CHAPTER 2

FALSIFIABILITY

HOW TO FOIL LITTLE GREEN MEN IN THE HEAD

In 1793, a severe epidemic of yellow fever struck Philadelphia. One of the leading doctors in the city at the time was Benjamin Rush, a signer of the Declaration of Independence. During the outbreak, Rush was one of the few physicians who were available to treat literally thousands of yellow fever cases. Rush adhered to a theory of medicine that dictated that illnesses accompanied by fever should be treated by vigorous bloodletting. He administered this treatment to many patients, including himself when he came down with the illness. Critics charged that his treatments were more dangerous than the disease. However, following the epidemic, Rush became even more confident of the effectiveness of his treatment, even though several of his patients had died. Why?

One writer summarized Rush's attitude this way: "Convinced of the correctness of his theory of medicine and lacking a means for the systematic study of treatment outcome, he attributed each new instance of improvement to the efficacy of his treatment and each new death that occurred despite it to the severity of the disease" (Eisenberg, 1977, p. 1106). In other words, if the patient got better, this improvement was taken as proof that bloodletting worked. If the patient died, it merely meant that the patient had been too ill for *any* treatment to work. We now know that Rush's critics were right: his treatments were as dangerous as the disease. In this chapter, we will discuss how Rush went wrong. His error illustrates one of the most important principles of scientific thinking, one that is particularly useful in evaluating psychological claims.

In this chapter, we focus in more detail on the third general characteristic of science that we discussed in Chapter 1: Scientists deal with solvable problems. What scientists most often mean by a *solvable problem* is a "testable theory." The way scientists make sure they are dealing with testable theories is by ensuring that their theories are falsifiable, that is, that they have implications for actual events in the natural world. We will see why what is termed the *falsifiability criterion* is so important in psychology.

THEORIES AND THE FALSIFIABILITY CRITERION

Benjamin Rush fell into a fatal trap when assessing the outcome of his treatment. His method of evaluating the evidence made it impossible to conclude that his treatment did not work. If the recovery of a patient meant confirmation of his treatment (and hence his theory of medicine), then it only seems fair that the death of a patient should have meant disconfirmation. Instead, he rationalized away these disconfirmations. By interpreting the evidence as he did, Rush violated one of the most important rules regarding the construction and testing of theories in science: He made it impossible to falsify his theory.

Scientific theories must always be stated in such a way that the predictions derived from them can potentially be shown to be false. Thus, the methods of evaluating new evidence relevant to a particular theory must always include the possibility that the data will falsify the theory. This principle is often termed the *falsifiability criterion,* and its importance in scientific progress has been most forcefully articulated by Karl Popper, a philosopher of science whose writings are read widely by working scientists (Magee, 1985).

The falsifiability criterion states that, for a theory to be useful, the predictions drawn from it must be specific. The theory must go out on a limb, so to speak, because in telling us what *should* happen, the theory must also imply that certain things will *not* happen. If these latter things *do* happen, then we have a clear signal that something is wrong with the theory: It may need to be modified, or we may need to look for an entirely new theory. Either way, we shall end up with a theory that is nearer to the truth. In contrast, if a theory does not rule out any possible observations, then the theory can never be changed, and we are frozen into our current way of thinking, with no possibility of progress. Thus, a successful theory is not one that accounts for every possible happening because such a theory robs itself of any predictive power.

Because we shall often refer to the evaluation of theories in the remainder of this book, we must clear up one common misconception surrounding the term *theory.* The misconception is reflected in the commonly used phrase "Oh, it's only a theory." This phrase captures what laypeople often mean when they use the term *theory:* an unverified hypothesis, a mere guess, a hunch. This is most definitely *not* the way the term *theory* is used in science. When scientists refer to theories, they do not mean unverified guesses.

A theory in science is an interrelated set of concepts that is used to explain a body of data and to make predictions about the results of future experiments. *Hypotheses* are specific predictions that are derived from theories (which are more general and comprehensive). Currently viable theories are those that have had many of their hypotheses confirmed. The theoretical structures of such theories are thus consistent with a large number of observations. However, when the database begins to contradict the hypotheses derived from a theory, scientists begin trying to construct a new theory that will provide a better interpretation of the data. Thus, the theories that are under scientific discussion are those that have been verified to some extent and that do not make many predictions that are contradicted by the available data. They are not mere guesses or hunches.

