Chapter 1 EDUCATIONAL GOALS AND THE NATURE OF SCIENTIFIC INQUIRY

> There was no doubt that atoms could explain some puzzling phenomena. But in truth they were merely one man's daydreams. Atoms, if they really existed, were far too small to be perceived directly by the senses. How then would it ever be possible to establish their reality? Fortunately, there was a way. The trick was to assume that atoms existed, then deduce a logical consequence of this assumption for the everyday world. If the consequence matched reality, then the idea of atoms was given a boost. If not, then it was time to look for a better idea. (Chown, 2001, p. 6)

Classroom inquiry begins when students explore nature and make puzzling observations. Why do some liquids change colors when mixed?

Source: Courtesy of the Research Institute in Mathematics and Science Education, Arizona State University.

Is this how people learn? Does learning start with daydreams? Does it involve tricks, assumptions, deductions, and so on? And if so, do teaching methods follow? The goal of this book is to find out. To do so we will first explore how scientists learn. We will then explore how students learn. We will then apply what we discover to developing teaching methods that enable students to learn not only science but also how to learn. After all, great teachers don't just hand students a few fish. They teach them how to fish.

APPLICABLE NSES STANDARDS

Standard A Science teachers plan an inquiry-based program. In doing so, they

- Develop a framework of yearlong and short term goals for students.

EXPLORING INSTRUCTIONAL ALTERNATIVES

Let's start with a little exploration. Someday you will be making decisions about how to teach certain science topics. Suppose, for example, you are a general science teacher and next week's topic is density. How will you go about introducing the topic? Rank the following alternatives in terms of how you perceive their effectiveness. Use 1 for the most effective and 5 for the least effective and be prepared to discuss your ranking and reasons with your classmates.

(a) Present a video in which 1-dm^3 blocks of various solid materials are carefully weighed and the volumes of 1-kg blocks are calculated from their dimensions, thus allowing comparison of two density determinations of each material.

(b) Arrange for a laboratory period in which students use rulers, calipers, graduated cylinders, and balances to determine the volumes and masses of solids of widely differing shapes and various materials for plotting on graphs of volume versus mass.

(c) Discuss with students their experiences with floating and sinking objects, including themselves when they swim or play in the water.

(d) Present an explanation and demonstrate ways in which various materials are weighed, their volumes are found by appropriate means, and you finally calculate the density of each material.

(e) Arrange for a laboratory period in which students accurately measure the density of carefully machined blocks and rods of materials whose volumes can be calculated easily from linear measurements.

Regardless of which approach you thought best, to make good instructional decisions you will need to carefully consider the broader goals of education. Why, for instance, are we teaching science in the first place? What do we really want our students to learn and be able to do as a consequence of our instruction?

THE GOALS OF AMERICAN EDUCATION

Freedom of Mind and the Ability to Reason

The acquisition of knowledge, values, attitudes, thinking processes, and creativity has been seen as a worthy educational objectives. But can a common thread be pulled from this diversity? In 1961 the Educational Policies Commission published *The Central Purpose of American Education*. In the commission's view, something it called "freedom of the mind" was of primary importance:

> Freedom of the mind is a condition which each individual must develop for himself. To be free, a man must be capable of basing his choices and actions on understandings which he himself achieves and on values which he examines for himself. He must be aware of the basis on which he accepts propositions as true. (pp. 3–4)

In other words, freedom of the mind requires understanding oneself, one's actions, and one's surroundings. How does one come to understand this? The commission's answer was contained in what it called the 10 rational powers—the powers of recalling, imagining, classifying, generalizing, comparing, evaluating, analyzing, synthesizing, deducing, and inferring. In the commission's view, without rational powers, people must accept the ideas, beliefs, and attitudes of others. In short, the commission argued that all of a school's varied objectives depend on the ability to think. But how can instruction be designed and implemented to help students become better thinkers?

Scientific Habits of Mind

In 1966 the Educational Policies Commission, recognizing the key role that science could play in developing the ability to think, published *Education and the Spirit of Science*. That document emphasized science as a way of thinking, a spirit of inquiry driven by a curiosity to understand nature. Although the commission recognized that no scientist may fully exemplify the spirit of science, it nevertheless identified the following scientific habits of mind:

1. Longing to know and to understand
2. Questioning of all things
3. Search for data and their meaning
4. Demand for verification
5. Respect for logic
6. Consideration of premises
7. Consideration of consequences

These habits of mind insist that students are not indoctrinated to think or act in a certain way. Rather, students must acquire the ability to make up their own minds, to develop freedom of the mind, and to learn to make decisions based upon reason and evidence. In this sense, the scientific values are the most complete expression of one of the deepest human values—the belief in human dignity.

Scientific Literacy

More recently the American Association for the Advancement of Science (AAAS, 1989) echoed the importance of effective thinking in terms of achieving scientific literacy. In its words,

> Scientific habits of mind can help people in every walk of life to deal sensibly with problems that often involve evidence, quantitative considerations, logical arguments,

> and uncertainty; without the ability to think critically and independently, citizens are easy prey to dogmatists, flimflam artists, and purveyors of simple solutions to complex problems. (p. 13)

Regrettably, the AAAS concluded that most Americans are not scientifically literate. To achieve scientific literacy the association advocated teaching science as science is practiced. In other words, teaching should (a) start with questions about nature, (b) engage students actively, (c) concentrate on the collection and use of evidence, (d) not separate knowing from finding out, and (e) de-emphasize the memorization of technical vocabulary.

