CHAPTER 1 SCIENTIFIC METHOD

1 SCIENTIFIC METHOD

1.1 THE APPEAL OF SCIENCE

The mission of *the Starship Enterprise* in the popular television series *Star Trek: The Next Generation* was "To boldly go where no one has gone before," a phrase which also aptly describes the mission of people who dedicate their lives to science. Some scientists, like those on the *Enterprise*, explore "strange, new worlds" at the farthest reaches of the universe; most chart new territories closer to home.

Psychologists, the scientists we focus on in this book, try to understand the world of the child (a world we once lived in but no longer know), or study how people organize the "blooming, buzzing, confusion" that surrounds us, or work to piece together a picture of what life is like for the animals that share the world with us. Because psychology is a relatively new science, there are many frontiers to explore.

Antoinette and John Lilly explored one such uncharted region in their pioneering research on communication in dolphins and whales. Their reports of the experience convey a sense of the excitement and challenge that attract people to science and keep them involved in it. In the following quotation, Antoinette Lilly describes the thrill of her first close encounter with a whale:

> *The joy of the next few minutes can only be described as absurd. . . . This whale's invitation to share her world gave me a glimpse through a cosmic crack between species ... a oneness of all living beings as we will know them someday in the future ... a place we have been before and will return to again ... a peaceful promise . . . the "peaceable kingdom." (Lilly, in Ferrucci, 1990, p. 239)*

John Lilly tells of the apprehension that also accompanied them on their foray into that unknown world:

> *This opening of our minds was a subtle and yet a painful process. We began to have feelings which I believe are best described by the word "weirdness." The feeling was that we were up against the edge of a vast uncharted region in which we were about to embark with a good deal of mistrust in the appropriateness of our equipment. (Lilly, in Ferrucci, 1990, p. 238)*

The recollections of these pioneers give us some sense of the intense and even uncanny emotions that accompany research on the frontiers of science. They also demonstrate that, despite popular stereotypes, scientists are not a special breed of people set apart from others by their superhuman rationality and robotlike detachment; and science is anything but the dull, methodical enterprise it is often assumed to be. Quite the contrary is true. Scientists are passionate people and scientific work is ignited and fueled by passion; otherwise it would not be possible for scientists to give so much of themselves to it.

For some scientists, inspiration comes from the promise science holds out of answering age-old questions of the meaning of life and/or our place in creation. Jean Piaget, the psychologist whose pioneering research revolutionized our understanding of children's thought, is one example:

> *I recall one evening of profound revelation. The identification of God with life itself was an idea that stirred me almost to ecstasy because it now enabled me to see in biology the explanation of all things and of the mind itself. The problem of knowledge . . . suddenly appeared to me in an entirely new perspective and as an absorbing topic of study. It made me decide to consecrate my life to the biological explanation of knowledge. (Piaget, 1952, p. 240)*

For others, science offers an opportunity to make one's life meaningful by doing work that makes a difference, and the potential of achieving a type of immortality—the chance to leave one's mark on the world. According to Thomas Kuhn:

> *What . . . challenges [the scientist] is the conviction that, if only he is skillful enough, he will succeed in solving a puzzle that no one before has solved or solved so well. (Kuhn, 1970, p. 38)*

Yet the possibility of making a contribution is not the only inspiration to scientific work. It's not just solving the puzzle that excites the scientist, it's finding just the right puzzle, and working on the puzzle too. George Kneller, a historian of science, puts it this way:

> *The scientist studies nature not simply because it is useful but because he delights in it. He sees beauty in the harmony of nature's parts which his mind is able to grasp. "Intellectual beauty," wrote Henri Poincare, "is sufficient unto itself, and it is for its sake, more perhaps than the future good of humanity, that the scientist devotes himself to long and difficult labors." (Kneller, 1978, p. 151)*

1.2 DISTINGUISHING FEATURES OF SCIENCE

1.2.1 Appeal to Evidence

Scientists work at discovering facts and inventing theories to explain them. A basic assumption underlying their efforts is that it is indeed possible to make sense of the events being considered. Science involves a continuous interplay between collecting observations and thinking about them; the aim is to develop formal principles to explain what has been observed.

Much of a scientist's time is spent carefully observing events, asking questions, formulating answers, and checking them by observing again. When things go well, explanatory principles and theories to account for their own and others' observations are the result. These explanatory principles and theories, in turn, lead to predictions, which are tested by making more observations.

The essence of the scientific method is the acquisition of facts and the testing of ideas by *appealing to the evidence*. No matter how much the scientist may want the results to turn out a particular way, a rule of scientific procedure is that judgment must be suspended until the evidence is in. This approach to acquiring knowledge—through observation and experimentation—is called *empirical*. All sciences that involve research are *empirical sciences*.

In England, in the middle of the 17th century, a group of thinkers committed to the empirical method refused to accept established truths that were backed up only by the authority of the church and state. These skeptics formed a society of revolutionists who vowed to "listen to the answers experiments give us and no other answers!" (de Kruif, 1926, p. 5). The society, called *The Invisible College,* met secretly to avoid the death penalty that would have been imposed on them if they had been discovered in their heretical activities. Robert Boyle, the chemist and physicist, and Isaac Newton, the physicist, were members of this group of scientists that later became the prestigious Royal Society of London (de Kruif, 1926, p. 7).

In the spirit of *The Invisible College*, scientists today are trained to be skeptical and to accept nothing on authority. Whereas the nonscientist may believe an idea to be true just because it feels right, because it seems logically correct, because it has always been believed, or because an authority claims it to be so, these bases for beliefs are unacceptable in science. The scientific community accepts as valid knowledge only those statements about events that are supported by convincing evidence.

> In science, statements about events must be backed up by empirical evidence.

Scientists aim to produce error-free results, the kind of evidence that will be convincing to themselves and other scientists. One way to reduce the likelihood of drawing conclusions based on accidental, one-time happenings is to systematically repeat one's observations and calculations. When researchers get a given result over and over, that is, when results can be *replicated,* the likelihood of error produced through accident is reduced.

> In science, errors in observing or in drawing conclusions must be avoided by systematically checking and repeating one's observations.

1.2.2 Rules of Evidence

Scientists are acutely aware of errors that can result from group and personal biases (from cultural assumptions, desires, values), readily acknowledging that their passion can foster blindness and self-deception. Antoine Lavoisier, the founder of modern chemistry, cautioned scientists against seeing and recording only evidence that fits their preconceptions (Ferrucci, 1990, p. 207). Charles Darwin, the naturalist who developed the theory of evolution, systematically wrote down observations that contradicted his ideas; he knew he would forget these most easily (Ferrucci, 1990, p. 207).

It is to counteract such sources of error that scientists have developed a community standard for testing and evaluating the truth of their assertions. To be accepted in science, research must meet agreed upon criteria for objectivity and precision. Only certain types of questions—those that are potentially answerable by collecting observations—qualify as scientific questions. Measures and procedures— instruments, research design, and data analysis— must comply with accepted practice in the field. Original methods and procedures have to be fully explained and justified.