The difference between the layperson's and the scientist's use of the term *theory* has been exploited throughout the 1980s and 1990s by some religious fundamentalists who want creationism taught in the public schools (Pennock, 1999). Their argument often is "After all, evolution is only a theory." This statement is intended to suggest the *layperson's* usage of the term theory to mean "only a guess." However, the theory of evolution by natural selection is not a theory in the layperson's sense (to the contrary, in the layperson's sense it would be called a fact). Instead, it is a theory in the *scientific* sense. It is a conceptual structure that is supported by a large and varied set of data (Berra, 1990; Dennett, 1995; Raymo, 1999; Ruse, 1999; Wilson, 1998). It is not a mere guess, equal to any other mere guess.

The Theory of Knocking Rhythms

A hypothetical example will show how the falsifiability criterion works. A student knocks at my door. A colleague in my office with me has a theory that makes predictions about the rhythms that different types of people use to knock. Before I open the door, my colleague predicts that the person behind it is a female. I open the door and, indeed, the student is a female. Later I tell my colleague that I am impressed, but only mildly so because he had a 50 percent chance of being correct even without his "theory of knocking rhythms." He says he can do better. Another knock comes. My colleague tells me it is a male under 22 years old. I open the door to find a male student whom I know to be just out of high school. I comment that I am somewhat impressed because our university has a considerable number of students over the age of 22. Yet I still maintain that, of course, young males are quite common on campus. Thinking me hard to please, my colleague proposes one last test. After the next knock, my colleague predicts, "Female, 30 years old, 5 feet 2 inches tall, carrying a book and a purse in the left hand and knocking with the right." After opening the door and confirming the prediction completely, I have quite a different response. I say that, assuming my colleague did not play a trick and arrange for these people to appear at my door, I am now in fact extremely impressed.

Why the difference in my reactions? Why do my friend's three predictions yield three different responses, ranging from "So what?" to "Wow"? The answer has to do with the specificity and precision of the predictions. The more specific predictions made a greater impact when they were confirmed. Notice, however, that the specificity varied directly with the falsifiability. The more specific and precise the prediction was, the more potential observations there were that could have falsified it. For example, there are a lot of people who are *not* 30-year-old females who are 5 feet 2 inches tall.

Note that implicitly, by my varied reactions, I signaled that I would be more impressed by a theory that made predictions that maximized the number of events which should *not* occur.

Good theories, then, make predictions that expose themselves to falsification. Bad theories do not put themselves in jeopardy in this way. They make predictions that are so general that they are almost bound to be true (for example, the next person to knock on my door will be less than 100 years old) or are phrased in such a way that they are completely protected from falsification (as in the Benjamin Rush example). In fact, a theory can be so protected from falsifiability that it is simply no longer considered scientific at all. Indeed, it was Popper's attempt to define the criteria that separate science from nonscience that led him to emphasize the importance of the falsifiability principle. There is a direct link here to psychology and to our discussion of Freud in Chapter 1.

Freud and Falsifiability

In the early decades of this century, Popper was searching for the underlying reasons that some scientific theories seem to lead to advances in knowledge and others lead to intellectual stagnation (Magee, 1985). Einstein's general relativity theory, for example, led to startlingly new observations (for instance, that the light from a distant star bends when it passes near the sun) precisely because its predictions were structured so that many possible events could have contradicted them and thus falsified the theory.

Popper reasoned that this is not true of stagnant theories and pointed to Freudian psychoanalysis as an example. Freudian theory uses a complicated conceptual structure that explains human behavior after the fact but does not predict things in advance. It can explain everything, but Popper argued, it is precisely this property that makes it scientifically useless. It makes no specific predictions. Adherents of psychoanalytic theory spend much time and effort in getting the theory to explain every known human event, from individual quirks of behavior to large-scale social phenomena, but their success in making the theory a rich source of after-the-fact explanation robs it of any scientific utility. Freudian psychoanalytic theory currently plays a much larger role as a spur to the literary imagination than as a theory in contemporary psychology (Robins & Craik, 1994; Robins, Gosling, & Craik, 1999). Its demise within psychology can be traced in part to its failure to satisfy the falsifiability criterion.

