CHAPTER 2

SCIENCE

Evaluating Claims to Knowledge

There once was a horse who could solve mathematical problems. In the early 20th century, a German teacher presented her horse, Clever Hans, to the public. The horse's trainer (or a member of the audience) would ask Hans what 5 + 3 is, and Hans would tap his hoof 8 times. If asked what 4 × 7 is, Hans would tap his hoof 28 times. Clever Hans quickly became famous as word of his abilities spread through amazed eyewitnesses. Can you imagine the stir caused by a horse with an uncanny ability to solve math problems?

It turns out that Hans was gifted, but not in exactly the way that most people thought. The eyewitnesses were not mistaken: Hans could tap out the answer to any math problem put to him. However, it was a mistake to infer that he arrived at the correct answers by performing calculations. It was not until Oskar Pfungst, a psychologist, tested Clever Hans's abilities under controlled conditions that the truth of the matter was revealed. Pfungst learned that Hans would only answer correctly if two critical conditions were met: The individual posing the problem must know the correct answer, and Hans must have an unobstructed view of this individual. When testing was arranged such that either of these conditions was not satisfied, Hans could no longer solve any problems. Remarkably, Hans had somehow mastered the subtle skill of reading nonverbal cues.

Hans was able to tell by the tone of someone's voice that a question was being asked of him, at which time he would begin tapping his hoof. He would then carefully watch the individual who had asked the question. When the correct answer was reached, this person would unknowingly and involuntarily tilt his or her head ever so slightly. This was Hans's signal to stop tapping his hoof, leaving observers to conclude that he had correctly performed mathematical calculations. Pfungst's application of scientific reasoning helped to eliminate an incorrect hypothesis—that Hans was doing math—while corroborating another, equally impressive hypothesis—that Hans was an astute reader of human nonverbal behavior.

Scientific Reasoning

Perhaps because of the way science is typically taught, many people regard science as simply a collection of facts. However, science is far more than the discoveries with which it is equated. It is a method, a process, a way of thinking that offers powerful tools for coming to know reality. Science cuts through wishful thinking by setting up observations under well-controlled conditions to determine what causal factors produce what effects. Science is unique among ways of knowing in that it contains a built-in, error-checking mechanism to root out faulty ideas. The case of Clever Hans and his extraordinary abilities shows precisely how scientific reasoning helps us move beyond premature conclusions to understand reality. By systematically manipulating the situation under which Hans was tested, Pfungst ruled out the hypothesis that Hans could add and multiply. In the end, the only tenable hypothesis was that Hans was reading nonverbal cues to produce his solutions.

The scientific enterprise seeks to continually refine and improve our understanding of reality through a series of successive approximations to the truth. When one interesting or impressive theory is disproved, it is often replaced by something equally or even more fascinating. Of course, how fascinating we find scientific theories to be has no bearing on their validity, which depends entirely on their consistency with all available evidence. The power of science derives from its ability to enhance our knowledge. It simultaneously reduces the threat of an unknown and mysterious world and paves the way for technological innovation. These rewards await anyone willing to engage in an open-minded skepticism, giving all ideas a fair chance but maintaining high standards for what constitutes persuasive evidence.

But what are the tools of scientific reasoning? James Lett (1990) offered a good overview of these tools when he outlined a series of six tests that a claim must pass to warrant belief. Each test reflects one essential component of scientific thinking that can protect us against foolish beliefs. The easy way to recall the six tests is by remembering the consonants of the word "FiLCHeRS": *F*alsifiability, *L*ogic, *C*omprehensiveness, *H*onesty, *R*eplicability, and *S*ufficiency

