effect, the knowledge that is produced by science is socially constructed. This is one of the many reasons for trying to promote the diversity of those involved in the scientific endeavor. The fact that scientists may not be value neutral means that we should be very clear about the limits of our knowledge and not encourage others to act upon knowledge that is not highly corroborated. Moreover, we should try to conduct our studies in such a way that someone who does not share our biases can determine if our arguments and evidence are reasonable.

It has been argued that science is predicated on two rules (Rauch 1993). First, no one gets the final say on any issue—all knowledge claims are, for the reasons outlined in this chapter, open to criticism. Second, no individual has a personal claim of authority about whether scientific statements are true or not. Taken together, these two rules create a social system that makes it possible that even though individual scientists will have biased perspectives, others, who hold different biases, will have incentives to check their work. As a community, scientists with many different biases will use the scientific method to check the claims that are being made in an attempt to reach a consensus that is independent of the biases held by individual scholars. This is yet another reason why having a diverse group of scientists is valuable.

The fourth myth, that politics cannot be studied in a scientific manner, can easily be dispelled by now. Our description of the scientific method clearly shows that this myth is false. The study of politics generates falsifiable hypotheses and hence generates scientific statements. These theories of politics can be tested just like any other scientific theory. We will further demonstrate that politics can be studied in a scientific manner in the remaining chapters of this book. The fact, though, that our subjects can read our work and change their behavior makes our job quite a bit harder than if we were working in one of the natural sciences.

CONCLUSION

In this chapter we have argued that it is useful to think about politics in a scientific manner. We have also tried to offer a clear view of what most practicing scientists have in mind when they use the word *science*. It is a fairly minimalist view. What unites all scientists is the idea that one ought to present one's ideas in a way that invites refutation (Popper 1962). It is incumbent upon the scientist to answer the question "What ought I to observe if what I claim to be true about the world is false?" This view of science recognizes that scientific knowledge is tentative and should be objective. Although it is certainly likely that our prejudices and biases motivate our work and will creep into our conclusions, the goal of science is to present our conclusions in a way that will make it easy for others to determine whether it is reasonable for people who do not share those prejudices and biases to view our conclusions as reasonable.

Key Concepts

argument *27*
categorical syllogism *27*
comparative method *35*
conclusion (in an argument) *27*
critical test *22*
deductive approach to learning *34*
falsifiable *18*
falsificationism *34*
inductive approach to learning *34*
invalid argument *27*
Method of Agreement *35*
Method of Difference *36*
necessary condition *36*
premise *27*
scientific method *19*
sufficient condition *36*
tautology *18*
theory *19*
uniformity of nature *20*
valid argument *27*

Problems

This section includes various questions designed to evaluate your comfort with some of the more important concepts, issues, and methods introduced in this chapter.

Logic: Valid and Invalid Arguments

1. Consider the following argument.

Major Premise: If a country has a strong economy, the government will be popular.

Minor Premise: The government is not popular.

Conclusion: Therefore, the country does not have a strong economy.

a. What form of categorical syllogism is this (affirming the antecedent/consequent or denying the antecedent/consequent)?

b. Is this a valid or an invalid argument?

2. Consider the following argument.

Major Premise: If the president commits a criminal act, then he can be impeached.

Minor Premise: The president does not commit a criminal act.

Conclusion: Therefore, the president cannot be impeached.

a. What form of categorical syllogism is this?

b. Is this a valid or an invalid argument?

3. Consider the following argument.

Major Premise: If a country employs proportional representation electoral rules, it will have many parties.

Minor Premise: The country does employ proportional representation electoral rules.

Conclusion: Therefore, the country will have many parties.

a. What form of categorical syllogism is this?

b. Is this a valid or an invalid argument?

4. Consider the following argument.

Major Premise: If a county has a participant culture, then democracy in that country will be stable.

Minor Premise: Democracy in country *X* is unstable.

Conclusion: Therefore, country *X* does not have a participant culture.

a. What form of categorical syllogism is this?

b. Is this a valid or an invalid argument?

5. Consider the following argument.

Major Premise: If Islam is incompatible with democracy, then Muslim majority countries are more likely to be dictatorships than democracies.

Minor Premise: Muslim majority countries are more likely to be dictatorships than democracies.

Conclusion: Therefore, Islam is incompatible with democracy.

a. What form of categorical syllogism is this?

b. Is this a valid or an invalid argument?

6. Consider the following argument.

Major Premise: If I work hard in this class, then I will get a good grade.

Minor Premise: I did not work hard in this class.

Conclusion: Therefore, I will not get a good grade.

a. What form of categorical syllogism is this?

b. Is this a valid or an invalid argument?

7. Consider the following argument.

Major Premise: If theory *T* is correct, all rich countries will be democracies.

Minor Premise: All rich countries are democracies.

Conclusion: Therefore, theory *T* is correct.

a. What form of categorical syllogism is this?

b. Is this a valid or an invalid argument?

c. If you wanted to demonstrate that theory *T* is wrong, what would you have to observe?

8. Come up with an example of your own categorical syllogism. Demonstrate why the argument is either valid or invalid.

Scientific Statements

9. A statement is scientific if it is falsifiable. Which of the following statements are scientific and why?

- Smoking increases the probability of getting cancer.
- A square is a two-dimensional figure with four equal straight sides and four right angles.
- The sun revolves around the earth.
- It always rains in England during the winter.
- Education spending increases under left-wing governments.
- Religious faith assures a person a place in the afterlife.
- Democracies are less likely to go to war than dictatorships.

- The unexamined life is not worth living.
- Voter turnout is higher among citizens living in rural areas than for citizens in urban areas.