The *National Science Education Standards*

Inquiry teaching is also prominent in the National Research Council's *National Science Education Standards* (NRC, 1996) (see (pp. xx–xxii)) as well as in its guidelines for science teacher education (NRC, 2001). For example, with regard to teaching methods the 1996 document had this to say:

> The Standards call for more than science as a process, in which students learn such skills as observing, inferring, and experimenting. Inquiry is central to science learning. When engaging in inquiry, students describe objects and events, ask questions, construct explanations, test those explanations against current scientific knowledge, and communicate their ideas to others. They identify their assumptions, use critical and logical thinking, and consider alternative explanations. In this way students actively develop their understanding of science by combining scientific knowledge with reasoning and thinking skills. (p. 2)

In summary, these national organizations are advocating teaching science not only as a way to produce more scientists but also as a way to develop citizens who think creatively and critically.

Teaching for Thinking

Thus, the central instructional question is this: *How can we teach science so that students develop scientific thinking skills—freedom of mind—scientific literacy?* The answer proposed in this book is by using teaching methods that allow students to explore nature and advance and test their own ideas. In other words, teaching should allow students to participate in inquiry—that is, in the knowledge construction process. Importantly, as we shall see, thinking skills play a key role in learning science content. Thus, teaching in ways that help students develop their thinking skills will pay off in terms of better content learning and better performance on high-stakes tests such as the SAT (formerly the Scholastic Aptitude Test), the ACT (formerly American College Testing) tests, and the new statewide science tests that have been mandated by the No Child Left Behind Act. This act, passed by Congress in 2002, mandated that by 2007 all students in Grades 10–12 be tested in science.

Consequently, if you want your students to do well on these tests, it behooves you to learn how to teach in ways that help students develop effective thinking skills. And as we shall see, thinking skills do not develop—nor do they function—independent of content. Therefore, we must carefully consider the relationship between content and process. Indeed, as we shall see, lessons in which students truly investigate nature and construct their own knowledge best facilitate both the learning of science content *and* the development of thinking skills.

In short, we do not want to simply give students a few fish. As mentioned, we want to teach them how to fish. To learn how to do this, however, we need to learn how people fish. Consequently, the remainder of this chapter and the next will explore how scientists fish for new knowledge. You may think that you already know quite a bit about how science is practiced. Experience suggests, however, that most prospective teachers, and even many experienced teachers, misunderstand several aspects of how scientific inquiry works. Hopefully you will find the following examples illuminating—perhaps even surprising. You should also find them helpful in setting the stage for the introduction of inquiry-based teaching methods in subsequent chapters.

HOW SCIENCE IS PRACTICED

Science is often defined as the attempt to explain nature. But how does one do this? Let's tackle this question by traveling to Africa's Serengeti Plain to see how biologists Lue Scheepers and Robert Simmons inquired into why giraffes have long necks.

You have probably heard that giraffes have long necks to feed on leaves in tall trees. This often-repeated explanation even includes the idea that short-necked giraffes can't reach the high-up leaves and starve. So after several generations, only long-necked giraffes have survived. Is this true? If true, where in trees would you expect to find giraffes most often feeding, particularly when food is scarce? Stop reading and generate an answer before reading on.

Now consider the following argument:

If . . . the "feeding-up-high" explanation is correct,

and . . . giraffes are surveyed to find out where they feed,

then . . . they should most often be observed feeding in the upper parts of the trees, particularly during the dry season when food is scarce.

Unfortunately for the feeding-up-high explanation, Scheepers and Simmons observed that giraffes most often feed by bending down (Figure 1.1). Even in the dry season, they spend nearly 40% of their time feeding at relatively low heights. Clearly this isn't the expected result based on the feeding-up-high explanation. Therefore, the explanation is not supported.

Scheepers and Simmons guessed that over several generations, long necks might have been acquired for courtship battles between males. During such battles, two males square off and swing their necks and heads at each other. The male with the longer neck generates

FIGURE 1.1 If the feeding-up-high explanation is correct, what should we see? Is that what we do see? What does this mismatch tell us about the explanation?

more force. So he wins and gets to mate with the female, while the shorter-necked loser is left to die. How might we test this "male-battle" explanation? Consider this argument:

> *If* . . . the feeding-up-high explanation is correct,
>
> *and* . . . the neck lengths of males and females are compared,
>
> *then* . . . both males and females should have equally long necks relative to their body size.

Alternatively:

> *If* . . . the male-battle explanation is correct,
>
> *then* . . . males should have relatively longer necks than females.

When analyzing their data, Simmons and Scheepers found that males do in fact have relatively longer necks than females. Because this is the expected result based on the male-battle explanation, that explanation was supported. Simmons and Scheepers also found that males have relatively larger and heavier heads than females. Although this result was not previously predicted, it does make sense in light of the male-battle explanation because a heavier head, like a longer neck, increases the impact force. Clearly, additional tests would be helpful. Nevertheless, the important point is that scientists use *If/and/then* reasoning to test, and perhaps reject, their proposed explanations.