Falsifiability

For a claim to have even the potential of being scientifically meaningful, it must in principle pose a hypothesis that could be disproved. That is, if indeed the claim is false, there must be some way to demonstrate its falsity. This is called the *falsifiability* of the claim, and it is essential because, without it, the claim is completely insulated from reality. Claims that are so completely shielded from unfavorable evidence that they cannot possibly be falsified are actually devoid of meaning. To demonstrate how vacuous an unfalsifiable claim is, Sagan (1995) suggested what would happen if he claimed that a fire-breathing dragon lives in his garage. Naturally, you would want to test this claim. When you ask to see it, you're told it is invisible. Perhaps you could use paint to make the dragon visible? No, it's incorporeal, so there is nothing to which the paint can adhere. What about putting flour on the floor to see its footprints? That won't work because it floats in the air. How about measuring the heat of its flaming breath? No such luck, it's heatless. You continue to pose reasonable tests that are met with increasingly evasive answers. Eventually you begin to wonder just what it means to speak of a dragon that is invisible, immaterial, hovering in the air, and breathing a heatless fire. How is this different from no dragon at all?

This example illustrates the problem posed by an unfalsifiable claim: Because it cannot be tested, the most substantial evidence supplied is somebody's word. This is simply unacceptable currency in scientific reasoning. The approach scientists take is to say, "Show me," and failure to meet this challenge with evidence relegates a claim to the realm of pseudoscience. Whereas pseudoscientists may shield their beliefs from testing by framing unfalsifiable hypotheses, scientists achieve progress by eliminating mistaken ideas one by one through careful experimentation. This requires not that hypotheses be false—nobody would waste time investigating beliefs known to be untrue—but that there must be some way to test the hypotheses, that they are falsifiable.

Logic

Naturally, any claim to knowledge must be logically sound. The soundness of a logical argument is based on satisfying two criteria. First, the premises on which the argument is based must all be true. Second, the proposed conclusion must validly follow from the premises. If either of these criteria is not met, the argument is defective.

For example, consider the argument used to contend that crop circles—huge geometric patterns pressed into farmer's fields—constitute evidence of extraterrestrial visitations:

1. Crop circles are extremely complex and numerous.
2. Human beings are incapable of such complexity on so grand a scale.
3. Therefore, crop circles are made by extraterrestrials

The first premise is certainly true, for elaborate crop circles have been observed all around the world for decades. The second premise, however, is not true. Two Englishmen, Doug Bower and Dave Chorley, confessed to having made crop circles for 15 years to fool gullible people (Schnabel, 1994). Not only did they claim to have done this, but they also eagerly showed reporters exactly how simple it is to make crop circles using nothing more than two-by-fours and some rope. Imagine how proud they must have felt to hear that their creations were considered too complex to possibly be of human origin. But suppose the second premise *were* true. Would the suggested conclusion logically follow from the premises? Not necessarily. Natural forces can play tricks on us—they are capable, for example, of fashioning intricate patterns like the pyramids on Mars—and it is conceivable that unusual winds or other natural phenomena could have caused strange patterns in wheat fields. Thus, this argument is shaky on several grounds and does not provide logically sound support for the claim that crop circles are evidence of extraterrestrial visitation.

Comprehensiveness

A claim to knowledge must account for *all* the pertinent data, not just select bits and pieces. One claim that fails this test is still widely known and involves the Blackout of '65, when large portions of the East Coast of the United States lost electrical power for several hours. Most people who were alive at the time believe that this was an occasion when an unusually large number of babies were conceived. There was a widespread human-interest story maintaining that a spike in the New York City birthrate occurred on Monday and Tuesday of the week falling nine months after the blackout, which suggests what people do when the lights go out. These data were correct, but only a small part of the relevant evidence was considered. It turns out that a similar spike in births actually occurred *every* Monday and Tuesday and still does to this day. The cause of this anomaly: Doctors prefer not to work on weekends. Therefore, induced labor and Caesarian section births tend to be scheduled for the beginning of the week. Although it is admittedly more interesting than the mundane reality of doctors' scheduling habits, the story of large numbers of babies being born nine months after a massive and extended blackout is simply not true. Whenever a claim cannot explain all the relevant evidence, it fails the comprehensiveness test.