1.2.3 Cycles of Discovery and Validation

So far, we have discussed the scientist as an individual collecting observations and formulating and testing principles for making sense out of those observations; but the great strength of science is that it is a *collective enterprise*.

Kneller (1978) has distinguished two phases in scientific work. The first he calls the *cycle of discovery*. In this phase, investigators work alone or in a team, collecting observations, formulating and testing tentative understandings. It is this activity that most people envision when they think of science, and it is this phase of scientific work that we have discussed so far in this chapter. But the image of science as solitary is misleading. For even in their solitary activities, scientists are mindful of the second phase of research, what Kneller calls the *cycle of validation*—the review of the work by other scientists.

In the cycle of validation, the scientist reports the research to scientific peers. In addition to carefully evaluating the merits of the work, these scientists may do research to replicate its findings or to test rival explanations of its results. Science depends on both types of activity—on committed researchers discovering and testing ideas, and on a community of scientists ready to question evidence and interpretations, and offer alternative interpretations and the evidence to support them.

The aim in science is to eliminate insights or observations that are available only to particular persons and to accept only knowledge that has been verified by others in the scientific community. In science, evidence must be public.

> Any trained scientist using the same methods must be able to reproduce the results.

When other scientists can repeat a procedure and replicate its results, the likelihood that false ideas will be detected and corrected is increased. The requirement that results be reproducible is an important check on systematic error produced by bias in observers, inadequacies in measuring instruments, or other peculiarities of the testing situation.

1.2.4 Progress

A scientist from just a century ago would surely be amazed at the great strides we have made in our understanding of human behavior—the facts are now different and antique theories have given way to new understandings. We no longer think you can read character in the bumps of the skull or in facial features. Modern psychologists no longer treat mental disorders by applying magnets to the body or by administering the "bath of surprise" (throwing a blindfolded patient into a river or lake).

The success of science in establishing new facts and developing theories to explain known facts is responsible, in part, for the respect we give to scientific work and to those who devote their lives to it. As a result of such advances, science has earned a reputation as perhaps our most trustworthy source of knowledge.

A scientist from the past would marvel not just at the knowledge we now have but also at the advances we've made in understanding how to answer questions in science, and the scientist would view these advances as anything but trivial. For scientists know that it's not their brilliance as a group (though many are) nor their superiority as observers (though many are) that makes science such a reliable source of knowledge. It is its methods.

1.3 Process of Scientific Research

1.3.1 Identifying a Research Problem

Empirical research begins with a research problem. This is true both in original research on a problem and in research conducted to assess the validity of that work. The problem is usually phrased in the form of a question. In psychology, we ask: What kinds of learning are animals capable of? Is intelligence inherited? How often do people dream? How does empathy develop in children? How should mental disturbances be classified? To be suitable for science, the question that is posed must be potentially answerable by the appeal to evidence.

1.3.2 From Observation to Explanation

Research in a new field, or on a new problem, or on previously unexamined phenomena, is likely to focus more on "getting the facts" than on testing a theory. At an early stage of the research, one's theories may even be seen as biases, untested assumptions and prejudices, that could "get in the way" of really seeing what is going on. For this reason, researchers sometimes try to set their preconceptions aside and let the data "speak for themselves." However, some philosophers of science question whether there really can be anything like pure fact-finding. They believe that the way we structure the events we encounter is colored by our basic assumptions and preconceptions, our theories about the world. Carl Hempel (1966) concluded that all research is guided by theory. He reasoned that whenever we are advanced enough in our thinking to select a particular set of events to relate, we are testing a theory, albeit an informal and perhaps poorly articulated one.

B. F. Skinner, whose pioneering research led to the principles of operant conditioning, disagreed. Skinner did not think that he had been guided in his research by theory, especially in the beginning.

Skinner described his discovery of "curves of extinction" as primarily a result of a fortunate accident. When his apparatus for delivering food pellets broke down, the rats in his experiment failed to receive a pellet for each correct response they made. The observations Skinner then made led to his famous "curves of extinction" of learned responses.

> *I am not saying that I would not have got around to extinction curves without a breakdown in the apparatus; Pavlov had given too strong a lead in that direction. But it is still no exaggeration to say that some of the most interesting and surprising results have turned up first because of similar accidents. (Skinner, 1959, p. 367)*

Whether we use the term "theory" or not, it's clear that in research on a new problem, the first observations are guided only by the vaguest expectations. But once the basic observations are in, they become the "raw material" used in formulating a generalization to account for what is observed. Skinner thought of this process as a matter of being in the right place, at the right time, when the right things happened. But philosophers of science, whose job it is to reflect on the process of scientific activity, use more formal terms, taken from logic, to describe this process. They call this type of reasoning *induction*.

> *Induction* refers to the process of reasoning from particular facts or individual cases to a general conclusion.

Whenever researchers observe particular events and then formulate a generalization to explain them, or make particular observations and draw general conclusions from them, they are using induction or inductive reasoning. Induction is involved in every research study; scientists make only certain observations yet draw general conclusions from them. Induction also is involved in all theory building.

The process of induction is mysterious. Studying the results of exploratory observations, the scientist hopes for the "flash of insight," the eureka experience, that will explain and systematize them. Unfortunately, no one knows the precise set of ingredients that leads to useful generalizations or hypotheses. To quote Hempel:

> *Scientific hypotheses and theories are not derived from observed facts, but invented in order to account for them. They constitute guesses at the connections that might obtain between the phenomena under study, at uniformities and patterns that might underlie their occurrence. "Happy guesses" of this kind require great ingenuity, especially if they involve a radical departure from current modes of scientific thinking. (Hempel, 1966, p. 15)*

But we do know some of the ingredients. Happy guesses do not come completely out of the blue. Skinner knew that he was onto an important principle of learning because he had been prepared for the discovery of extinction curves by studying the work of the famous Russian physiologist Ivan Pavlov. The prepared mind can see meanings and possibilities in events when other minds, less prepared, cannot.

Happy guesses also seem to require total absorption, sometimes even obsession, with solving the problem. Thomas Edison was so involved in his experiments that he forgot to attend his own wedding. Marie Curie reported that while she and her husband were engaged in the work that led to the discovery of radium, they lived "with a single preoccupation, as if in a dream" (Ferrucci, 1990, p. 226). Happy guesses sometimes emerge only after months or years of hard work. For this reason, scientists must be able to endure uncertainty, even confusion, over long periods of time.

Historians of science point to other ingredients that are important in this creative process. Thomas Kuhn (1970) concluded that revolutionary ideas in science come most often from people new to a field or from young people, whose backgrounds and academic training allow them to see things in a fresh way. Jean Piaget is a good example.