Notice that the part of the *If/and/then* argument after each *If* is a proposed explanation where the word *explanation* means to make clear the cause or reason of—to account for. Thus, an *explanation* is a tentative *cause* for a puzzling observation. Scientists often refer to proposed explanations as **hypotheses**. The term *hypothesis* sometimes is also used to label tentative descriptive statements—that is, answers to descriptive "who," "what," "when," and "where" questions. However, in this present context, hypotheses will generally refer to possible causes—possible explanations—possible answers to a causal question (e.g., Why do giraffes have long necks?).

The part of *If/and/then* arguments that follows the *and* is sometimes called the **planned test**—or imagined test. In the giraffe example the planned test involved observing giraffes to see where in the trees they most often feed. The part after each *then* is an **expected result** (sometimes called a **prediction**). Expected results are statements of what *should* happen in the future, what one *should* observe in the future, *if* the hypothesis is correct and *if* the test is conducted as planned. In other words, expected results state what reasonably follows from the hypothesis and its planned test.

To make sure that this last point is clear, note that the term *expectation* (or *prediction*) has at least two meanings. For example, suppose during the past few months you have taken 20 math quizzes and earned an A on each one. What grade do you *expect* you will earn on your next math quiz? Based on the past, you could reasonably *expect/predict* that you will earn another A. The statement that the next grade will be an A is an expectation/prediction. But in this case deriving this expectation involves extending (extrapolating) a pattern of past events into the future. This is *not* the sort of expectation involved in hypothesis testing. During hypothesis testing, expectations/predictions are derived not by extrapolation but by a thought process called **deduction**. For example, if giraffes have long necks to feed in tall trees, then it follows (via deduction) that they should feed most often in the upper parts of trees, particularly when food is scarce. That is, it deductively follows that they should feed up high as opposed to the notion that they should feed up high because that is where they have been feeding during the past 20 days. The statement that they should feed up high is an expectation just like the statement that you will earn an A on your next math quiz. However, in one case the expectation comes from extrapolation, and in the other case it comes from deduction.

When we add the observed result and the conclusion to one of Scheepers and Simmons's *If/and/then* arguments, we get something that looks like this:

> *If* . . . the feeding-up-high hypothesis is correct,
>
> *and* . . . giraffes are surveyed to find out where in trees they most often feed (planned test),
>
> *then* . . . they should spend most of their time feeding in the upper parts of the trees, particularly during the dry season when food is scarce (expected result).
>
> *But* . . . giraffes spend most of their time feeding down low, even in the dry season (observed result).
>
> *Therefore* . . . the hypothesis is contradicted (conclusion).

Notice that the last statement in the argument following the *Therefore* is the **conclusion**. A conclusion is *not* an observed result. Instead, a conclusion tells us about the status of the tested hypothesis. In general, conclusions state whether the hypothesis was supported based on a match between expected and observed results or contradicted based on a mismatch.

In summary, hypothesis testing is a bit odd in the sense that to test a hypothesis you have to suppose, for the time being, that it is correct. You have to do this so you can test it and perhaps find that it is *not* correct!

Avoiding Bias and the Need for Alternative Hypotheses

Let's stay on the Serengeti Plain to take an even closer look at the nature of hypothesis testing. That closer look will consider the potential problem of bias and how to avoid it. We will also consider the source of hypotheses as well as kinds of evidence that can be used in their test.

While watching cheetahs chase gazelles (see Figure 1.2) British biologist Tim Caro noticed that gazelles often leap high in the air with their legs stiffly pointed backward. This puzzling behavior, called *stotting*, led Caro to ask, Why do gazelles stott? Does stotting somehow help them escape? Caro thought that stotting was worth being curious about because it seemed odd that gazelles would jump, slowing themselves down and making themselves more conspicuous while running for their lives. What explanations for stotting can you suggest?

Prior to Caro's research, another biologist proposed an explanation for stotting. He thought that stotting gives gazelles an advantage, despite appearances to the contrary,

Figure 1.2 Why do gazelles, such as this one, stott when being chased by cheetahs?

because stotting enables them to better spot cheetahs. Caro knew about this explanation but was unconvinced. Consequently, he generated several additional explanations. One was that a gazelle stotts to warn other gazelles of danger. Another was that an adult gazelle stotts to draw attention from its more vulnerable offspring. Still another was that stotting signals the cheetah that the gazelle is very fit, thus telling the cheetah that the gazelle will be difficult to catch. Consequently, the cheetah should give up the chase.

So Caro didn't generate one explanation. He generated several. Generating several alternatives at the outset is important because it helps one be more open-minded later on. People who fail to consider alternatives often become **biased**. In other words, they often get stuck believing in an incorrect explanation, particularly if the explanation comes from an authority or obtains initial supporting evidence. Note also that Caro's alternative hypotheses didn't come from direct observation. Instead they came from his prior knowledge of similar situations. For example, he knew that Arctic ground squirrels draw attention to themselves to protect their offspring. So he suspected that gazelles might do the same.

Having used his prior knowledge to generate several hyphotheses, Caro needed to test them. Let's see how he used correlational evidence to test the explanation that stotting signals the cheetah that the gazelle is very fit and hence will be difficult to catch.