Honesty

It should go without saying that claims to knowledge must be evaluated honestly and without self-deception. However, this is often easier said than done. Judgment can sometimes become cloudy, particularly when a cherished belief is on the line. For example, a study by two psychologists found that children who spent more time in day care received lower grades, worse behavioral ratings by parents and teachers, and lower standardized test scores in the third grade than did children who spent less time in day care (Vandell & Corasantini, 1991). This study was at least as methodologically rigorous as comparable studies on related topics, yet it met with fierce objections from peer reviewers at the top journal in their field, *Developmental Psychology*. Given that these reviewers probably work long hours and may be members of dual-career couples, it seems likely that some or all of their own children spent considerable time in day care. Might this have produced some distaste for the results of the study? It is impossible to know for certain, but it is conceivable that the nature of the findings made a fair, honest evaluation of the study difficult to procure. In light of the many ways that we can deceive ourselves, passing the test of honesty is more challenging than it may appear.

Replicability

A claim to knowledge should be based on consistent results observed in multiple experiments, preferably conducted in multiple laboratories. A fluke might occur in a single experiment, but it is highly unlikely that the same problem would plague repeated experiments in the same way. Therefore, consistent results across repeated tests are far more trustworthy than are inconsistent results, or flukes.

Perhaps the clearest failures to successfully replicate have come from research on extrasensory perception (ESP). ESP researchers typically conduct a huge number of tests, and they analyze their data in myriad ways. By nothing more than chance alone, every so often a surprising pattern emerges that gives the appearance of outstanding performance. The key is to replicate this finding to guarantee that it is not a statistical anomaly, a mere fluke. Despite more than a century of concerted effort spent on ESP investigations, researchers have failed to demonstrate even a single replicable phenomenon. Several terms in the ESP literature reveal just what a failure this effort has been so far (Gilovich, 1991):

- *Psi missing.* This term is used when someone performs *worse* than would be expected by chance alone. Believers in ESP, or "psi" phenomena, argue that this poor performance is evidence for the paranormal because it deviates from chance. But are *very bad* predictions really the type of evidence that supports the beliefs associated with supernatural mental abilities?
- *Experimenter effect.* This effect occurs when individuals are able to perform feats of ESP *except* in the presence of a skeptical observer. Believers in ESP take this to mean that "negative energy" of some sort inhibits mental phenomena. Perhaps what it really indicates is that the experimental conditions allow for cheating—either conscious or unconscious—when no skeptic is present to keep everyone honest.
- *Decline effect.* This effect occurs when someone who initially performs well begins to do worse and winds up at purely chance-level performance. Believers in ESP refuse to recognize this effect as an obvious *failure to replicate.* Flukes such as "beginner's luck" quickly disappear upon repeated testing.

Sufficiency

Finally, a claim to knowledge must be backed by enough evidence to warrant belief. Consider three related points when evaluating the sufficiency of available evidence:

1. *The burden of proof is on the claimant.* This principle is similar to the operation of our criminal justice system: Someone making a claim to knowledge is in the position of a prosecutor, who is responsible for mustering sufficient evidence for that claim. We need not accept the responsibility for *disproving* the claim because if insufficient evidence is provided, this alone suggests that the belief is unwarranted.
2. *Extraordinary claims require extraordinary evidence.* This fairly commonsensical principle means that the more fantastic the claim, the more persuasive the evidence must be to muster belief in that claim.
3. *Evidence based on authority is inadequate.* Do other types of evidence exist besides somebody's say-so? No matter *how* reputable this individual may be, history teaches us that anyone can be wrong. No less a respected scientist than Albert Einstein, for example, seems to have been wrong about an important scientific theory of his time: Einstein rejected the uncertainty principle of quantum mechanics, arguing that "God does not play dice." Physical experiments, however, have verified the predictions of the uncertainty principle with great precision.