Piaget earned his doctorate in zoology before becoming a developmental psychologist. Right from the beginning, Piaget (1952) asked different questions about intelligence than were standard in the field. Piaget's first job as a psychologist required him to test schoolchildren's intelligence. The standard procedure was to count the number of correct responses on an intelligence test. But Piaget was curious about the reasoning behind the children's answers; so he went beyond the usual practice, asking the children to explain each of the answers that they gave. Piaget also was able to use the observation skills that he had acquired in studying sparrows and mollusks to make systematic observations of his own children's behaviors in infancy and childhood. These methodological innovations led eventually to a revolutionary theory of intellectual development.

Piaget is only one of many innovators in psychology who came from other disciplines. Indeed, many of the pioneers in research methods that we discuss in this book were trained in fields other than those in which they made their major contributions.

Paul Feyerabend (1975) provides a thought-provoking explanation for why newcomers so often make important innovations in science. Feyerabend believes that scientific revolutions result from criticism of the basic prejudices and assumptions of a discipline. This can't happen, he says, from within the discipline because "prejudices are found by contrast, not by analysis" (Feyerabend, 1975, p. 31). He asks:

> *How can we possibly examine something we are using all the time? How can we analyze the terms in which we habitually express our most simple and straightforward observations and reveal their presuppositions? How can we discover the kind of world we presuppose when proceeding as we do? The answer is clear: we cannot discover it from the inside. We need an external standard of criticism, we need a set of alternative assumptions or, . . . an entire alternative world, we need a dream-world in order to discover the features of the real world we think we inhabit. (Feyerabend, 1975, pp. 31-32)*

Kuhn's and Feyerabend's observations suggest that a nonconforming personality may be an asset in science. Indeed, Frank Sulloway (1996), a historian of science, found that great theoretical advances often are made by the youngest child in a family, the last born, the family rebel.

1.3.3 Hypotheses

Early in this century, psychologists offered competing answers to the question "How do animals learn?" One influential theorist, Edward

Thorndike, thought that animals learned only by blind, trial-and-error. In Thorndike's view, unintelligent stimulus response connections were blindly "stamped-in" whenever an animal's response led to satisfaction. Wolfgang Kohler, a proponent of the Gestalt theory of learning, thought that animals were more intelligent than Thorndike gave them credit for. He thought animals responded to "gestalts," that is, to the whole character of situations, to the relationships between stimuli rather than to the absolute properties of stimuli, as Thorndike's theory asserted.

These ideas of Thorndike and Kohler are tentative answers to research questions. Philosophers of science call such ideas *hypotheses* when the focus is on using the idea as a guide to empirical research. Hempel offers the following definition of a hypothesis:

> *[The hypothesis is] whatever statement is under test, no matter whether it purports to describe some particular fact or event or to express a general law or some other, more complex proposition. (Hempel, 1966, p. 19)*

It is the hypothesis, the tentative answer to the research question, not the question itself, that determines which observations to make in a study (Kneller, 1978). Research that is guided by a hypothesis is said to involve *hypothesis testing*. Most of the research published in psychology journals tests hypotheses.

So far, we have discussed how hypotheses are developed inductively. Both Thorndike's and Kohler's hypotheses about animal learning came about in this way— by observing what animals do and trying to make sense of it. But this picture of hypothesis formation is incomplete. Hypotheses also are derived from theories and suggest themselves in the process of research. A theory, like Kohler's Gestalt psychology, leads to many specific hypotheses. Unexpected findings in research prompt the refinement of working hypotheses and the replacement of older hypotheses by newer ones.

Early in a program of research, hypotheses are more likely to be rough guesses at how the events of interest might be related. The researcher might wonder, for example, how sensory deprivation affects critical thinking or whether the self-esteem of introverts is different from that of extroverts. As the research progresses and more is learned about the phenomena being studied, hypotheses become more specific.

1.3.4 Designing a Test of the Hypothesis

Thorndike developed his hypothesis, that animals form stimulus-response connections, by watching chickens learn to escape from an enclosure. Kohler came up with his hypothesis, that animals take account of the relationships between stimuli they encounter, by watching apes and other animals in a variety of problem-solving situations. All Thorndike and Kohler could observe were animals making particular responses. The processes they hypothesized to underlie these responses could not be observed; Thorndike saw no "stimulus-response connections," nor Kohler any "gestalts." These hypothetical processes were invented to make sense out of the concrete behaviors they did observe.

The difference between explanatory concepts and the events they explain becomes important in hypothesis testing. To test a hypothesis, the researcher must make inferences about what will be observed in a concrete test situation. Hempel called such predictions the *test implications* of a hypothesis.

> The *test* implications of a hypothesis are if . . . then statements based on the assumption that the hypothesis is true.

They are predictions that if certain conditions hold true, certain other events also will hold true.

The logical process used to derive test implications is called *deduction* by philosophers of science:

> Deduction refers to the process of reasoning from a premise to a logical conclusion or from a general principle to specific observed events.

When researchers derive a particular hypothesis from a theory, or predict what will happen in a particular test situation on the basis of a previously established principle, they are using deductive reasoning.

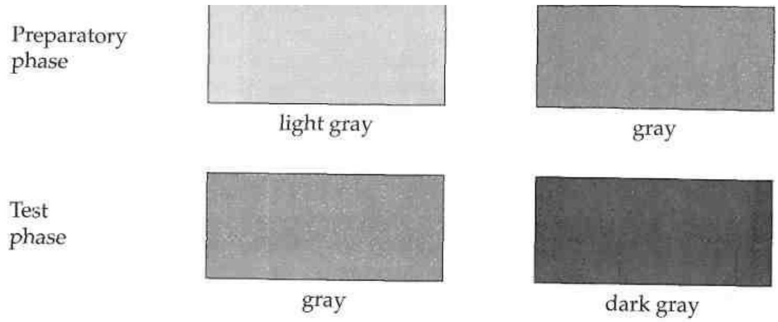
One of Kohler's subjects discovers how to stack boxes to retrieve a banana.

To test his hypothesis, Kohler had to devise a concrete test situation that would demonstrate that his animals had learned by perceiving gestalts and not by forming stimulus-response connections. He first had to reframe his hypothesis as an *if . . . then statement* of the sort: Given that animals learn by perceiving the relationships between stimuli, if they are placed in this type of learning situation, they will behave as follows.

Kohler (1925, in Heidbreder, 1961) decided to test hens in a discrimination learning experiment. In the preparatory phase of his experiment, grain was placed on two pieces of paper of different shades of gray.

When the hens pecked at grain on the darker of the two papers, they were allowed to eat it; when they pecked at grain on the lighter gray paper, they were shooed away. After several hundred trials, the hens pecked quite consistently at the darker paper, only rarely pecking at the lighter one.

Then came the test of the hypothesis. Grain was once again spread on papers of two shades of gray. But in this test, the stimuli were "transposed;" the previously rewarded "dark gray" paper was now the lighter of two papers, being paired with a paper of an even darker gray.



