The Generality of Hypothetico-Deductive Reasoning: Making Scientific Thinking Explicit

Author(s): Anton E. Lawson
Source: The American Biology Teacher, 62(7):482-495.
Published By: National Association of Biology Teachers
DOI: http://dx.doi.org/10.1662/0002-7685(2000)062[0482:TGOHDR]2.0.CO;2
URL: http://www.bioone.org/doi/full/10.1662/0002-7685%282000%29062%5B0482%3ATGOHDR%5D2.0.CO%3B2

The Generality of Hypothetico-Deductive Reasoning: Making Scientific Thinking Explicit

Anton E. Lawson

LEWIS (1988) made the case that today's biology is essentially hypothetico-deductive in nature. In Lewis's view, biology is not a science in which one gathers all the facts, classifies them and then somehow crystallizes them into theory. In Lewis's words, "This erroneous view of method plus the absence of a proper definition of theory misguided my early development in science" (p. 362). The failure of Lewis's teachers and textbook authors to portray science as a hypothetico-deductive enterprise not only misguided his early development as a scientist, but some teachers and textbook authors continue to misguide today's students as well. Therefore, in an effort to better understand the extent to which not only biology, but also geology, physics and chemistry are hypothetico-deductive in nature, examples of the use of hypothetico-deductive thinking in these disciplines have been sought, as have examples of use of the method in solution of practical human problems.

The purpose of this paper is to present these examples so that teachers can use them to explicate how biology, as well as other sciences, is largely hypothetico-deductive in nature. Indeed, the examples will show that hypothetico-deductive thinking is not at all new to science as it can be found in research from the Middle Ages. A key element in making hypothetico-deductive thinking explicit to students is to cast it in the form of *If . . . and . . . then . . . And/But . . . Therefore . . .* arguments. Making thinking explicit in this way is helpful because scientists seldom make their hypothetico-deductive reasoning explicit. In fact, their writing often obfuscates thinking. Further, as pointed out by Gibbs and Lawson (1992), most biology textbook writers do little better, if at all, at explicating the use of hypothetico-deductive reasoning. Let us start with early theories of blood flow and the classic research of William Harvey to see how his thinking can be cast in the form of hypothetico-deductive arguments.

Anton E. Lawson is Professor of Biology at Arizona State University, Tempe, AZ 85287-1501.

Early Theories of Blood Flow & William Harvey's Research

Based in part on his reading of Hippocratic theory, and in part on dissections and experiments on living animals, the Roman physician Galen (130–201 A.D.) proposed a theory of blood flow that became widely accepted. According to Galen, blood flowed from its origin in the liver to the heart and to a few other organs. Galen believed that both the right and left heart chambers contained blood and that blood oozed from the right chamber to the left through tiny unseen holes in the wall separating the two chambers (the cardiac septum). He also thought that some blood arrived in the left chamber from the lungs and that blood flowed out to the organs in vessels where most of it was consumed. And like Hippocrates, he imagined that some blood flowed back in the same vessels.

Galen's theory of blood flow was virtually unquestioned for nearly fifteen hundred years until 1628 when the English physician William Harvey (1578–1657) published a book titled *On the Motion of the Heart and Blood in Animals*. Harvey's book contained a revolutionary theory of blood circulation. But more importantly, Harvey's book not only presented a theory, it also presented tests of the theories' postulates, as well as tests of the alternatives. For this reason Harvey's book is generally regarded as an example of science at its best.

Like Hippocrates and Galen before him, Harvey was impressed by an analogy. But Harvey's guiding analogy was not the ebb and flow of tides. Rather, his was circular planetary orbits and the belief that large-scale planetary patterns should be echoed in smaller-scale physiological systems. Hence, Harvey set out to find circular patterns of blood flow. Harvey's proposed circulation theory can be summarized by the postulates listed in Table 1.

Notice that, unlike Hippocrates and Galen, Harvey is proposing that blood flows in two circular forked paths. One path directs blood from the heart to either the right or left lung and then back again. And the other directs blood from the heart to either the upper

Table 1. Postulates of William Harvey's Blood Circulation Theory (after Lewis 1988).