The Use of Correlational Evidence

If Caro's "very-fit" hypothesis is correct, what relationship, if any, should exist between stotting and getting caught? Consider this:

> *If* . . . the very-fit hypothesis is correct,
>
> *and* . . . several chases are observed, some in which stotting takes place and some in which it doesn't,
>
> *then* . . . stotting gazelles should get caught less often than nonstotting gazelles. Presumably the stotting gazelles won't get caught as often because the cheetahs have received the message that the gazelle is very fit. So they will give up the chase.

Take a look at Caro's observed numbers in Table 1.1. Do the observed numbers support the hypothesis? If the hypothesis is correct, the nonstotters should get caught more often than the stotters. Did this happen? Of the 24 nonstotters, 5 got caught (21%). Of the seven stotters, none got caught (0%). So it looks like the nonstotters got caught about 21% of the time, while the stotters got caught 0% of the time. The stotters always got away! So stotting seems to help. *Therefore,* the very-fit hypothesis is supported. But caution is certainly in order as the data don't rule out alternatives. For example, perhaps stotting confuses the cheetah. How could this alternative be tested?

This present test involves finding a corelationship or **correlation** between the values of two variables. As you may know, this type of evidence is referred to as **correlational evidence**. When using correlational evidence, one looks to see if the values of the two variables are "linked." If they are linked as predicted, then the hypothesis is supported—but certainly

TABLE 1.1 Does the Evidence Support the Very-Fit Hypothesis? Caro's Observed Results

Stott	*yes*	0	7
	no	5	19
		yes	*no*
		Caught	

not proven (more will be said about this in Chapter 2). Suppose, for example, someone hypothesizes that cell phones cause brain tumors. What correlational evidence could you use to test this hypothesis?

Importantly, finding a correlation between two variables in the absence of a prior hypothesis doesn't tell you which variable is the cause and which is the effect, or if a causal relationship even exists. Fortunately, there is another way to test hypotheses that involves "manipulating" nature. Such manipulations are called **experiments**, and they produce **experimental evidence**. Experimental evidence has an advantage over correlational evidence because one can more clearly establish cause and effect.

TESTING HYPOTHESES USING EXPERIMENTS

Let's travel to the Pacific Northwest where biologists used experiments to test hypotheses about the curious homing behavior of silver salmon. Silver salmon are born in the cool, quiet headwaters of freshwater streams in the Pacific Northwest. Young salmon swim downstream to the Pacific Ocean where they grow and mature sexually. They then return to swim upstream, often jumping incredible heights up waterfalls to ultimately lay eggs or deposit sperm in the streams' headwaters before dying. By tagging young salmon, biologists discovered that mature salmon actually return to precisely the same headwaters where they were born some years earlier! The discovery of this descriptive pattern (what scientists would call the **law** of salmon navigation and reproduction) raised a very interesting causal question: How do salmon find their home streams? In other words, what *causes* them to end up in their home streams? Before reading on, take a few minutes and see how many alternative hypotheses you can come up with.

A number of alternative hypotheses can be proposed. For instance, people often navigate by sight. Perhaps salmon do as well. Returning salmon may recall objects, such as large rocks, they saw while swimming downstream. Studies of migratory animals also suggest possibilities. For example, migratory eels are known to be enormously sensitive to dissolved chemicals. Perhaps salmon are as well. Perhaps they swim a short distance into various streams until they find the one that smells right, and they then follow the chemical path home. Also homing pigeons are known to navigate using the Earth's magnetic field. In one experiment, pigeons wearing little magnets (to disrupt the magnetic field) were not as successful at finding their way home as another group of pigeons wearing nonmagnetic metal bars. Perhaps salmon are also sensitive to the magnetic field and use it to find their home streams. Thus, by borrowing explanations from possibly similar contexts—by using **analogies**—we have generated three alternative hypotheses:

1. Salmon use sight.
2. Salmon smell chemicals in their home streams.
3. Salmon use the Earth's magnetic field.

A key point is that the use of analogies, sometimes called analogical transfer, is a creative process. Thus, when trying to explain something, scientists must be creative. They need to brainstorm and generate several alternatives, even if some may seem silly. Scientists can't be afraid of being wrong because that's the way they learn. They often learn from their "mistakes," from their rejected hypotheses.

What Combinations of Hypotheses Exist?

In addition to the three hypotheses listed above, others remain. Indeed none of the three may be correct. Or perhaps salmon use two of the three methods, or perhaps all three. Generating all possible combinations of the hypotheses gives us these possibilities:

1. The sight hypothesis is correct.
2. The smell hypothesis is correct.
3. The magnetic-field hypothesis is correct.
4. Both the sight and the smell hypotheses are correct.
5. Both the sight and the magnetic-field hypotheses are correct.
6. Both the smell and the magnetic-field hypotheses are correct.
7. All three hypotheses are correct.
8. Not one of the three hypotheses is correct. In other words, some other hypothesis is correct.
9. Some combination of other hypotheses is correct.

What Reasoning Guides Experiments?

Having generated all possible combinations of likely explanations, the next task is to test them. But this time we will test the hypotheses by manipulating nature—that is, by doing experiments. However, before we discuss how experiments were conducted on salmon, let's briefly discuss the reasoning that guides experiments in general.