The question was whether the hens would now peck the paper of the same gray they had learned to peck in the preliminary training (as Thorndike's hypothesis would predict), or whether they would peck now at the darker paper, supporting Kohler's hypothesis that the hens had learned to respond to the darker of two stimuli. In most cases, the hens pecked at the darker paper, supporting Kohler's hypothesis.

Kohler was able to devise a test situation in which his theory and Thorndike's led to different predictions. Philosophers of science (Hempel, 1966) call such tests *crucial experiments:*

> In a *crucial experiment,* a test situation is envisioned in which two well established hypotheses derived from two competing theories predict mutually exclusive outcomes. The single test then simultaneously provides support for one of the hypotheses and lack of support for the other.

We can use Kohler's experiment, considered a classic experiment on learning, to illustrate the decision making that goes into devising a test of the hypothesis. Decisions have to be made about subjects, the research design, as well as the study's apparatus, procedures, and measures.

Finding just the right test of a hypothesis requires ingenuity. Researchers must anticipate that other scientists, working with other theories, will scrutinize every aspect of the study and be all too ready to suggest alternative explanations for its results. Every effort must be made, therefore, to set up the test situation so that if the hypothesis is true, the evidence supporting it will be unambiguous.

1.3.4.1 Subjects.

Researchers must decide whether to test animal or human subjects, what age the subjects should be, and how many should be studied. Information on the criteria used to select subjects should be detailed in research reports. Although only certain subjects can be tested in the study, researchers hope to draw conclusions that will *generalize* (that is, apply) to animals or people other than those used in the research.

> *"Appropriate identification of research participants is critical to the science and practice of psychology, particularly for generalizing the findings, making comparisons across replications, and using the evidence in research syntheses and second data analyses." (Publication Manual of the American Psychological Association, 2010, p. 29)*

Kohler used apes in much of his research. He most likely tested chickens in the "transposition" experiment because Thorndike had formulated his stimulus-response hypothesis observing these birds. By testing chickens himself, Kohler was able to show that his results, showing more complex learning processes than Thorndike's, applied to the very animals that Thorndike studied.

In addition to the question of generalization, subject selection depends on the hypothesis being tested. Hypotheses about children require the researcher to test children; hypotheses predicting differences in the behaviors of men and women require testing these groups.

In later chapters, we present a variety of research designs that offers researchers the option of studying individuals or groups of subjects. We also discuss strategies for deciding how many subjects to test.

1.3.4.2 Research design.

Researchers also must decide on a *research design*.

> A *research design* is a general strategy for collecting observations in research.

There are two general types of designs to choose from: *experiments* and *observational studies* (also called *passive-observational studies)*. The main difference between them is the amount of control the researcher can exercise over the test conditions.

> In an *experiment,* the researcher manipulates the test situation so as to create the precise conditions required for testing the hypothesis. Potential rival explanations of the results are eliminated by holding conditions associated with them constant.

It was possible for Kohler to devise an experimental test of his hypothesis, but many questions of interest to psychologists cannot be studied experimentally. Sometimes the concepts referred to in the hypothesis cannot be created (we cannot change a person's body build, temperament, or birth order); sometimes ethical considerations prevent it (it is not ethically permissible to cause phobias or depression in people, for example).

The researcher can do an observational study when the conditions required for testing the hypothesis cannot be created or when the research question or hypothesis demands it.

> In a *passive-observational study* the hypothesis must be tested nonexperimentally, by seeking out, or waiting for, cases where the specified conditions are realized by nature, and then checking whether [the event] does indeed occur. (Hempel, 1966, p. 20)

Observational studies are common in psychology. For example, Piaget (1954) used this type of design to establish the stages of intellectual development in children.

In designing a study, researchers must anticipate what will happen in the cycle of validation, studying the planned research to see whether holes could be poked in the evidence it will yield. They must ask, for example, whether there are events or conditions, other than those specified by the hypothesis, that vary in such a way that a critic might argue that these, rather than the hypothesized events or conditions, led to the results.

> To the extent possible, potential rival explanations of the results should be eliminated by holding the conditions associated with them constant. This strategy is called *controlling for* rival hypotheses.

The more familiar researchers are with the conditions or events that influence the behavior under study, the greater the likelihood that potential rival explanations will be identified and eliminated by appropriate controls.

Kohler anticipated that if he did not vary the positions of the gray papers that the chickens pecked, his results would be open to a rival explanation. A critic might argue that the chickens had learned to peck at a particular location, rather than at the darker of two stimuli. This argument would be tenable, for example, if the darker paper was always on the right or left side. To control for position, Kohler kept the percent of trials on which the darker paper was in the two positions equal throughout the learning trials. By this means, he eliminated position as a potential rival explanation of his results.

Researchers do not create the test conditions in an observational study, but select for them or wait for them to occur. Because events in nature occur in combinations, it is not always possible to control for potential rival explanations of results in observational studies. For this reason, the conclusions drawn from such studies often are less clear-cut than those from experiments. Whenever feasible, therefore, experiments are preferred over observational studies for testing hypotheses.

1.3.4.3 Apparatus, procedures, and measures.

Decisions also must be made on the particulars of the test situation—where to test subjects, with what apparatus, instructions, manipulations, and measures. Such choices will depend on the hypothesis being tested and on the available technology and traditions in a particular area of research. You can increase your likelihood of finding just the right situation for testing a hypothesis by reading the published studies on the problem.

To test a hypothesis, its theoretical terms must be translated into specific procedures. For some concepts, this will mean manipulating conditions or events in a particular way (like Kohler did); for others, it will mean finding the right measuring technique. Developing precise procedures for manipulating and measuring theoretical concepts in research is called *operationally defining* them.

> *Operational definitions are part of a movement in science called operationism. Operationism is the demand that all theoretical terms in science—that is those that do not refer to something directly observable—be given operational definitions (Leahey, 1994, p. 519)*

The term *operational definition* was coined by the physicist Percy Bridgeman in 1927 and introduced to psychology in 1935 by S. S. Stevens, a psychologist whose work we discuss in Chapter 4, Measurement.

The goal in operationally defining concepts is to specify them with sufficient precision so that others trained in the field can understand and use them in their own research. In this way, it is hoped that subjective, individual understandings of concepts will be eliminated and objective public procedures and measures put in their place. When precise operational definitions are used, scientists can think clearly about problems, and other researchers can understand and replicate the defined features of studies.

To illustrate, Skinner (1938) operationally defined "hunger" in his experimental subjects by varying the number of hours of food deprivation they experienced prior to testing. His rats were allowed "to feed freely once a day for a definite length of time. . . . After about a week of this procedure a high and essentially constant degree of hunger is reached each day just before the time of feeding" (Skinner, 1938, p. 56). With such an operational definition, there is little misunderstanding of what is meant by hunger.

To take another example, intelligence might be operationally defined as the score a person achieves on a standard intelligence test, like the Stanford-Binet or Wechsler Intelligence Test; a different operational definition of intelligence would be needed to study intelligence in newborns or chimpanzees.