10. Some statements are nonscientific because they are tautologies and some because they refer to inherently unobservable phenomena. Come up with an example of both types of nonscientific statement.

11. Sometimes it is hard to know whether a statement is scientific or not. Much depends on how we define certain terms. Consider the following statement.

- All good students get high grades.

Whether this statement is falsifiable depends on how we define "good students." On the one hand, if we define good students as those who get high grades, then this statement becomes tautological or true by definition—no observation could falsify it. This is easy to see if we swap in our definition of good students in the statement above. If we did this, we would have "All students with high grades get high grades." It should be obvious that this statement could never be falsified because it is impossible to ever find a student with high grades who does not have high grades! With this particular definition of good students, the statement above is not scientific. On the other hand, if we define good students as those who work hard, then the statement above is scientific. This is easy to see if we swap in our new definition of good students. If we did this, we would have "All students who work hard get good grades." It should be obvious that this statement could be falsified. It would be falsified if we observed a student who worked hard but received a low grade.

Consider the following statement.

- All mainstream US senators agree that the House bill is unacceptable.

    a. Is this statement scientific if "mainstream US senators" are defined as those who find the House bill unacceptable?

    b. Is this statement scientific if "mainstream US senators" are defined as those who share a middle-of-the-road ideology?

Now consider the following statement.

- If the Affordable Care Act (sometimes referred to as "Obamacare") is successfully implemented, then health care outcomes in the United States will improve.

    a. Is this statement scientific if the "successful implementation of the Affordable Care Act" is defined in terms of improved health care outcomes?

    b. Is this statement scientific if the "successful implementation of the Affordable Care Act" is defined in terms of the Affordable Care Act actually being passed in Congress?

Necessary and Sufficient Conditions

12. Consider the following statements. After looking at the structure of each statement, would you say that the conditions shown in boldface type are necessary or sufficient to produce the effects shown?

- If a person contracts measles, then she was **exposed to the measles virus**.
- If a **democracy is rich**, then it will stay a democracy.
- If **a democracy has a participant culture**, then it will stay a democracy.
- A country cannot maintain democracy unless **it has a participant culture**.
- Countries have many parties only when they employ proportional electoral rules.
- Countries always have few parties when they employ majoritarian electoral rules.
- Students will receive a good grade only if they work hard.

Model Building in the Scientific Method

13. It has frequently been observed that students coming into a lecture hall tend to fill up the rear of the hall first (Lave and March 1975; Schelling 1978). Here are two possible explanations, or models, that predict this kind of behavior.

    Minimum Effort Theory: People try to minimize effort; having entered at the rear of the hall, they sit there rather than walk to the front.

    "Coolness" Theory: General student norms say that it is not cool to be deeply involved in schoolwork. Sitting in front would display interest in the class, whereas sitting in the rear displays detachment.

    a. Make up two facts (that is, derive two specific predictions) that, if they were true, would tend to support the Minimum Effort Theory. Do the same thing for the "Coolness" Theory.

    b. Make up a critical fact or experiment (specific prediction) that, if it were true, would tend to support one theory and contradict the other.

    c. Propose a third theory to explain student seating results and explain how you might test it against the other two theories.

14. It has frequently been observed that democracies do not go to war with each other. This has come to be known as the Democratic Peace.

    a. Make up two theories or models that would account for this observation.

    b. Generate a total of three interesting predictions from the two models and identify from which model they were derived.

    c. Find some critical fact/situation/observation/prediction that will distinguish between the two models. Be explicit about how it simultaneously confirms one model and contradicts the other.

15. A casual look around the world reveals that some governments treat their citizens better than other governments do.

    a. Make up two theories or models that would account for this observation.

b. Generate a total of three interesting predictions from the two models and identify from which model they were derived.

c. Find some critical fact/situation/observation/prediction that will distinguish between the two models. Be explicit about how it simultaneously confirms one model and contradicts the other.

Implicit Bias

In the chapter, we pointed out that scientists are not always value neutral. As with any individual, scientists are likely to have biases that influence how they evaluate, interpret, and act in the world around them. If they are aware of their biases, they can take steps, if they so choose, to minimize the impact of their biases. However, a growing body of research indicates that all individuals have biases of which they are unconscious. These "implicit biases" are activated involuntarily and without an individual's conscious control or even awareness. These implicit biases can be positive or negative, they often have to do with things like race, gender, age, and appearance, and they develop over the course of one's lifetime in response to the direct and indirect messages that one receives from different sources. Due to their different cultural backgrounds, individuals from different countries or regions can often have different implicit biases. What the research tells us, though, is that implicit biases are pervasive and that everyone has them. It also indicates that our implicit biases do not always line up with our conscious attitudes and beliefs. Among other things, implicit biases have been shown to affect which groups gain most from medical research (Romm 2014) and who gets hired to do science in the first place (Moss-Racusin et al. 2012). Unlike explicit biases, implicit biases are not accessible through introspection and, as a result, are much harder to overcome.

Scientists at *Project Implicit* have developed an Implicit Association Test to determine the extent to which people have implicit biases. Take one or more of these tests at https://implicit.harvard.edu/implicit/takeatest.html to see how these scientists try to capture people's implicit biases with respect to things like race, age, and gender.

16. Explain the strategy used by the researchers to try to measure implicit bias. That is, how does the test work?

17. Although the researchers in this example are psychologists, not political scientists, they face the same problem that you'll see again and again in later chapters—how can we measure concepts that are hard to observe? Do you find the approach taken by the scientists who designed the Implicit Association Test to be reasonable? Why or why not? If you find this approach to be unsatisfying, can you think of an alternative?