Suppose I have two golf balls, a Titleist and a Top-Flite, and want to find out which is bouncier. Suppose you tell me to drop them to find out. So I drop the Titleist from over my head and the Top-Flite from my waist. To which you reply, "That is not fair. You have to drop them from the same height." "OK," I say. So I drop them from the same height. But this time I drop the Titleist on cement and the Top-Flite on carpet. To which you reply, "That is still not fair. They both have to hit the cement." So I drop them from the same height onto cement. But before letting go, I spin one and not the other. So once again you tell me that the experiment is not fair. And so on. The point is the experiment must be "fair" in the sense that both balls have to be treated the same way for us to discover which ball is bouncier.

Consider a slightly more complicated situation. Take a look at the pendulums shown in Figure 1.3. When the weights hanging on the ends of the pendulums are released, the pendulum on the right swings back and forth much faster than the one on the left. Why does it swing faster? What possibilities can you suggest? You may have noticed that the right pendulum has a shorter string. It also has a heavier weight. And it was pulled out and released from farther out than the left pendulum. So we have at least three possible reasons (causes) for why the right pendulum swings faster:

1. It might swing faster because of its shorter string.
2. It might swing faster because of its heavier weight.
3. It might swing faster because of its more distant release point.

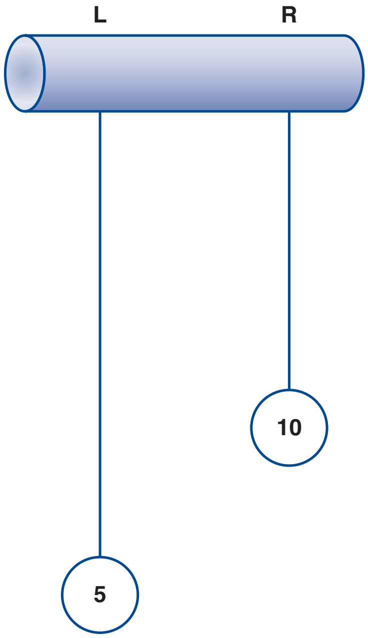


FIGURE 1.3 Diagram of pendulums

Of course, some combination of these three factors (variables) may be the reason. Or perhaps some other variable that we have not thought of may be the reason.

Let's start with hypothesis 1. How could we test the hypothesis that the right pendulum swings faster because it has a shorter string? Let's call it the "string-length" hypothesis. Clearly we would need to swing two pendulums, one with a shorter string and one with a longer string. If we could be certain that the only difference between the two pendulums was string length, and if we do find that the shorter pendulum swings faster, then we could conclude that changing string length causes a difference in swing speed. On the other hand, if we conduct an experiment in which both string length and release point vary and we find a difference in swing speed, then we cannot be sure which variable (string length or release point) caused the difference in swing speed. In other words, we need to conduct an experiment in which we change the values (short vs. long) of one variable (string length) while keeping the values (heavy vs. light weight, near vs. far) of the other variables (weight, release point) the same. Such an experiment is called a **controlled experiment**. So the reasoning guiding the controlled experiment that tests the string-length hypothesis looks like this:

> *If* . . . the string-length hypothesis is correct,
>
> *and* . . . we change string length while keeping the values of all other potentially causal variables constant (planned test),
>
> *then* . . . swing speed should vary (prediction).

In this experiment, the string length varies from one trial to the next. So "string length" is the **independent variable** (sometimes called the *input* or *manipulated variable*, the possible *cause*, the *stimulus*, or the variable x in mathematical expressions) because it is being varied independently of the other possible causal variables. And, according to the hypothesis, swing speed should depend on string length. So "swing speed" is called the **dependent variable**. Other terms used to label the dependent variable include the *outcome variable*, the *effect*, the *responding variable*, or simply the *response* (usually designated the variable y in mathematical expressions). Thus, the point of a controlled experiment is to discover if changing the values of one variable (the independent variable) really does cause a change in the values of another variable (the dependent variable). Or said the other way around, the point of a controlled experiment is to discover if the values of one variable (the dependent variable) really do *depend* on the values of another variable (the independent variable). And in this case they do. Therefore the string-length hypothesis is supported. You might want to actually conduct the experiment yourself to see if you get the same result. And what about the other independent variables—pendulum weight and release point? It turns out that these variables don't make a difference. So the weight hypothesis and the release-point hypothesis are not supported. Instead they are contradicted.

Using Controlled Experiments to Test Hypotheses About Salmon Navigation

Let's now consider how American biologist A. D. Hasler conducted controlled experiments to test the hypotheses about salmon navigation. To test the sight hypothesis, Hasler first captured salmon that had just returned to two streams near Seattle, Washington. The streams, shown in Figure 1.4, were the Issaquah and East Fork. Hasler then tagged the captured fish, identifying which had come from the Issaquah and which from the East Fork.