The researcher tries to create a manipulation that is self-evident (like Skinner's), or, if the events must be measured rather than created, the best measure possible. When the phenomena of interest can be observed directly (for example, which paper the hens pecked in Kohler's experiment), and when physical measurements are used, this is relatively easy. In such cases, demonstrating the adequacy of the measure is a matter of showing that the instruments are accurate.

The clock used to assess the number of hours of food deprivation would be certified to be accurate to + or - a certain number of seconds per year. Measurements of height or weight would be made with instruments calibrated to match standard measures kept at the Bureau of Standards in Washington, DC. Within the limits of measurement error of the instrument, which normally would be small, height as measured would correspond to height as understood by others.

Researchers face a more difficult task in establishing the accuracy of psychological measures. Investigators who purchase an intelligence test, for example, would find no mention of its accuracy in the test manual, discovering instead a discussion of the test's *reliability* and *validity*.

> The reliability of a measuring instrument is a numerical index of the extent to which it yields consistent results from one occasion to the next.

To quote Anne Anastasi, a specialist in measurement:

> *Reliability refers to the consistency of scores obtained by the same persons when reexamined with the same test on different occasions, or with different sets of equivalent items, or under other variable examining conditions. (Anastasi, 1988, p. 109)*

A metal ruler would yield a highly reliable measure of a person's height, since repeated measurement would yield the same results. A less reliable measure would fail to produce comparable results with replication. Adequate reliability is a minimum requirement for a measuring instrument.

Intelligence tests are highly reliable but there is concern about their validity. The *validity* of a measure has to do with whether it measures what it is supposed to measure. Again, quoting from Anastasi:

> *The validity of a test concerns **what** the test measures and **how well** it does so. It tells us what can be inferred from test scores. . . . The validity of a test cannot be reported in general terms. No test can be said to have "high" or "low" validity in the abstract. Its validity must be established with reference to the particular use for which the test is being considered. (Anastasi, 1988, p. 139)*

To establish the validity of an intelligence test, its developer might compare how similar the results of the new test are to those obtained using established intelligence tests, or, alternatively, how scores on the new test relate to grade point average or other behaviors thought to be related to intelligence.

Usually a great deal of research must be done to establish the validity of a measuring instrument in psychology. For this reason, whenever possible, researchers in psychology use measures with established validity rather than developing their own measures; the validity of a new test must be documented before it is considered acceptable for research. The reliability and validity of psychological measures are discussed more fully in Chapter 12, Planning the Study.

1.3.5 Drawing Conclusions

Once a research design and the details of procedure have been worked out, the planned observations are collected and analyzed. The observations must be recorded in some symbolic form—as words, frequencies of particular behaviors, or numerical scores (e.g., time, scores on a test). Scientists use the term *data* (plural) to refer to the recorded observations that are analyzed to reach conclusions in a study. The term *results* refers to the outcome of the analysis of the data.

Researchers must be concerned with what conclusions to draw from their observations. In research involving hypothesis testing, the question is whether the data support or refute the hypothesis; if the results turn out as predicted, the hypothesis gains credibility. If the results are not in line with predictions, the hypothesis, and ultimately the theory from which it is derived, are rendered less credible. In the words of the famous statistician, R. A. Fisher:

> *The severest test of a theory is to build upon it a system of inferences, for if any rigorously logical inference is found to be untrue the theory fails. If, on the contrary, facts previously unsuspected are inferred from the theory, and found on trial to be true, the theory is undoubtedly strengthened. (Fisher, in Box, 1978, p. 216)*

The credibility of a hypothesis is increased with each successful test, especially when the same result is achieved by many researchers testing the hypothesis under different circumstances. The credibility of a hypothesis, however, does not depend only on confirming predictions; it also depends on how well potential rival hypotheses have been controlled.

Although good researchers take pains to avoid drawing incorrect conclusions, the possibility of error can never be eliminated. Even if the results support the hypothesis consistently, we never can know for sure that future tests will yield the same consistent results. For example, Kohler's study did not resolve the issue of whether animals respond to the absolute properties of stimuli or to the relationship between them. For a discussion of subsequent studies, see Klein (1996). No matter how exacting the methods, scientific conclusions are not without fault.

When the results are uniform, all pointing in the same direction, it is relatively easy to reach conclusions. But such results are rare in psychology. Usually some observations argue for one conclusion, others for the opposite conclusion. Various strategies to help researchers draw conclusions in such cases, including a variety of statistical tests, are considered in later chapters of this book.

1.3.6 Evaluation by the Scientific Community

Scientists form a community bound together by shared knowledge and agreed-upon rules of operation and standards of evaluation. In the second phase of research, *the cycle of validation*, researchers communicate their findings to other scientists through informal discussion, presentations at conferences, and publication in professional journals. To be accepted for presentation at conferences or for publication in journals, research papers must meet the standards of referees and editors qualified to evaluate the work.

Referees judge the importance of the research problem addressed (does the paper add substantively to scientific knowledge?), the appropriateness of its arguments and analyses, and the adequacy of the evidence on which its conclusions are based. In psychology, the subjects, measures and manipulations, and research design would receive careful scrutiny.

Research papers must be written according to an accepted format and give appropriate credit to other scientists. The American Psychological Association's publication manual (2010) specifies the content, organization, and style that is now standard in the field (see Chapter 13, Communicating Research).

Once published or presented at a conference, the research is subjected to a second wave of evaluation. Scientists who attend the conference or read the journal evaluate the research as the referees did; but their judgments may be quite different. They may write papers criticizing the research or conduct studies challenging its hypotheses and testing rival theories. If the findings are validated in this cycle, they slowly become established knowledge in the field.

1.4 A CASE STUDY IN PSYCHOLOGICAL RESEARCH

So far in this chapter, we have discussed the vocabulary and procedures shared by psychologists. We now turn to a case study of research, the story of how a scientific theory and a form of treatment based on that theory were invented and evaluated, to illustrate how these terms are applied.

The experiments that follow, perhaps the first ever performed in clinical psychology, evaluated a popular and controversial theory—that obstructions in the flow of an invisible fluid, "animal magnetism," in people's bodies caused pain and other symptoms.

1.4.1 The Cycle of Discovery: Mesmer and Animal Magnetism

1.4.1.1 A theory invented.

Anton Mesmer, inventor of the theory of animal magnetism, was a wealthy Viennese physician who studied philosophy, theology, and law before settling on a career in medicine. Mesmer introduced the idea of animal magnetism in 1765 in the thesis he submitted to obtain his medical degree.

In this thesis, Mesmer hypothesized the existence of a universal force of nature that penetrated and surrounded all things. In the physical world, Mesmer believed, this force took the form of mineral magnetism, gravity, and electricity; in human beings, it took the form of animal gravity or, as he came to call it, animal magnetism. Animal magnetism was a "subtle fluid," which could not be seen or felt but could be shown to exist only through its effects.