The evidence in support of a claim must be satisfactory in all three of these ways to pass the sufficiency test. Keep in mind one final statement: "Absence of evidence does not constitute evidence of absence." For example, those who believe that UFO sightings are good evidence of alien visitation argue that if you cannot prove them wrong, they must be right. Remember, though, that your inability to provide an explanation for each alleged UFO sighting does not mean that there *is* no natural explanation. Indeed, a long list of well-understood phenomena and manmade machines can be and have been misperceived as extraterrestrial spacecraft by human observers; including these (Sagan, 1995):

- Conscious fraud or hoaxes, such as objects suspended on strings against a dark night sky or faked photographs or videos
- Conventional or unconventional aircraft, perhaps spotted during military testing flights and therefore unconfirmed for security reasons
- High-altitude balloons, such as the ones misinterpreted as a crashed saucer at Roswell, New Mexico, in 1947
- Planets seen under unusual atmospheric conditions
- Luminescent insects
- Optical mirages
- Lenticular clouds

- Ball lightning
- Meteors and green fireballs
- Satellites, nose cones, or rocket boosters reentering the atmosphere, two or three of which are destroyed on reentry into the Earth's atmosphere each day, many visible to the naked eye

You need not determine which of these, or other, explanations is the cause of any particular UFO sighting. Rather, as noted earlier, it is the claimant's responsibility to provide sufficient proof that what he or she saw could *not* be explained by any of these sources. Especially in light of the many things that might be seen and misunderstood in the sky, those who proclaim alien visitation via UFOs bear the burden of proof and must provide extraordinary evidence for such an extraordinary claim.

If a claim passes all six FiLCHeRS tests, you can tentatively place some measure of confidence in the belief. To the extent that one or more tests are inconclusive or failed, you would be wise to exercise caution. Those who routinely demand the highest standards of evidence will tend to form the most accurate beliefs and thereby keep from making foolish decisions.

ANTISCIENTIFIC BELIEFS

You might think that the benefits of scientific reasoning speak for themselves or that the verifiable knowledge and astounding technological innovations that they generate adequately support them. However, there are individuals who directly attack the methods of science or, knowingly or otherwise, hold beliefs that are hostile toward science.

Postmodernism

In some corners of the academic world, there remain individuals who endorse the extreme forms of certain closely related styles of thinking—such as postmodernism and cultural relativism—that hold that science is merely "one way of knowing," no more valid than any other. Although a thorough refutation of this belief is beyond the scope of this book, it is worthwhile to discuss several of the most serious weaknesses of postmodernism. (See Norris, 1993, for an extended critique; Sokal & Bricmont, 1998, for a review of the abuse of science and math; or Englebretsen, 1997, for a brief discussion of the degradation in standards of scholarship that characterize postmodernism in academia.)

Postmodern thinking originated in the field of literary criticism, where scholars decided that a text (a term that includes any written material or image intended to convey a message) has no fixed meaning but can be interpreted within different contexts. That is, the intended meaning of the author or creator of a text is beside the point; all that matters is one's personal reaction to it. Two of the central tenets of extreme postmodernism are these: (1) There is no such thing as an external reality. Rather, we each construct our own personal reality. (2) Logical contradictions between interpretations pose no problem, because knowledge is valid only to the extent that we choose to believe that it is valid.

In 1996, Alan Sokal wrote a virtually impenetrable paper entitled "Transgressing the Boundaries: Toward a Transformative Hermeneutics of Quantum Gravity." In it, he applied the tools of postmodern literary criticism to complex issues in physics, highlighting the subjective nature of the world in a critical assault on science. Peer reviewers responded favorably at one of the premiere postmodernist journals, *Social Text*, where the paper was accepted for publication. Some time after the article appeared, Sokal revealed something shocking: It was a hoax. In detail, he showed how he had intentionally misused scientific and mathematical concepts, structured illogical and nonsensical arguments, and reached unjustified conclusions.

How, Sokal asked, could the reviewers and editors of one of the top postmodern journals be so easily fooled? Could it be that they were personally satisfied with his conclusions and emotionally gratified by his gratuitous citations to the work of postmodern writers? In a subsequent book elaborating upon this hoax, *Fashionable Nonsense: Postmodern Intellectuals' Abuse of Science*, Sokal and Bricmont (1998) reviewed the disturbing consequences of a reckless, "anything goes" attitude toward scholarship in which science, mathematics, and anything else smacking of objectivity or a verifiable external reality is scorned and abused.