As an eminent critic has argued, "Incorrect but widely dispersed ideas about the mind inevitably end by causing social damage. Thanks to the once imposing prestige of psychoanalysis, people harboring diseases or genetic conditions have deferred effective treatment while scouring their infantile past for the sources of their trouble" (Crews, 1993, p. 65). Take, for example, the history of Gilles de la Tourette syndrome. This is a disorder characterized by physical tics and twitches that may involve any part of the body, as well as vocal symptoms such as grunts and barks, echolalia (involuntary repetition of the words of others), and coprolalia (compulsive repetition of obscene words). Tourette syndrome is an organically based disorder of the central nervous system and is now successfully treated with the drug haloperidol (Bower, 1990, 1996a). Throughout history, individuals with Tourette syndrome have been persecuted—earlier as witches by religious authorities, and in more modern times by being subjected to exorcisms (Hines, 1988). Importantly, understanding of the cause and treatment of the disorder was considerably hampered from 1921 to 1955, when explanations and treatments for Tourette syndrome were dominated by psychoanalytic conceptualizations (see Kushner, 1999; Shapiro, Shapiro, Bruun, & Sweet, 1978). Author after author presented unfalsifiable psychoanalytic explanations for the syndrome. The resulting array of vague explanations created a conceptual sludge that obscured the true nature of the syndrome and probably impeded scientific progress toward an accurate understanding of it. For example, according to one author:

> [Tourette syndrome is] a classic example of the retrogressive effect of psychoanalysis on the investigation of brain disease. La Tourette had attributed the disease to a degenerative process of the brain. After Freud's theories became fashionable in the early decades of the present century, attention in such conditions was deflected from the brain.... The consequence of this retrograde movement was that patients tended to be referred to psychiatrists (usually of a psychoanalytic persuasion) rather than to neurologists, so that physical examinations and investigation were not performed. (Thornton, 1986, p. 210)

Shapiro et al. (1978) described one psychoanalyst who thought that his patient was "reluctant to give up the tic because it became a source of erotic pleasure to her and an expression of unconscious sexual strivings." Another considered the tics "stereotyped equivalents of onanism.... The libido connected with the genital sensation was displaced into other parts of the body." A third considered the tic a "conversion symptom at the anal-sadistic level." A fourth thought that a person with Tourette syndrome had a "compulsive character, as well as a narcissistic orientation" and that the patient's tics "represent[ed] an affective syndrome, a defense against the intended affect." The summary by Shapiro et al. of the resulting theoretical situation demonstrates quite well the harmful effects of ignoring the falsifiability criterion:

> Psychoanalytic theorizing of this kind in effect leaves no base untouched. Tics are a conversion symptom but not hysterical, anal but also erotic, volitional but also compulsive, organic but also dynamic in origin.... These psychological labels, diagnoses, and treatments were unfortunately imposed on patients and their families, usually with little humility, considerable dogmatism, and with much harm.... These papers, because of their subsequent widespread influence,

> had a calamitous effect on the understanding and treatment of this syndrome. (pp. 39–42, 50, 63)

Progress in the treatment and understanding of Tourette syndrome began to occur only when researchers recognized that the psychoanalytic "explanations" were useless. These explanations were enticing because they seemed to explain things. In fact, they explained everything—after the fact. However, the explanations they provided created only the illusion of understanding. By attempting to explain everything after the fact, they barred the door to any advance. Progress occurs only when a theory does not predict *everything* but instead makes specific predictions that tell us—in advance—something specific about the world. The predictions derived from such a theory may be wrong, of course, but this is a strength, not a weakness.

The Little Green Men

It is not difficult to recognize unfalsifiable conceptualizations when one is detached from the subject matter and particularly when one has the benefit of historical hindsight (as in the Benjamin Rush example). It is also easy to detect unfalsifiable conceptualizations when the instance is obviously concocted. For example, it is a little known fact that I have discovered the underlying brain mechanism that controls behavior. You will soon be reading about this discovery (in the *National Enquirer*, available at your local supermarkets). In the left hemisphere of the brain, near the language areas, reside two tiny green men. They have the power to control the electrochemical processes taking place in many areas of the brain. And, well, to make a long story short, they basically control everything. There is one difficulty, however. The green men have the ability to detect any intrusion into the brain (surgery, X rays, etc.), and when they do sense such an intrusion, they tend to disappear. (I forgot to mention that they have the power to become invisible.)

I have no doubt insulted your intelligence by using an example more suitable to elementary school students. However, consider this. As an instructor in psychology, I am often confronted by students who ask me why I have not lectured on all the startling new discoveries in extrasensory perception (ESP) and parapsychology that have been made in the past few years. I have to inform these students that most of what they have heard about these subjects has undoubtedly come from the general media, rather than from scientifically respectable sources. In fact, some scientists have looked at these claims and have not been able to replicate the findings. I remind these students, who not uncommonly have already completed a methodology course, that replication of a finding is critical to its acceptance as an established scientific fact and that this is particularly true in the case of results that contradict either previous data or established theory.