It is not possible to say exactly where Mesmer's theory came from. No doubt its invention involved a complex interplay of induction and deduction. Discoveries in other branches of science certainly set the stage for Mesmer's thinking. To quote John Darnton, a historian:

> *Science had captivated Mesmer's contemporaries by revealing to them that they were surrounded by wonderful, invisible forces: Newton's gravity, made intelligible by Voltaire; Franklin's electricity, popularized by a fad for lightning rods and by demonstrations in the fashionable lyceums and museums of Paris; and the miraculous gases of the Charlieres and Montgolfieres that astonished Europe by lifting man into the air for the first time in 1783. Mesmer's invisible fluid seemed no more miraculous. (Darnton, 1968, p. 10)*

Mesmer was particularly impressed with Newton's theory of gravity, especially his discussion of how the moon's gravitational pull caused the tides. Drawing an analogy between the physical and animal "machines," Mesmer *deduced* that the heavenly bodies influence not only the oceans but human bodies as well:

> *There is almost no change which happens in the heavenly bodies without its influencing the fluids and solids of our earth in agreement. Then, who would deny that the animal machine would, in these circumstances, be agitated to a certain degree by the same causes? The animal is a part of the earth and is composed of fluids and solids, and when the proportion and the equilibrium of these fluids and solids are modified to a certain degree, very perceptible effects will occur from this. (Mesmer, in Bloch, 1980, p. 13)*

Once Mesmer developed the idea of animal magnetism, he went on to explore its importance to health. Knowing that his ideas would carry no weight with the medical faculty without support, Mesmer searched the literature for *evidence* that the human body was affected, like the oceans, by the positions of the heavenly bodies. He was not disappointed.

The medical literature contained many reported cases of pain, fever, hemorrhages, epileptic seizures, madness, and nervous disorders becoming worse during the new and full moon. Mesmer even found evidence of people's faces being disfigured, as though by a tidal pull, during certain phases of the moon:

> *A curious case published by Kerkring is worth mentioning; that of a French woman endowed with a very pretty fat-cheeked face during full moon, but whose eyes, nose, and mouth would turn to one side during the decreasing of the moon. She was then turned so ugly that she could not go out into the world until the full moon returned and she regained the beauty of her face. (Mesmer, in Bloch, 1980, pp. 1516)*

Satisfied that his observations justified the theorized links between the state of the human body and the positions of the sun and moon (the process of *induction)*, Mesmer encouraged his readers to consider the medical applications of these ideas. He promised to dedicate himself to finding out how the medium for such effects, animal magnetism, could be used to restore health and vitality.

1.4.1.2 Mesmer's observational and experimental research.

Some years later, the opportunity arose for Mesmer to test his theory empirically. The *subject* was a Fraulein Oesterline, a 29-year-old woman who suffered from an incredible array of disabling problems:

> *[The patient] had undergone terrible convulsive attacks since the age of two. She had an hysterical fever to which was joined, periodically, persistent vomiting, inflammation of various visceral organs, retention of urine, excessive toothaches, earaches, melancholic deleriums, opisthotonos, lypothymia, blindness, suffocation, and several days of paralysis and other irregularities. (Mesmer, in Bloch, 1980, p. 26)*

Since Mesmer had tried all the standard procedures, to no avail, the possibility of a treatment based on his thesis suggested itself. But Mesmer first wanted to determine whether his patient's symptoms were affected by the positions of the heavenly bodies, as his thesis predicted. This first investigation would have to be an *observational study*, since Mesmer could not control the positions of the heavenly bodies. To test his hypothesis, Mesmer systematically recorded his patient's symptoms and the positions of the sun and moon over several months. As in all *hypothesis testing*, this study involved *deduction;* a hypothesis was used to predict what would occur when specific observations were made.

Mesmer's findings were encouraging. He reported that as the study progressed he began to see regularity in Fraulein Oesterline's symptoms; gradually, by taking account of the positions of the sun and moon, he could foresee his patient's relapses and even predict how long they would last.

His observations led to refinements in his thinking about the causes of illness. He now hypothesized that obstructions in the flow of animal magnetism led to disease. Formulating the *test implications* of this hypothesis, Mesmer reasoned that *if* he placed magnets at strategic points on his patient's body, *then* her symptoms would be washed away. He wrote:

> *Magnetic matter, by virtue of its extreme subtlety and its similarity to nervous fluid, disturbs the movement of the fluid in such a way that it causes all to return to the natural order, which I call the harmony of the nerves. (Mesmer, in Bloch, 1980, p. 29)*

Mesmer applied the magnets to his patient's chest and feet and waited for the *results*. Almost immediately, she reported "a burning and piercing pain" in her body, accompanied by sweating on the side of her body that had been paralyzed. Shortly thereafter, she had a convulsion and was freed of her symptoms. The next day, some symptoms returned, so Mesmer *replicated* the treatment

The magnetic treatment

producing the same breathtaking support for his hypothesis. Although Fraulein Oesterline had several relapses, she finally was cured. She later married Mesmer's stepson and bore several children.

Excited and inspired by his research results, Mesmer conducted other experiments to discover the principles of animal magnetism and the extent of its power. Based on his experiments, Mesmer concluded that magnetic fluid acted very much like electricity. Bottles could be filled with it, just like Leyden jars could be filled with electricity. Like electricity, the fluid could be magnified to produce painful jolts. Anything that could be touched could be magnetized -- paper, bread, wool, stones, glass, porcelain cups, water, dogs, people.

Mesmer also concluded that people could be magnetized from a distance. This observation led him to modify his theory. Although he believed that magnets conducted the fluid, he no longer saw them as necessary for the cure. Instead, Mesmer now thought that magnetic effects were caused by the diffusion of animal magnetism from the magnetist, where it was highly concentrated, to the patient, where it was depleted. The rush of animal magnetism produced the convulsions, or "crises," which produced the cure.