Next, he randomly split the tagged Issaquah fish into two groups and blindfolded all the fish in one group. He then repeated the procedure for the tagged East Fork fish. The blindfolded Issaquah and East Fork fish became the **experimental group**. Hasler then released the blindfolded salmon along with some nonblindfolded salmon from both streams about three quarters of a mile below where the streams join (marked site of release). The nonblindfolded salmon were the **comparison group**—sometimes called the control group. Finally, the tagged fish were recaptured in traps about a mile above the junction as they swam back up the streams (marked recapture site). The following summarizes Hasler's reasoning:

If . . . the sight hypothesis is correct,

and . . . blindfolded salmon and nonblindfolded salmon from the two streams are released below the fork where the two streams join,

then . . . the blindfolded salmon should *not* be recaptured in their home streams as often as the nonblindfolded salmon.

To establish this possible link between the salmon's ability to see and where they are recaptured, all the other ways that the two groups of fish differ must be the same (i.e., held constant). For example, another variable is the amount of time the fish are kept out of water. Clearly it would not be fair to keep the blindfolded salmon out of water longer than the nonblindfolded

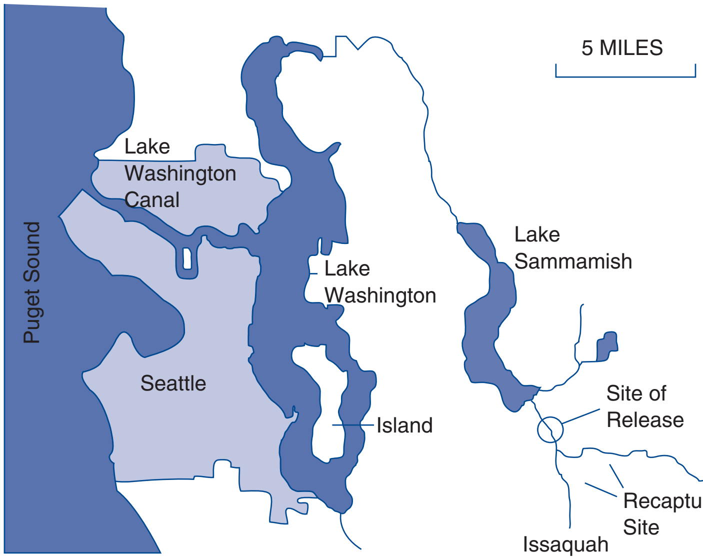


FIGURE 1.4 Issaquah and East Fork streams showing release and recapture sites

salmon because doing so might hurt their homing ability. With this in mind, we can see that a controlled experiment (i.e., a fair test) of the sight hypothesis requires that experimental and comparison groups differ in only one way or be treated differently in only one way.

In Hasler's controlled experiment, the salmon's ability to see varies between the experimental and the control group. So "ability to see" is the independent variable. And, according to the sight hypothesis, where the salmon are recaptured should depend on sight. So "recapture site" is the dependent variable. Thus, as mentioned, the point of a controlled experiment is to discover if changing the values of the independent variable really does *cause* a change in the values of the dependent variable.

Suppose having conducted Hasler's controlled experiment, we discover that the sighted salmon are better at returning home than the blindfolded salmon. This is the predicted result based on the sight hypothesis, so the result would support the hypothesis. However, as mentioned, we need to be careful. Perhaps during the experiment, the blindfolded salmon were hindered in returning, not by lack of sight but by their inability to swim with blindfolds. Or perhaps blindfolding the fish simply shocked them and disrupted their swimming ability. At any rate, one must try to avoid these potential problems (i.e., uncontrolled independent variables) as much as possible. But as you may have guessed, one can never be absolutely certain that all such problems have been eliminated. So caution is needed even when interpreting experimental results.

On the other hand, suppose we conduct the experiment and find that both groups are equally successful at returning home. In this case, the sight hypothesis would be contradicted. In other words, now we can be reasonably sure that salmon navigate some other way. However, again we need to be cautious. Overlooked independent variables might be operating. For example, perhaps the blindfolded salmon could see under their blindfolds. This is how blindfolded magicians see. Or perhaps the blindfolds were not thick enough to block out all the light. Or perhaps the blindfolds were effective and the salmon do use sight when they can, but when they can't, they use some other sense to navigate, such as smell.

In short, the reasoning involved in experimentally testing hypotheses follows the *If/and/then* pattern just like before. But it also involves identifying and attempting to hold constant the values of independent variables. This is crucial. However, because we can never be certain that all of the independent variables have been identified and/or controlled, even after conducting experiments, conclusions must remain somewhat tentative. In sum, scientific arguments and evidence can be convincing *beyond a reasonable doubt* but not *beyond all possible doubt*. As mentioned, we will return to this point in Chapter 2.

As it turned out, when Hasler conducted his experiment, he found that the blindfolded salmon were just as successful as the nonblindfolded salmon at finding their home streams. Therefore, the sight hypothesis was contradicted. So Hasler moved on to test the smell hypothesis.

Testing the Smell Hypothesis

To test the smell hypothesis, Hasler captured and tagged salmon from the two streams and randomly divided the Issaquah fish into two groups. He inserted cotton plugs coated with petroleum jelly in the noses of one group to block their smelling ability (Figure 1.5). These fish became the experimental group. He left the noses of the comparison/control group unplugged. Hasler then randomly split the East Fork fish into two groups and plugged the noses of one group as he had done with the East Fork fish. Finally, he released all the fish to the release point. As the fish returned upstream, they were recaptured in traps above the streams' junction.