Perhaps we can make one of the most convincing arguments for the existence of an external reality by considering the possibility of there being no such thing. It is doubtful that anyone truly believes in an existence so empty as this, for it fails to explain the coherence of the world as we experience it, provides no grounds for making choices, and attaches no consequences to any of our actions. There is a certain order to the information that reaches each of us, and our actions produce somewhat predictable results. This is what helps us to make decisions and plan for the future, and the lack of an external reality would produce an existence utterly devoid of direction.

Genuine belief in the nonexistence of external reality also renders education meaningless. If there is no reality—nothing "out there" to know about—then we cannot teach or learn anything beyond our own direct personal experience. Given that the relatively few individuals who claim to endorse this extreme version of postmodernism are employed by academic institutions, this situation is both amusing (because they appear to have chosen a futile direction for their lives' work) and troubling (because they are exposing students to such unfounded ideas).

The tolerance for logical contradictions presents another puzzle. If it is the case that "truth is relative," that what is true for one person might be false for another, how can a society function at all? For example, imagine a criminal justice system in which a defendant could cling to a not-guilty plea on the grounds that this was most harmonious with his or her personal reality, while a jury could evaluate the evidence and decide that—in their personal realities—the defendant is guilty. Clearly, criminal justice cannot tolerate such logical contradictions.

Let's allow postmodern thinking into medicine for a moment. Here's the situation: One doctor believes that you have a malignant tumor that must be removed, whereas another doctor believes that your tumor is benign and will cause no harm. Can we accept the truth of both of these beliefs? Of course not; they contradict one another. Your tumor will or will not cause you physical harm, will or will not endanger your life. It makes no sense at all to entertain the possibility that both doctors' beliefs might be "true for them."

The fuzzy thinking that embraces contradictions breaks down in any real-world scenario. The attempt to relieve the sting of error by elevating all ideas to an equally acceptable status through sheer force of will is a reckless practice with dangerous consequences. Nobody has ever demonstrated a theoretical or practical advantage to looking the other way when two ideas contradict one another. Why, then, should this absurd practice be tolerated at all? The clear message of any real logical contradiction is that at least one of the ideas is wrong.

In contrast to the path-of-least-resistance attitude, feel-good reliance on wishful thinking, and political correctness of extreme postmodernism is a central tenet of the exploration of human reasoning undertaken here: A fundamental goal of critical thinkers is the attempt to know reality as it is. This entails seeking out better and better approximations to the truth by constructing logical arguments based on scientific data.

Although, as mentioned earlier, it is beyond the scope of this book to deal with all of the unlikely premises and faulty logic of postmodernism, I invite readers to judge for themselves which style of thinking is more likely to stimulate theoretical or practical advances in our knowledge of the world. Whereas science has an error-checking mechanism, postmodernism does not provide a way to detect or rule out invalid ideas. For example, if you meet a postmodern thinker, ask him or her to construct an argument that trephining—drilling holes in the skull to let out evil spirits—is inadvisable as medical care, that storks do not deliver babies, or even that the earth is not flat. Within the confines of extreme postmodernism, no such arguments are possible, because anything goes. Everyone's ideas are granted "personal validity," so critical evaluations or general conclusions are impossible. Whereas pseudoscience can cause harm by masquerading as science without delivering on its promises, postmodernism poses an even deeper threat. As our population grows and we become increasingly dependent on technology in many spheres of our lives (e.g., food production and distribution, communications and information technology, physical and mental health care, monitoring the global environment), a rejection of scientific reasoning is the surest way to stunt human progress and put us on a path back to the dark ages.