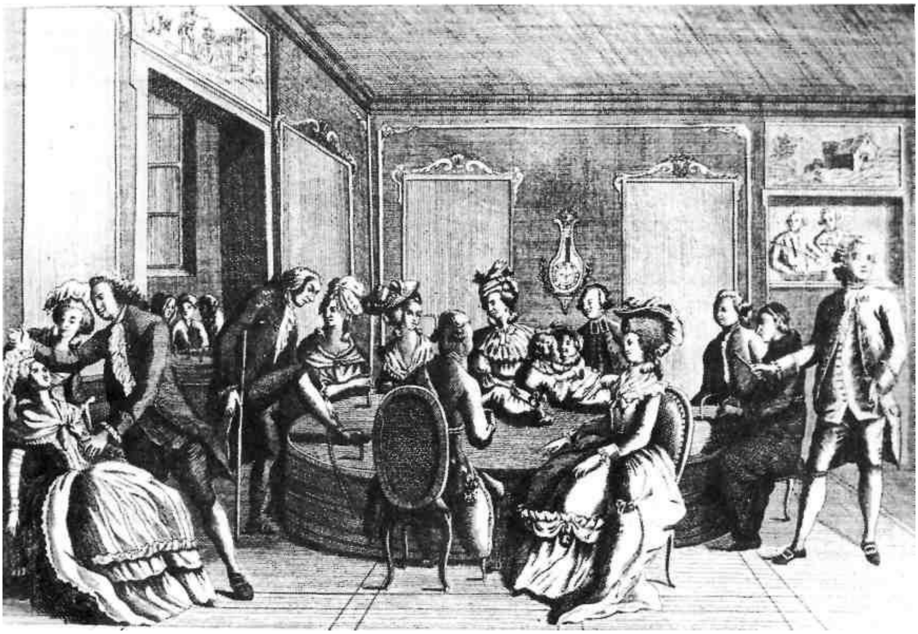
1.4.2 The Cycle of Validation: King Louis XVI's Royal Commission

By 1784, animal magnetism had become the rage in Paris. Parisians talked of little besides magnetism and flocked to the salons where Mesmer "operated" on them. Many great philosophers and scientists saw magnetism as a medical breakthrough, a means of curing illnesses that no one before had treated successfully. To them, animal magnetism offered a wonderful alternative to traditional treatments, which often left patients worse off for having seen a physician than they would have been without.

But there were skeptics as well. Many scientists saw animal magnetism as a figment of Mesmer's overactive imagination and Mesmer as an unfortunate victim of self-deception. Others, less charitable, believed that Mesmer was a charlatan, a fraud, interested less in relieving people of their illnesses than in relieving them of their money.

1.4.2.1 The research problem: assessing animal magnetism. Fortunately, there was a way to settle this difference of opinion. French law required that all new medical treatments be evaluated by the government, in the same way that we require that drugs be tested and approved before they are prescribed. Animal magnetism was to go on trial, and the trial, by special commission, was to decide whether animal magnetism was real and useful.

The Royal Commission, established by King Louis XVI, included France's most prominent physicians and scientists. Benjamin Franklin, ambassador to France from the United States and the world's leading expert on electricity, headed it. J. S. Bailly, the famous astronomer; Antoine Lavoisier, the founder of modern chemistry; and J. I. Guillotin, for whom the "humane" instrument of death used during the French Revolution was named, were among the commissioners.



A public session at Mesmer's salon. Patients, seated around the *baquet,* a wooden tub containing iron filings and magnetized water, use its iron rods and ropes to conduct magnetic fluid to afflicted parts of their bodies. A smaller *baquet* in the back room is for less well-to-do patrons. Mesmer, in the left foreground, assists a patient in crisis (Tatar, 1978).

The commission's report is a most unusual scientific document, which provides a rare glimpse into the step-by-step thinking of a distinguished group of scientists. The report describes how the commissioners gathered facts regarding animal magnetism, formulated hypotheses to explain its workings, decided on research procedures, gathered data, and arrived at conclusions.

1.4.2.2 Hypothesis.

Although the commissioners had been charged with finding out whether Mesmer's magnetic fluid was real and useful, they chose to concentrate on the question of magnetism's reality.

> *The question of its existence is first in order; that of its utility it were idle to examine, till the other shall have been fully resolved. The animal magnetism may indeed exist without being useful, but it cannot be useful if it do not exist. (Report, 1785, p. 29)*

As they made their fact-finding *observations*, the commissioners grew more and more doubtful of the existence of animal magnetism and increasingly confident that animal magnetism's effects resulted from suggestion rather than from an invisible fluid. Mesmer's experimental design, they believed, left open the possibility that it was his patients' faith in the magnetic treatment alone that produced the odd sensations and the relief from symptoms that they experienced.

1.4.2.3 Test implications of the hypothesis.

But deciding on the reality of the magnetic fluid was no easy task. Given the invisibility and intangibleness of the fluid, the commissioners concluded that the only way it could be studied was to follow Mesmer's lead—the treatment had to be studied through its effects. But what sorts of effects should they examine and using what kind of testing situation?

One way to evaluate the magnetic treatment would be to assess its effectiveness in curing patients. To do this, the commissioners would have to administer the treatment repeatedly over an extended period of time. They rejected this option on methodological grounds. Given enough time, they reasoned, nature cured many diseases; so if patients recovered following a course of magnetic treatment, it would be impossible to decide whether their cures resulted from the treatment or from a spontaneous recovery, occasioned only by the passage of time.

Because the commissioners wanted their experiments to be "decisive and unanswerable," they chose to study the immediate sensory and behavioral effects of the magnetic treatment—the pain and other sensations, and especially the convulsions, or crises, which Mesmer claimed brought about the cures.

The commissioners also had to decide who to test. People varied in their susceptibility to magnetism. Some exhibited early and extreme "crises;" others, like the commissioners, experienced nothing. Since the commissioners were interested in explaining the cause of these convulsions, they decided to test patients who were known to respond to the magnetic treatment by exhibiting them.

When magnetism was put on trial in France, Mesmer wanted no part in it. Instead, his most celebrated student and protege, Charles Deslon, opened his salon so that the commissioners could learn firsthand what went on in the public sessions that had become so popular in Paris's high society. The commissioners decided against using the public sessions as a setting for their research, because too many events were happening all at once to allow for accurate observations. They reasoned that if their research was to be decisive, they would have to isolate subjects and create precisely the right conditions for testing their hypothesis. They therefore chose *experiments* over *observational studies*.

The designs that the commissioners selected were brilliant, anticipating alternative ways of interpreting and explaining results and eliminating them as possibilities by means of *experimental controls*. They hoped to design a series of *crucial experiments* that would enable them to decide conclusively between two *hypotheses:* their own hypothesis, that the effects of the magnetic treatment were due to suggestion, and Mesmer's hypothesis, that the effects were caused by the invisible fluid.

Their first step was to translate their hypothesis into *test implications*, as follows: *If* magnetism is applied without a subject's awareness, *then* there will be no effects; conversely, *if* a subject believes that magnetism is being applied when it is not, *then* there will be effects. Mesmer's hypothesis would predict the opposite results.

The commissioners conducted many experiments using different subjects. Although the procedural details varied, their research design was always the same— individual subjects were given several different treatments, and their responses to them were compared. Sometimes the magnetic treatment was given without the subject's awareness; sometimes the subject was led to believe that magnetism was being applied when it was not. The commissioners recorded the presence or absence of convulsions under each condition.

We will look at only two of the many experiments they conducted.

Experiment 1

The first experiment took place in two rooms separated by a doorway that was covered by paper. The subject, a seamstress, and Deslon's patient, was led to one of the rooms where she joined a commissioner and a woman who supposedly wanted some sewing done for her. Both people were seated already when the patient arrived, leaving only one chair, located right in front of the doorway, for her.