I further admit that many scientists have lost patience with ESP research. Although one reason is undoubtedly that the area is tainted by fraud, charlatanism, and media exploitation, perhaps the most important reason for scientific disenchantment is the existence of what Martin Gardner (1972) once called the catch-22 of ESP research.

It works as follows: A "believer" (someone who accepts the existence of ESP phenomena before beginning an investigation) claims to have demonstrated ESP in the laboratory. A "skeptic" (someone who doubts the existence of ESP) is brought in to confirm the phenomena. Often, after observing the experimental situation, the skeptic calls for more controls (controls of the type we will discuss in Chapter 6), and though these are sometimes resisted, well-intentioned believers often agree to them. When the controls are instituted, the phenomena cannot be demonstrated (see Alcock, 1981, 1990; Druckman & Swets, 1988; Hines, 1988; Humphrey, 1996; Hyman, 1992, 1996; Milton & Wiseman, 1999; Wiseman, Beloff, & Morris, 1996). The skeptic, who correctly interprets this failure as an indication that the original demonstration was due to inadequate experimental control and thus cannot be accepted, is often shocked to find that the believer does not regard the original demonstration as invalid. Instead, the believer invokes the catch-22 of ESP: Psychic powers, the believer maintains, are subtle, delicate, and easily disturbed. The "negative vibes" of the skeptic were probably responsible for the disruption of the "psi powers." The powers will undoubtedly return when the negative aura of the skeptic is removed.

Now comes the surprise. My student does not chuckle at this but says, "Yeah, it makes sense. I can see how the skeptic might give off negative interferences that would mess up ESP." Slightly stunned but maintaining my composure, I gently remind the student of the funny example of the little green men I used in class a few weeks ago. I point out that, although the whole class thought my argument about the little green men was ridiculous, the believer's argument is logically analogous. ESP operates just as the little green men do. It's there as long as you don't intrude to look at it carefully. When you do, it disappears. If we accept this explanation, it will be impossible to demonstrate the phenomenon to any skeptical observers. It appears only to believers. Of course, this position is unacceptable in science. We do not have the magnetism physicists and the nonmagnetism physicists (those for whom magnetism does and does not work). At this point, the student looks a little sheepish. He or she mumbles agreement with me, then usually quickly leaves the room. I, of course, never know whether the agreement is an indication of true understanding or the result of embarrassment and a wish to escape the situation. Perhaps it is a little of both, for the student has been asked to do one of the most difficult things in the world: confront a strongly held belief with contradictory evidence. It can be quite disconcerting to rigorously apply the criteria of scientific explanation to long-held beliefs. But this is just what a psychologist must do. No one ever said it would be easy.

Not All Confirmations Are Equal

The principle of falsifiability has important implications for the way we view the confirmation of a theory. Many people think that a good scientific theory is one that has been repeatedly confirmed. They assume that the amount of confirming evidence is critical in the evaluation of a theory. But falsifiability implies that the number of times a theory has been confirmed is not the critical element. The reason is that, as our example of the "theory of knocking rhythms" illustrated, not all confirmations are equal. Confirmations are more or less impressive depending on the extent to which the prediction exposes itself to potential disconfirmation. One confirmation of a highly specific, potentially falsifiable prediction (for instance, a female, 30 years old, 5 feet 2 inches tall, carrying a book and a purse in the left hand and knocking with the right) has a greater impact than the confirmation of 20 different predictions that are all virtually unfalsifiable (for instance, a person less than 100 years old).

Thus, we must look not only at the quantity of the confirming evidence, but also at the quality of the confirming instances. Using the falsifiability criterion as a tool to evaluate evidence will help the research consumer resist the allure of the nonscientific, all-explaining theory that inevitably hinders the search for a deeper understanding of the nature of the world and the people who inhabit it. Indeed, such theoretical dead ends are often tempting precisely because they can never be falsified. They are islands of stability in the chaotic modern world.

Popper often made the point that "the secret of the enormous psychological appeal of these theories lay in their ability to explain everything. To know in advance that whatever happens you will be able to understand it gives you not only a sense of intellectual mastery but, even more important, an emotional sense of secure orientation in the world" (Magee, 1985, p. 43). However, the attainment of such security is not the goal of science, because such intellectual security would be purchased at the cost of intellectual stagnation. Science is a mechanism for continually challenging previously held beliefs by subjecting them to empirical tests in such a way that they can be shown to be wrong. This characteristic often puts science—particularly psychology—in conflict with so-called folk wisdom or common sense.