The "What Works Is Different for Everybody" Fallacy

Another antiscientific notion is captured by a popular catchphrase of many pseudosciences: "What works is different for everybody." That is, different treatments are proposed to work for different people. The unstated implication is that scientists fail to recognize this point and pseudoscientists do. In fact, scientific research provides a mechanism for studying individual differences. For example, if you suspect that a particular treatment works better for women than for men, you can tabulate treatment effects separately for each sex and compare them. In this way, the scientific method allows researchers to examine any potentially influential factor and provide increasingly detailed information regarding what type of person is most likely to benefit from what type of treatment for what type of problem.

All of this specificity is proclaimed but utterly unsubstantiated by practitioners of pseudoscience. Where is the evidence indicating which treatment to choose for a given person? As Chapter 3 will discuss further, the tendency of pseudoscientists is to escape into the fog of holism. But what does one *do* with an assessment of the "whole person"? Without guidelines for systematic decisions based on reliable individual differences, it's anybody's guess. Individuating information is sorely lacking because it can *only* be obtained through carefully designed scientific research.

The truth is that there is no evidential support for the holistic prescriptions of pseudosciences. But think about what a wonderful cover-up for ignorance the "what works is different for everybody" ruse can be. It allows practitioners to diagnose and treat health conditions in any manner that they like, and any inconsistencies or failures can be chalked up to enigmatic and unspecified individual differences. When your own health is on the line, however, it seems well advised to stick with a practitioner who can justify his or her treatment plans with research evidence and offer you a prognosis based on data, not a combination of anecdotes and wishful thinking.

The "Probability Is Irrelevant to the Unique Individual" Fallacy

Another variant on the alleged need to personalize treatments involves the claim that because scientists study groups of patients, scientific results have no bearing on the individual case. Although the premise is usually true—scientists do tend to study groups—the conclusion is false: Scientific results are *crucial* for understanding the individual case. In fact, the only way to uncover general principles that afford sound predictions for new individuals is to search through data collected from large groups for patterns, trends, or other regularities. The only way to offer rational advice is to generalize from previous knowledge. If there are no general principles, then nobody could offer advice that would predict the future better than chance.

The conclusion that the patterns discovered in research are irrelevant to unique individuals is absurd. How else could we learn? Is each new case to be an exception to the rules? (If so, how did the rules emerge in the first place?) Is it necessary to consider additional variables? If so, which ones? More to the point, how can you *know* that other variables are important if they have not been studied? Rather than relying on guesswork, we must perform research to determine whether additional variables are in fact important.

Despite this logic and the fact that denying it only introduces additional difficulties, many people seem to feel that probability is irrelevant. "It makes no sense," the argument goes, "to apply probabilities to the single case. The patient before me is a unique individual." However, a simple thought experiment may convince you that probabilities are relevant to each and every single case (Meehl, 1973).

Imagine that you are going to play Russian roulette once, and you are offered your choice of either of two revolvers to use. The first contains one bullet in its cylinder, with five empty chambers; the second contains five bullets with one empty chamber. You will choose one gun, put it to your head, and pull the trigger once. This is a unique event, a single case that will never be repeated—you will either live or die. *If you truly believe that probability is irrelevant to the single case, you should not care which gun you choose.* I have yet to meet anyone so cavalier. The *only* rational way to select a gun is to consider the probability of dying and choose a one-sixth chance as less risky than a five-sixths chance. That everybody chooses the gun with just one bullet shows that, deep down, we all clearly recognize that probability is of the utmost importance in making a smart decision about each and every single case that comes before us. When scientific research evidence—which is inherently probabilistic, like the difference between the two hypothetical guns—is available to guide decision making, ignoring it would violate the most fundamental principles of medical ethics. Throwing up one's hands and saying that "probability is irrelevant to the unique individual" introduces chaos. Clearly, we all know better when our own life is on the line. Regardless of what a person may say or write, it seems doubtful that anyone sincerely and consistently believes that probability is irrelevant to the unique individual.