Once seated and involved in conversation, a different commissioner then magnetized her through the paper for 30 minutes from a distance of 18 inches. Since Mesmer claimed that magnetic fluid could pass through doors and walls, the paper over the doorway would be no obstacle to its flow. The patient, unaware that she was being magnetized, experienced no special effects; on the contrary, she appeared cheerful and reported that she felt fine. The magnetic treatment produced no effects under this condition.

Next, the commissioner who had just magnetized the patient through the door came in and asked her if she would agree to be magnetized. Once she agreed, he did so, again from 18 inches, the same distance as in the previous test; but there was one additional wrinkle. Although the subject thought that she was being magnetized in the standard fashion, the commissioner actually performed the magnetism in a way that Mesmer's theory would predict would lead to no effects. Nevertheless, the patient went into convulsions, once again supporting the commissioners' psychological hypothesis.

Experiment 2

The second experiment, done with a different subject, involved some porcelain cups. It was well known that convulsions resulted if people who were susceptible to magnetism came close to or touched magnetized porcelain cups. In the experiment, several cups, none of which actually was magnetized, were presented to the subject. The first cup produced no unusual sensations, but the second cup did; by the time the patient touched the fourth cup, she went into full-blown convulsions. Since none of the cups was magnetized, these results supported the hypothesis that the convulsions were due to suggestion; magnetism couldn't have produced the convulsions because they occurred in its absence.

Once the convulsions stopped, the patient asked for water, giving the commissioners an opportunity to test for the effects of magnetism in the absence of suggestion. Without her awareness, the commissioners put the water in a magnetized porcelain cup. According to Mesmer's theory, the patient should have had convulsions, or some other unusual symptoms, as she held the cup and drank. But just the opposite occurred; contrary to the theory, the water appeared to soothe her, supporting the hypothesis that magnetism, without suggestion, had no effects.

1.4.2.4 Drawing conclusions.

The commissioners believed that their experiments were conclusive. In every replication, and in every separate condition of their experiments, the results supported the commissioners' explanation and failed to support Mesmer's hypothesis. The commissioners concluded:

> *Having demonstrated by decisive experiments, that the imagination without the magnetism produces convulsions, and that the magnetism without the imagination produces nothing; they [the commissioners] have concluded with an unanimous voice respecting the existence and the utility of the magnetism, that the existence of the fluid is absolutely destitute of proof. (Report, 1785, pp. 105-6)*

1.4.2.5 Evaluation by the scientific community.

Despite the attempts of the magnetists to save their theory by offering alternative explanations of the commission's findings, the commissioners' report was the beginning of the end of animal magnetism. The popularity of the treatment waned after the report and its use all but died out during the French Revolution.

The research design used by the commissioners, in which a single subject is exposed consecutively to different treatments, became standard practice in psychology until the early 20th century, when it was supplemented by other research designs. In addition, the commissioners laid the groundwork for a methodologically important procedure, the *placebo,* that still is used in modern medical and psychological experiments.

> In a *placebo treatment,* subjects are given a treatment that appears the same as the experimental treatment but lacks its "active" ingredient.

Without this methodological advance, like Mesmer, we would be unable to sort out the effects of suggestion, or faith in the treatment, from the effects of the treatment itself.

Political cartoon showing magnetists fleeing at the sight of the Royal Commission's report. Benjamin Franklin is holding the report. The magnetists are shown as asses, the symbol of the quack.

By demonstrating the powerful effects of suggestion, the commission's investigation of animal magnetism also ultimately led to psychological explanations of nervous disorders. We now believe that many of Mesmer's patients suffered from conversion disorder, a psychological condition characterized by dramatic bodily symptoms with no anatomical basis (paralysis without injury, blindness with no damage to the eye).

Although Mesmer's theory was discredited, he deserves an important place in the history of psychology. Mesmer's bold theory helped make sense, in a logical way, of many medical "facts" of his day that were not accounted for by any other theories. In addition, the theory he invented accomplished dramatic cures of symptoms that had resisted other forms of treatment. Mesmer's work set the stage for revolutionary changes in the treatment of psychological problems, first by suggestion and hypnosis and eventually with modern psychotherapies.

Many people would conclude that Mesmer's theory and research deserve only to be forgotten, and as quickly as possible. But in the introduction to their report, the thoughtful men of science who made up the commission disagreed, offering the following compelling argument why ideas like Mesmer's warrant our continuing attention:

> *Perhaps the history of the errors of mankind, all things considered, is more valuable and interesting than that of their discoveries. Truth is uniform and narrow; it constantly exists, and does not seem to require so much an active energy, as a passive aptitude of soul in order to encounter it. But error is endlessly diversified; it has no reality, but is the pure and simple creation of the mind that invents it. In this field the soul has room enough to expand herself, to display all her boundless faculties, and all her beautiful and interesting extravagancies and absurdities. (Report, 1785, pp. xvii-xviii)*

1.5 Key Terms

Evidence

Empirical sciences

The Invisible College

Replicated results

Cycle of discovery vs. cycle of validation

Induction vs. deduction

Research problem

Hypothesis

Hypothesis testing research

Test implications of a hypothesis

Crucial experiment

Subjects

Generalizable results

Research design

Experiments vs. observational studies (passive-observational studies)

Controlling for rival hypotheses

Apparatus, manipulations, and measures

Operational definition

Operationism

Reliability

Validity

Data, results, observations

Placebo treatment

1.6 Key People

Antoinette and John Lilly

Jean Piaget

George Kneller

Charles Darwin

B. F. Skinner

Ivan Pavlov

Wolfgang Köhler

Edward Thorndike

Percy Bridgeman

S. S. Stevens

Anne Anastasi

Anton Mesmer

Benjamin Franklin

1.7 REVIEW QUESTIONS

1. Identify and discuss four distinguishing features of science.

2. Why is it important to replicate the results of a study?

3. Identify and discuss the two phases of scientific work that Kneller distinguished.

4. What general steps are involved in scientific research?

5. What did Hempel see as the role of theory in research?

6. Why did Skinner think that his research was not guided by theory?

7. Distinguish between induction and deduction, the two processes of reasoning used in scientific work.

8. How do philosophers and historians of science explain why newcomers to a field often make important innovations?

9. What is the difference between the hypothesis and the test implications of the hypothesis?

10. Describe Köhler's experiment testing his Gestalt hypothesis of how animals learn against Thorndike's stimulus-response hypothesis.

11. Distinguish between experiments and passive-observational studies.

12. How did Skinner operationally define hunger in his experiments with rats?

13. Distinguish between the reliability and the validity of a measure.

14. What did R. A. Fisher think was the severest test of a theory?

15. What is the purpose of publishing research in scientific journals that require articles to be evaluated by referees and editors?

16. Describe the evidence Mesmer used in developing his theory of animal magnetism.

17. Describe one of the experiments that Franklin's commission did to evaluate Mesmer's theory.

18. What happened to Mesmer's animal magnetic cure for illness in the years following the commission's report?