Tables 1.2 and 1.3 show Hasler's results. Do they support the smell hypothesis? If so, why? If not, why not? Of course other hypotheses remain. But let's save these for another day.

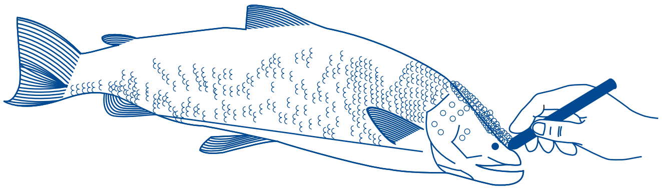


Figure 1.5 Will plugging a salmon's nose keep it from finding its home stream?

Table 1.2 Observed Results for Comparison Fish With Unplugged Noses

	Recapture Site	
	Issaquah	*East Fork*
Initial Capture Site		
Issaquah	46	0
East Fork	8	19

Table 1.3 Observed Results for Experimental Fish With Plugged Noses

	Recapture Site	
	Issaquah	*East Fork*
Initial Capture Site		
Issaquah	39	12
East Fork	16	3

BASIC AND APPLIED RESEARCH

Why Do People Test Alternative Explanations?

Why did the scientists go to all the trouble to test alternative hypotheses about long-necked giraffes, stotting gazelles, and navigating salmon? One answer is simply that they were curious. Curiosity often motivates science. The intent is to better understand nature. Such research is called **basic research**.

Sometimes, however, hypotheses are tested with practical purposes in mind. Suppose, for example, that your car is making a strange clicking sound. So you take it to an auto mechanic. The mechanic's task is to discover the cause. To do so, one must generate and test alternative hypotheses. Also, consider what happens when you go to a medical doctor with an illness. The doctor's first task is to answer the **causal question:** What is causing your symptoms? To find the cause, the doctor must generate alternative hypotheses and test them as quickly as possible so that you can be treated and get well. In some cases this may be relatively simple; for example, if the cause is a bacterial infection and antibiotics are prescribed and taken, then you should get well. And after taking the antibiotics, you do get well. Therefore the bacterial infection hypothesis is supported, and everyone is happy. On the other hand, if you do not get well, then perhaps a bacterial infection was not the cause, and the doctor will have to generate and test other hypotheses until you are cured. Medical researchers have generated and tested hypotheses in hopes of finding the causes and possible cures of diseases such as typhoid, syphilis, tuberculosis, polio, cancer, and more recently AIDS. The intent of hypothesis testing in all of these cases from the auto mechanic to the medical researcher is to find a cause so that some practical action can be taken. Hypothesis testing for practical reasons is called **applied research**.

Whether driven by curiosity or practicality, the central issue is finding causes. Generating and testing hypotheses to answer causal questions is central to nearly every aspect of adult life. Unfortunately, however, we seldom have eyewitness accounts to help establish cause-effect relationships. And even when we do, eyewitness accounts are often unreliable. It turns out that people often see what they expect to see, rather than what really happened. You may not have thought of a murder trial as an instance of hypothesis generation and test, but that is exactly what it is. The central causal question is, Who killed the victim (i.e., who caused the death)? If reliable eyewitness testimony is unavailable, the prosecutor's job is to use **circumstantial evidence** and sound reasoning to convince the jury that the "defendant-is-the-killer" hypothesis is correct. On the other hand, the defense's job is to generate and support alternative hypotheses, such as drug dealers killed the victim or a bolt of lightning killed the victim. To get an acquittal, all the defense has to do is get one juror to accept the plausibility of only one of these alternative hypotheses. Doing so would mean that the defendant-is-the-killer hypothesis, in that juror's mind, has not been *proven beyond a reasonable doubt.*

Summary

- The central goal of American education is to teach students how to think. To do so, science must be taught as a process of creative and critical inquiry.
- Basic to science is the generation and test of explanations. The initial generation of several alternative explanations encourages an unbiased test because one is less likely to be committed to any specific explanation. Explanations are tested by supposing that the explanation is correct and by planning some test that allows the deduction of one or more predictions. Data are then gathered and compared with predictions. A good match provides support for the explanation, while a poor match contradicts the explanation and may lead to its rejection.
- Explanations can be tested using circumstantial, correlational, or experimental evidence.
- Although both tentative descriptions and explanations are tested by use of *If/and/then* reasoning, descriptions and explanations are not the same thing. Descriptions (sometimes called laws) tell us about nature in terms of identifiable patterns, while explanations attempt to identify causes for such patterns.
- Generating hypotheses is a creative process based on perceived similarities (analogies) between the present situation and past situations.
- Use of experimental evidence relies on the manipulation of causal variables and is an effective way to test hypotheses because the experimenter can perform controlled experiments.
- People test hypotheses to find out why things happen, to find which of the possible causes is/are the actual cause(s). People want to know the actual causes of things to satisfy their curiosity—basic research—or so that the new knowledge can be put to practical use—applied research.
- Inquiry is open ended as answering one causal question often raises another.