Voodoo Science and Legal Standards

Postmodernism and the reasoning fallacies described earlier, in addition to many variations on these themes, constitute attacks on scientific reasoning. Think for a moment about who stands to gain from the rejection of science as the best available tool that we have for evaluating the accuracy of knowledge. Science consists of well-established procedures for sorting cold facts from warm, fuzzy fictions. Is it any wonder then that many people will seek to ignore or dispute the scientific method when its results threaten their interests?

In his book *Voodoo Science: The Road from Foolishness to Fraud,* Robert Park (2000) distinguishes among four types of what he calls *voodoo science,* each of which deviates from genuine science in important ways. First, there is *pseudoscience,* the focal point of this book, that involves many of the outward appearances of science but lacks its rigorous methodology and skeptical reasoning. Second, there is *pathological science,* in which people deceive themselves. Individuals engaged in pathological science may be unaware that they are designing studies or evaluating data in a biased fashion, giving all benefit of the doubt to their pet theory. Third, there is *fraudulent science,* cases in which people fabricate or selectively report information with the intention of deceiving others. Fourth, there is *junk science,* theories that are based on what might be possible rather than what has been tested and supported. Park notes that junk science often finds its way into courts of law. It is instructive to consider what the courts have done to curb the introduction of junk science into civil and criminal proceedings because this tells us much about the nature of science and its proper role in modern society.

In 1923, a U.S. circuit court ruled in a case that involved the question of whether a witness could be qualified as an expert to introduce polygraph (or "lie detector") evidence at trial. In the court's decision, this witness was not allowed to present polygraph test results because the technique was not "sufficiently established to have gained general acceptance in the particular field in which it belongs" (*Frye* v. *United States;* for more on polygraph techniques and their utility, see Chapters 10 and 13). This *general acceptance* criterion assisted judges in making determinations about what was acceptable scientific evidence until the U.S. Supreme Court furnished a more explicit set of guidelines 70 years later.

In the case of *Daubert* v. *Merrell Dow Pharmaceuticals, Inc.,* the plaintiff argued that Bendectin, an FDA-approved drug used to treat morning sickness in pregnant women, had caused two particular children's birth defects. As reviewed in Foster and Huber (1997), the available data from many large-scale epidemiological studies provided no evidence that children born to mothers who used Bendectin were at higher risk for birth defects than were children whose mothers had not used this drug. At issue throughout the appeals process was whether the plaintiff could introduce expert witnesses to present their interpretation of this evidence. In an affidavit for the case, one prospective witness, Shanna Swan, stated:

> It is my opinion that one cannot conclude that there is in fact no statistical association between Bendectin and limb reduction defects. More specifically one cannot draw a conclusion from those studies that Bendectin does not cause limb reduction defects [Swan, 1992].

Before examining the Court's ruling, note that this statement betrays a failure to understand the importance of making falsifiable claims to operate in the scientific realm. Scientific research can assess the *extent* of a potential risk, but it can never prove that a drug (or other treatment) is without risk. The reason is that any study contains a margin of error, which is often largely determined by the size of the study. One study might suggest that a particular adverse effect occurs anywhere from one third *less* often to three times *more* often when using a drug, as compared to a control condition in which people did not use the drug. The range from one third to three represents the margin of error. To reduce the uncertainty and obtain a better estimate of whether the drug influences the risk of an adverse effect, one would have to perform a larger study or combine the data across multiple studies. By convention, epidemiologists consider a risk to be practically significant when it constitutes at least a doubling of the original risk. In the case of rare events such as birth defects, detecting even a doubled risk requires studies with many thousands of births.

With regard to Bendectin, the available data provided no evidence that there was an increase of this magnitude. Experts disagreed about the actual margin of error, but the findings appeared to rule out practically significant risks. Swan's statement, however, asks scientists to do the impossible, to prove that there is *no* risk. The problem with this is that one cannot altogether eliminate the margin of error in scientific research. For example, even if millions of births were studied and the margin of error suggested that there was not even a 10 percent increase in risk of birth defects, it would remain possible that an even larger study might detect a 5 percent increase or a 1 percent increase. There is no way to be certain that the increase is precisely 0 percent, which Swan appears to require. We cannot afford to perform studies with millions of people to test for all possible positive and negative effects of every proposed treatment. We have to do the best we can with limited time, money, and other resources, to narrow the margin of error and reach informed judgments using the best available data. If the FDA insisted that drug companies demonstrate evidence of 0 percent risk increase to gain approval for sale, it would be unable to approve *any* medications—past, present, or future.