Falsifiability and Folk Wisdom

Psychology is a threat to the comfort that folk wisdom of the type discussed in Chapter 1 provides because, as a science, it cannot be content with explanations that cannot be refuted. The goal of psychology is the empirical testing of alternative behavioral theories in order to rule out some of them. Aspects of folk wisdom that are explicitly stated and that do stand up to empirical testing are, of course, welcomed, and many have been incorporated into psychological theory. However, psychology does not seek the comfort of explanatory systems that account for everything after the fact but predict nothing in advance. It does not accept systems of folk wisdom that are designed never to be changed and that end up being passed on from generation to generation. It is self-defeating to try to hide this fact from students or the public. Unfortunately, some psychology instructors and popularizers are aware that psychology's threat to folk wisdom disturbs some people, and they sometimes seek to soothe such feelings by sending a false underlying message that implies, "You'll learn some interesting things, but don't worry—psychology won't challenge things you believe in strongly." This is a mistake, and it contributes to confusion both about what science is and about what psychology is.

Science *seeks* conceptual change. Scientists try to describe the world as it really is, as opposed to what our prior beliefs dictate it should be. The dangerous trend in modern thought is the idea that people must be shielded from the nature of the world, that a veil of ignorance is necessary to protect a public unequipped to deal with the truth. Psychology is like other sciences in rejecting the idea that people need to be shielded from the truth. Biologist Michael Ghiselin (1989) argued further that we all lose when knowledge is not widespread:

> We are better off if we have healthy neighbors, and it would be utter folly to monopolize the supply of medicine in order to be more healthy than they are. So too with knowledge. Our neighbor's ignorance is as bad for us as his ill health, and may indeed be the cause of it. Industry and all the rest of us benefit from a supply of skilled labor. We rely upon others for their skill and expertise. (p. 192)

Psychologists feel, like Ghiselin, that we all lose when we are surrounded by others who hold incorrect views of human behavior. Our world is shaped by public attitudes toward education, crime, health, industrial productivity, child care, and many other critical issues. If these attitudes are the products of incorrect theories of behavior, then we are all harmed.

The Freedom to Admit a Mistake

Scientists have found that one of the most liberating and useful implications of the falsifiability principle is that, in science, making a mistake is not a sin. One of Karl Popper's most influential books has a title that is a beautiful, short summary of how science progresses: *Conjectures and Refutations* (1963). According to Popper, the most useful theoretical predictions (conjectures) are those with very specific implications, those that expose themselves to falsification (refutations). But of course, such predictions will in fact sometimes *be* falsified. Thus, we can say that the falsified hypotheses are mistakes in the sense that they do not reflect the way the world is, but the correcting of such mistakes ultimately brings us closer to the truth. The reason is that falsified hypotheses provide information that scientists use to adjust their theories so that these theories accord more closely with the data. Philosopher Daniel Dennett (1995) has said that the essence of science is "making mistakes in public" (p. 380). By the process of continually adjusting theory when data do not accord with it, scientists eventually arrive at theories that better reflect the nature of the world. Falsification is a mechanism of theoretical change, not to be avoided as it is in everyday life.

In fact our way of operating in everyday life might be greatly improved if we could use the falsifiability principle on a personal level. This is why the word *liberating* was used in the opening sentence of this section. It has a personal connotation that was specifically intended, for the ideas developed here have implications beyond science. We would have many fewer social and personal problems if we could only understand that, when our beliefs are contradicted by evidence in the world, it is better to adjust our beliefs than to deny the evidence and cling tenaciously to dysfunctional ideas. Physicist J. Robert Oppenheimer argued:

> There's a point in anybody's train of reasoning when he looks back and says, "I didn't see this straight." People in other walks of life need the ability to say without shame, "I was wrong about that." Science is a method of having this happen all the time. You notice a conflict or some oddity about a number of things you've been thinking about for a long time. It's the shock that may cause you to think another thought. That is the opposite of the worldly person's endless web of rationalization to vindicate an initial error. (Dos Passos, 1964, pp. 150–151)

How many times have you been in a hot argument with someone when right in the middle—perhaps just as you were giving a heated reply and defending your point of view—you realized ("Oh, my God!") that you were wrong about some critical fact or piece of evidence? What did you do? Did you back down and admit to the other person that you had assumed something that wasn't true and that the other person's interpretation now seemed more correct to you? Probably not. If you are like most of us, you engaged in an "endless web of rationalization to vindicate an initial error." You tried to extricate yourself from the argument without admitting defeat. The last thing you would have done was admit that you were wrong. Thus, both you and your partner in the argument became a little more confused about which beliefs more closely tracked the truth. If refutations never become public (as they do in science), if both true and false beliefs are defended with equal vehemence, and if the correct feedback about the effects of argument is not given (as in this example), there is no mechanism for getting beliefs more reliably in sync with reality. This is why so much of our private and public discourse is confused and why the science of psychology is a more reliable guide to the causes of behavior than is so-called common sense.