Key Terms

analogies

applied research

basic research

biased

causal question

circumstantial evidence

comparison group

conclusion

controlled experiment

correlation

correlational evidence

deduction

dependent variable

expected result

experimental evidence

experimental group

experiments

hypotheses

independent variable

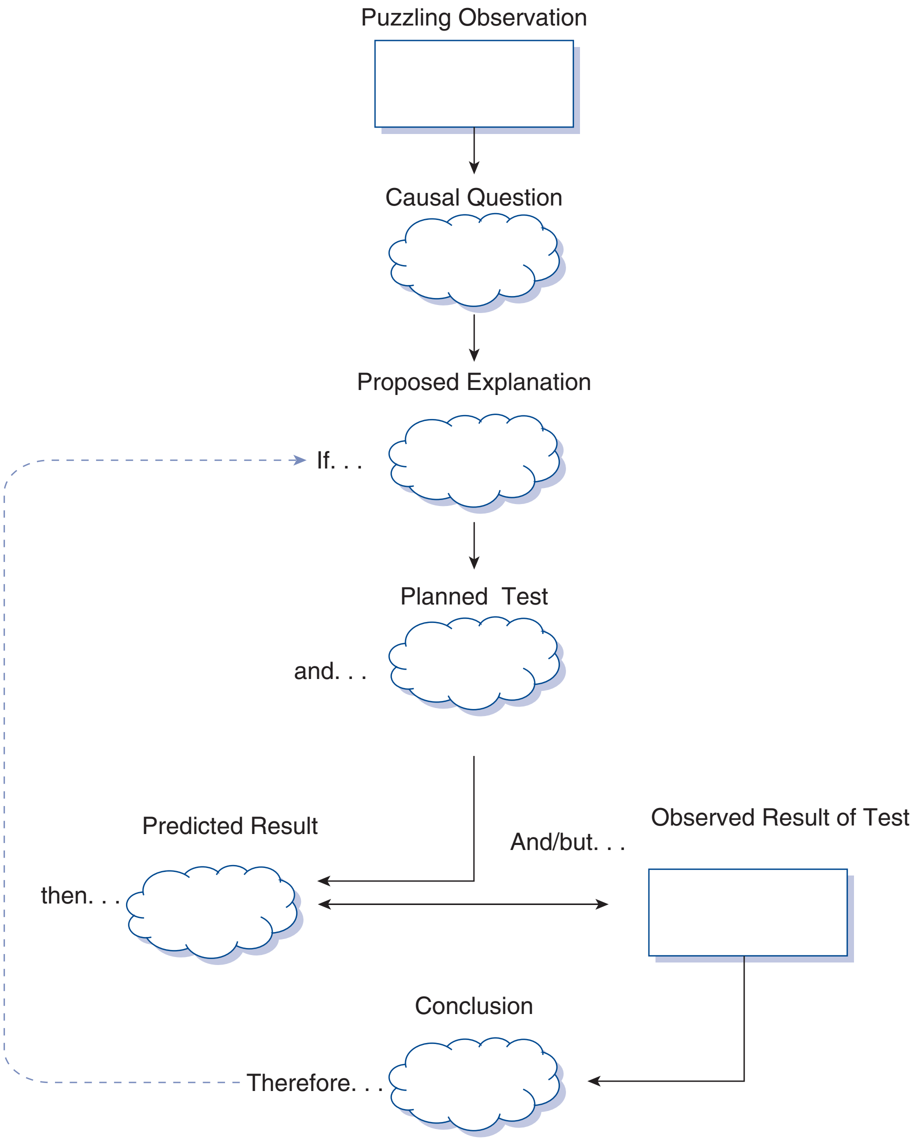
law

planned test

prediction

Application Questions/Activities

1. During the next day or two, note several objects, events, or situations that raise causal questions. For example, suppose while on the way to class you notice an automobile accident and ask, What caused the accident? Or perhaps you notice a spot of yellow grass in the middle of someone's green lawn and ask, What caused the yellow spot? Make a list of five such causal questions. Pick one of the causal questions listed above. Generate two alternative hypotheses to answer it. How could your hypotheses be tested? Use the pattern of *If/and/then* reasoning to generate expected results based on your hypotheses and some imagined test conditions. Based on your initial observations, generate a list of five descriptive questions. For example, some descriptive questions regarding the auto accident might be, Was anyone hurt? How many people were in the cars? How fast were the cars going?
2. In general, how does a description of an event differ from an explanation? Provide an example.
3. Check the newspaper during the next few days for at least two articles that discuss scientific studies. Identify the causal questions raised. What hypotheses were generated? What sort of evidence (i.e., circumstantial, correlational, experimental) was used to test the hypotheses? Were the hypotheses supported or not supported? Were you convinced by the arguments and evidence? Explain.
4. Given a group of students, name five ways in which the individuals in the group are likely to vary. Name two values for each of the named variables.
5. Fill in the boxes and "clouds" in the following figure to construct a complete argument regarding Scheepers and Simmons's test of the feeding-up-high hypothesis. The boxes represent observations while the "clouds" represent unobservable (imagined) elements of the hypothesis-testing process. Start by writing, "Giraffes have long necks" in the top box. Next, the causal question can be stated like this: "Why do giraffes have long necks?" Write this question in the first "cloud." Now write the hypothesis in the second "cloud." Now complete the figure.


Puzzling Observation
Causal Question
Proposed Explanation
If. . .
Planned Test
and. . .
Predicted Result
then. . .
And/but. . .
Observed Result of Test
Conclusion
Therefore. . .