In part because it recognized the importance of making falsifiable claims, the U.S. Supreme Court ultimately ruled that Swan's testimony, as well as that of others prepared to interpret the epidemiological evidence on behalf of the plaintiff, was inadmissible. In doing so, the Court suggested that judges should assume the role of "gatekeepers," determining whether witnesses would be allowed to testify in accordance with a number of flexible guidelines that can be helpful to distinguish legitimate scientific evidence and expertise from junk science. The Court's criteria, which were based on its own careful study of the philosophy of science, can be summarized as follows (Grove & Barden, 1999):

1. Is the proposed theory on which the testimony is to be based *testable*? Theories that are not falsifiable, that cannot be tested, should not be admissible as science in court.

2. Has the proposed theory been tested using valid and reliable procedures and with *positive results?* Just because a theory can be tested does not guarantee that it will survive such tests—supportive empirical evidence is required.
3. Has the theory been subjected to *peer review?* The Court recognized peer review as an essential quality-control mechanism to separate real science from junk science.
4. What is the known or potential *error rate* of the scientific theory or technique? In addition to demonstrating that a theory or technique is more accurate than chance, it is highly desirable to have a trustworthy estimate of how often it will be mistaken.
5. What *standards,* controlling the technique's operation, maximize its validity? This criterion relates to the Court's concern that there be a logically relevant connection between the theory and the facts involved in a particular case. For legal purposes, one must demonstrate the applicability of a scientific theory or technique to the case at hand.
6. Has the theory been *generally accepted* as valid in the relevant scientific community? This is similar to the requirement from the earlier *Frye* case, but it is now one of several criteria to be considered.

The Court emphasized that these criteria serve as examples of the types of questions that judges should ask when attorneys seek to admit scientific expert testimony. These criteria were not meant to be exhaustive but to highlight the most important issues that should be thoroughly considered in most cases. The reasoning of the Court recognizes that with regard to all four types of voodoo science, the methods of science and scientific reasoning pose threatening obstacles to those who wish to make unsupported claims. It sought to restrict those who wish to find a profitable niche while believing whatever they choose and behaving in whatever manner they like.

Robyn Dawes (1994) uses the term *argument from a vacuum* to refer to beliefs that are served up without sufficient evidential support. Pseudosciences typically rely on such arguments from a vacuum. In contrast to the freewheeling abandon of such arguments, science constrains and focuses our thinking by imposing the limits of known reality. But some people find it bothersome to be limited in this way. They would like to operate completely free of the bounds of reality, and denying the legitimacy of science allows them to do this.

Your safest defense against unfounded claims is to learn some of the rudiments of scientific reasoning. You cannot always count on others to look out for your best interests and must therefore protect yourself. The FiLCHeRS criteria (falsifiability, logic, comprehensiveness, honesty, replicability, and sufficiency) constitute one valuable set of scientific reasoning skills that you can profitably adopt into your own everyday thinking with a little practice. Similarly, the *Daubert* criteria established by the U.S. Supreme Court suggest useful ways to evaluate the merits of claims based on alleged scientific support. Outside of courts of law, however, we must act as our own gatekeepers to distinguish genuine science from its many imposters. Only by applying the highest standards of evidence can you form and retain the most accurate beliefs. If you have lax standards, you will be mistaken—and taken advantage of—more often.

InfoTrac College Edition

If your instructor ordered the version of this book with InfoTrac College Edition, go to: http://www.infotrac-college.com/ and enter the following search terms:

day care and psychology
extrasensory
falsifiability
postmodernism