Many scientists have attested to the importance of understanding that making errors in the course of science is normal, and that the real danger to scientific progress is our natural human tendency to avoid exposing our beliefs to situations in which they might be shown to be wrong. Nobel Prize winner Peter Medawar (1979) wrote:

> Though faulty hypotheses are excusable on the grounds that they will be superseded in due course by acceptable ones, they can do grave harm to those who hold them because scientists who fall deeply in love with their hypotheses are proportionately unwilling to take no as an experimental answer. Sometimes instead of exposing a hypothesis to a cruelly critical test, they caper around it, testing only subsidiary implications, or follow up sidelines that have an indirect bearing on the hypothesis without exposing it to the risk of refutation.... I cannot give any scientist of any age better advice than this: *the intensity of the conviction that a hypothesis is true has no bearing on whether it is true or not.* (p. 39; italics in original)

Many of the most renowned scientists in psychology have followed Medawar's advice. In an article on the career of noted experimental psychologist Robert Crowder, one of his colleagues, Mahzarin Banaji is quoted as saying that "he is the least defensive scientist I know. If you found a way to show that his theory was wobbly or that his experimental finding was limited or flawed, Bob would beam with pleasure and plan the demise of his theory with you" (Azar, 1999, p. 18). Azar (1999) describes how Crowder developed a theory of one component of memory called precategorical acoustic storage and then carefully designed the studies that falsified his own theory.

But the falsifying attitude doesn't always have to characterize each and every scientist for *science* to work. Jacob Bronowski (1973, 1977) often argued in his many writings that the unique power of science to reveal knowledge about the world does *not* arise because scientists are uniquely virtuous (that they are completely objective; that they are never biased in interpreting findings, etc.) but instead it arises because fallible scientists are immersed in a process of checks and balances—in a process in which other scientists are always there to criticize and to root out the errors of other scientists. Psychologist Ray Nickerson (1998) has made the same point by arguing that is not necessary for every scientist to display the objectivity of Robert Crowder. Like Bronowski, Nickerson believes that it is "not so much the critical attitude that individual scientists have taken with respect to their own ideas that has given science its success...but more the fact that individual scientists have been highly motivated to demonstrate that hypotheses that are held by some other scientists are false" (p. 32).

In essence, what we are talking about here is a type of intellectual honesty—and a particular type of open-mindedness that scientists value most. In the language of the general public, open-mindedness means being open to possible explanations for a phenomenon. But in science it means that and something more. Philosopher Jonathan Adler (1998) teaches us that science values more highly another aspect of open-mindedness: "What truly marks an open-minded person is the willingness to follow where evidence leads. The open-minded person is willing to defer to impartial investigations rather than to his own predilections.... Scientific method is attunement to the world, not to ourselves" (p. 44).

Thoughts Are Cheap

Our earlier discussion of the idea of testing folk wisdom leads us to another interesting corollary of the falsifiability principle: Thoughts are cheap. What we mean here, of course, is that certain *kinds* of thoughts are cheap. Biologist and science writer Stephen J. Gould (1987) illustrated this point:

> Fifteen years of monthly columns have brought me an enormous correspondence from nonprofessionals about all aspects of science.... I have found that one common misconception surpasses all others. People will write, telling me that they have developed a revolutionary theory, one that will expand the boundaries of science. These theories, usually described in several pages of single-spaced typescript, are speculations about the deepest ultimate questions we can ask—what is the nature of life? the origin of the universe? the beginning of time? But thoughts are cheap. Any person of intelligence can devise his half dozen before breakfast. Scientists can also spin out ideas about ultimates. We don't (or, rather, we confine them to our private thoughts) because we cannot devise ways to test them, to decide whether they are right or wrong. What good to science is a lovely idea that cannot, as a matter of principle, ever be affirmed or denied? (p. 18)

The answer to Gould's last question is "No good at all." The type of thoughts that Gould is saying are cheap are those that we referred to earlier in our discussion of Karl Popper's views: grand theories that are so global, complicated, and "fuzzy" that they can be used to explain anything—theories constructed more for emotional support ("an emotional sense of secure orientation in the world" to quote Popper again) because they are not meant to be changed or discarded. Gould was telling us that theories like this are useless for scientific purposes, however comforting they may be. Science is a creative endeavor, but the creativity involves getting conceptual structures to fit the confines of empirical data. This is tough. These types of thoughts—those that explain the world as it actually is—are not cheap. Probably this is why good scientific theories are so hard to come by and why unfalsifiable pseudoscientific belief systems pervade our society.

James Gleick, author of a popular book on chaos theory in mathematics and science, wrote about how, after the publication of his book, he was inundated with letters from purveyors of "cheap thoughts": unfalsifiable grand theories of the universe. In an article titled "Uh-oh, Here Comes the Mailman" (1990), Gleick wrote, "Many of my correspondents these days have conceived completely new cosmologies. It seems that when you write about science, you quickly get onto an international mailing list for psychic discoveries, mathematical proofs, stock market, and grand theories of everything" (p. 32).

Gleick (1990) viewed this response as an indicator of the public's misunderstanding of how science works:

> What's going on here? We're supposed to be living in an era of sophistication about science, an era of universal education, of public-television science specials, of daily newspaper science pages.... Still, whether they are sadly crackpot or wholesomely curious, some people seem to need something different. False knowledge is remarkably durable.... Otherwise intelligent people believe in ESP and parapsychology.... Newspaper and book publishers who should know better have no qualms about feeding their readers' hunger for horoscopes. (p. 32)

Gleick speculated that the reason for this sad state of affairs is that people do not fully understand that certain types of ideas are cheap. As Gould said, any verbally fluent and intelligent person can come up with an unfalsifiable "theory of the universe." Scientific theories, on the other hand, do not come cheap. Gleick (1990) suggested that so many pseudoscientific beliefs exist among the public because:

> science is hard. It is imperfect. People find it easier to absorb weird ideas from science than to understand the process that leads to the ideas. The grand theories and themes sound magical sometimes; but real scientists have to sweat their way through piles of real data from a night at the big telescope or a day in the laboratory. When they make a theory they can't afford to indulge a taste for voodoo. They are about to be proved right or wrong. Another pile of data is on the way. (p. 32)

In this passage, Gleick reinforced Gould's point in the earlier quote: Theories in science make contact with the world. They are falsifiable. They make specific predictions. But I think that Gleick was wrong in one respect. Yes, doing science is hard. Actually coming up with the theories that are truly scientific explanations is a prodigious task. But understanding the general logic by which science works is *not* so difficult. Indeed, there are books about the logic of scientific thinking that have been written for children (Kramer, 1987). So, while science may be hard for *scientists*, there is at least one sense in which it should not be so hard for the general public. The public, however, must be willing to learn a few general principles—no more than the few principles in this book. But the catch is that these principles must be well learned. So let's get on with it.

ERRORS IN SCIENCE: GETTING CLOSER TO THE TRUTH

In the context of explaining the principle of falsifiability, we have outlined a simple model of scientific progress. Theories are put forth and hypotheses are derived from them. The hypotheses are tested by a variety of techniques that we shall discuss in the remainder of this book. If the hypotheses are confirmed by the experiments, then the theory receives some degree of corroboration. If the hypotheses are falsified by the experiments, then the theory must be altered in some way, or it must be discarded for a better theory.

Of course, saying that knowledge in science is tentative and that hypotheses derived from theories are potentially false does not mean that everything is up for grabs. There are many relationships in science that have been confirmed so many times that they are termed *laws* because it is extremely doubtful that they will be overturned by future experimentation. It is highly unlikely that we shall find one day that blood does not circulate or that the earth does not orbit the sun. These mundane facts are not the type of hypotheses that we have been talking about. They are of no interest to scientists precisely because they are so well established. Scientists are interested only in those aspects of nature that are on the fringes of what is known, that is, those things that are not so well confirmed that there is no doubt about them.

This aspect of scientific practice—that scientists gravitate to those problems on the fringes of what is known and ignore things that are well confirmed (so-called laws)—is very confusing to the general public. It seems that scientists are always emphasizing what they don't know rather than what is known. This is true, and there is a very good reason for it. To advance knowledge, scientists must be at the outer limits of what is known. Of course, this is precisely where things are uncertain. But science advances by a process of trying to reduce the uncertainty at the limits of knowledge. Psychologist Robert McCall (1988) discussed the public misunderstanding that results from this characteristic of science:

> Scientists are taught to question, even to quibble, and to ruminate over the solidity of evidence and the interpretation of findings that appear contradictory in the hope of discerning a nugget of truth.... Although such quibbling is an essential part of the scientific process, it can make scientists look indecisive and incompetent to journalists and to the public. What society and some reporters fail to understand is that by definition, professionals on the edge of knowledge do not know what causes what. Scientists, however, are privileged to be able to say so, whereas business executives, politicians, and judges, for example, sometimes make decisions in audacious ignorance while appearing certain and confident. (p. 88)

It should also be emphasized that, when scientists talk about falsifying a theory based on observation and about replacing an old, falsified theory with a new one, they do not mean that all the previous facts that established the old theory are thrown out (we shall talk about this at great length in Chapter 8). Quite the contrary, the new theory must explain all of the facts that the old theory could explain *plus* the new facts that the old theory couldn't explain. So the falsification of a theory does not mean that scientists have to go back to square one. Science writer Isaac Asimov illustrated the process of theory revision very well in a very readable essay titled "The Relativity of Wrong" (1989), in which he wrote about how we have refined our notions of the earth's shape. First, he warned us not to think that the ancient belief in a flat earth was stupid. On a plain (where the first civilizations with writing developed), the earth looks pretty flat, and Asimov urged us to consider what a quantitative comparison of different theories would reveal. First, we could express the different theories in terms of how much curvature per mile they hypothesized. The flat-earth theory would say that the curvature is 0 degrees per mile. This theory is wrong, as we know. But in one sense, it is close. As Asimov (1989) wrote:

> About a century after Aristotle, the Greek philosopher Eratosthenes noted that the sun cast a shadow of different lengths at different latitudes (all the shadows would be the same length if the earth's surface were flat). From the difference in shadow length, he calculated the size of the earthly sphere and it turned out to be 25,000 miles circumference. The curvature of such a sphere is about 0.000126 degrees per mile, a quantity very close to 0 per mile, as you can see.... The tiny difference between 0 and 0.000126 accounts for the fact that it took so long to pass from the flat earth to the spherical earth. Mind you, even a tiny difference, such as that between 0 and 0.000126 can be extremely important. The difference mounts up. The earth cannot be mapped over large areas with any accuracy at all if the difference isn't taken into account and if the earth isn't considered a sphere rather than a flat surface. (pp. 39–40)

But science, of course, did not stop with the theory that the earth was spherical. As we discussed earlier, scientists are always trying to refine their theories as much as possible and to test the limits of current knowledge. For example, Newton's theories of gravitation predicted that the earth should not be perfectly spherical, and indeed this prediction has been confirmed. It turns out that the earth bulges a little at the equator and that it is a little flat at the poles. It is something called an *oblate spheroid*. The diameter of the earth from North Pole to South Pole is 7,900 miles, and the equatorial diameter is 7,927 miles. The curvature of the earth is not constant (as in a perfect sphere); instead, it varies slightly from 7.973 inches to 8.027 inches to the mile. As Asimov (1989) noted, "The correction in going from spherical to oblate spheroidal is much smaller than going from flat to spherical. Therefore, although the notion of the earth as a sphere is wrong, strictly speaking, it is not as wrong as the notion of the earth as flat" (p. 41). Asimov's example of the shape of the earth illustrates for us the context in which scientists use such terms as *mistake, error,* or *falsified.* Such terms do not mean that the theory being tested is wrong in every respect, only that it is incomplete. So when scientists emphasize that knowledge is tentative and may be altered by future findings, they are referring to a situation like this example. When scientists believed that the earth was a sphere, they realized that, in detail, this theory might someday need to be altered. However, the alteration from spherical to oblate spheroidal preserves the "roughly correct" notion that the earth is a sphere. We do not expect to wake up one day and find that it is a cube.

SUMMARY

What scientists most often mean by a *solvable problem* is a *testable theory.* The definition of a testable theory is a very specific one in science: It means that the theory is potentially falsifiable. If a theory is not falsifiable, then it has no implications for actual events in the natural world and hence is useless. Psychology has been plagued by unfalsifiable theories, and that is one reason why progress in the discipline has been slow.

Good theories are those that make specific predictions, and such theories are highly falsifiable. The confirmation of a specific prediction provides more support for the theory from which it was derived than the confirmation of a prediction that was not precise. In short, one implication of the falsifiability criterion is that all confirmations of theories are not equal. Theories that receive confirmation from highly falsifiable, highly specific predictions are to be preferred.

Even when predictions are not confirmed (i.e., when they are falsified), this falsification is quite useful to theory development. A falsified prediction indicates that a theory must either be discarded or altered so that it can account for the discrepant data pattern. Thus, it is by theory adjustment caused by falsified predictions that sciences such as psychology get closer to the truth.