1. Blood circulates continuously due in part to contractions of the heart.
2. The heart contains one-way valves. Thus, circulating blood passes from the heart's lower-right chamber (right ventricle) to the lungs (via the pulmonary arteries), then back to the heart's upper-left chamber (left atrium), via the pulmonary veins, and from there into the heart's lower-left chamber (left ventricle).
3. From the left ventricle, blood is forced into the aorta and through its branches and subbranches to all parts of the body except the lungs.
4. From the arteries' smallest branches, blood flows through tiny unseen vessels (capillaries) into the smallest veins.
5. The veins contain one-way valves to prevent backward blood flow. Thus, due to contractions of nearby muscles, blood is squeezed from the smallest veins into larger and larger veins into the largest veins and then into the heart's upper-right chamber (right atrium).
6. The heart's right atrium then periodically forces blood into the right ventricle.

or lower body and then back again. Notice also that Harvey's circulation theory requires that arteries directing blood away from the heart connect to veins directing blood back. But in 1628, neither Harvey nor anyone else had observed such connections.

How Did Harvey Test His Theory?

To argue in favor of his fifth postulate and against the ancient ebb and flow idea, Harvey conceived of an experiment that is both elegant and astounding in its simplicity and importance. The hypothetico-deductive reasoning behind his experiment can be reconstructed as follows:

If **. . .** blood flows in veins only toward the heart because of the presence of one-way valves (Postulate 5),
and **. . .** a tourniquet is tied around the upper arm (as shown in Figure 1), a finger is pressed on the vein at point G, and then slid toward the hand down to point H (as shown in Figure 1B) (planned test),
then **. . .** the vein between points G and H should bulge only part way down toward H (expected result). The vein should bulge only down to a point at which the blood encounters a valve (i.e. at point O in Figure 1), which will presumably retard its flow back toward the hand (theoretical rationale).

On the other hand:

If **. . .** blood flows in veins in both directions (Galen's theory)
then **. . .** the vein between points G and H should bulge the entire way (expected result). The vein should bulge the entire way because both forward and backward flow presumably are normal occurrences (alternative theoretical rationale).
And **. . .** the result of Harvey's experiment is the one expected on Postulate 5 (observed result).
Therefore **. . .** Postulate 5 is supported, that is, it appears that blood flows in veins in only one direction (conclusion).

Can the one-way blood flow through the heart also be demonstrated? To answer this question Harvey planned a series of simple experiments again based on hypothetico-deductive reasoning as follows:

If **. . .** blood passes through the heart in only one direction (Postulates 2 and 6)
and **. . .** a probe is pushed through the right ventricle to the pulmonary artery, through the pulmonary vein to the right atrium, and through the right atrium to the right ventricle, and the same probe is then pushed in the opposite directions (planned tests),
then **. . .** the probe should pass through easily in the first three tests but should not pass easily, if at all, in the next three tests (expected results). The probe should pass through easily in the first three structures because these are the directions the one-way valves presumably act to direct blood. Conversely, the probe should not pass through easily in the other three structures because these are the directions in which the values presumably act to block blood (theoretical rationale).
And **. . .** when Harvey carried out his tests, all six turned out as expected based on Postulates 2 and 6 (observed result).
Therefore **. . .** Harvey claimed additional support for his circulation theory (conclusion).

How Did Harvey Test Galen's Claim About Septum Holes?

Harvey also tested Galen's claim that blood oozes through tiny holes in the cardiac septum. To do so he conducted an experiment that can be represented by the following hypothetico-deductive reasoning:

If **. . .** blood oozes through the cardiac septum from the right to left ventricle (Galen's theory),
and **. . .** water is pumped into the right ventricle of a dog's heart with its pulmonary artery tied shut (planned test),
then **. . .** water should ooze through the septum and collect in the left ventricle (expected result).
But **. . .** unfortunately for Galen's theory, no water pumped into the right ventricle collected in the left ventricle. Instead, the right ventricle ballooned up (observed result).
Therefore **. . .** Galen's theory was not supported (conclusion).

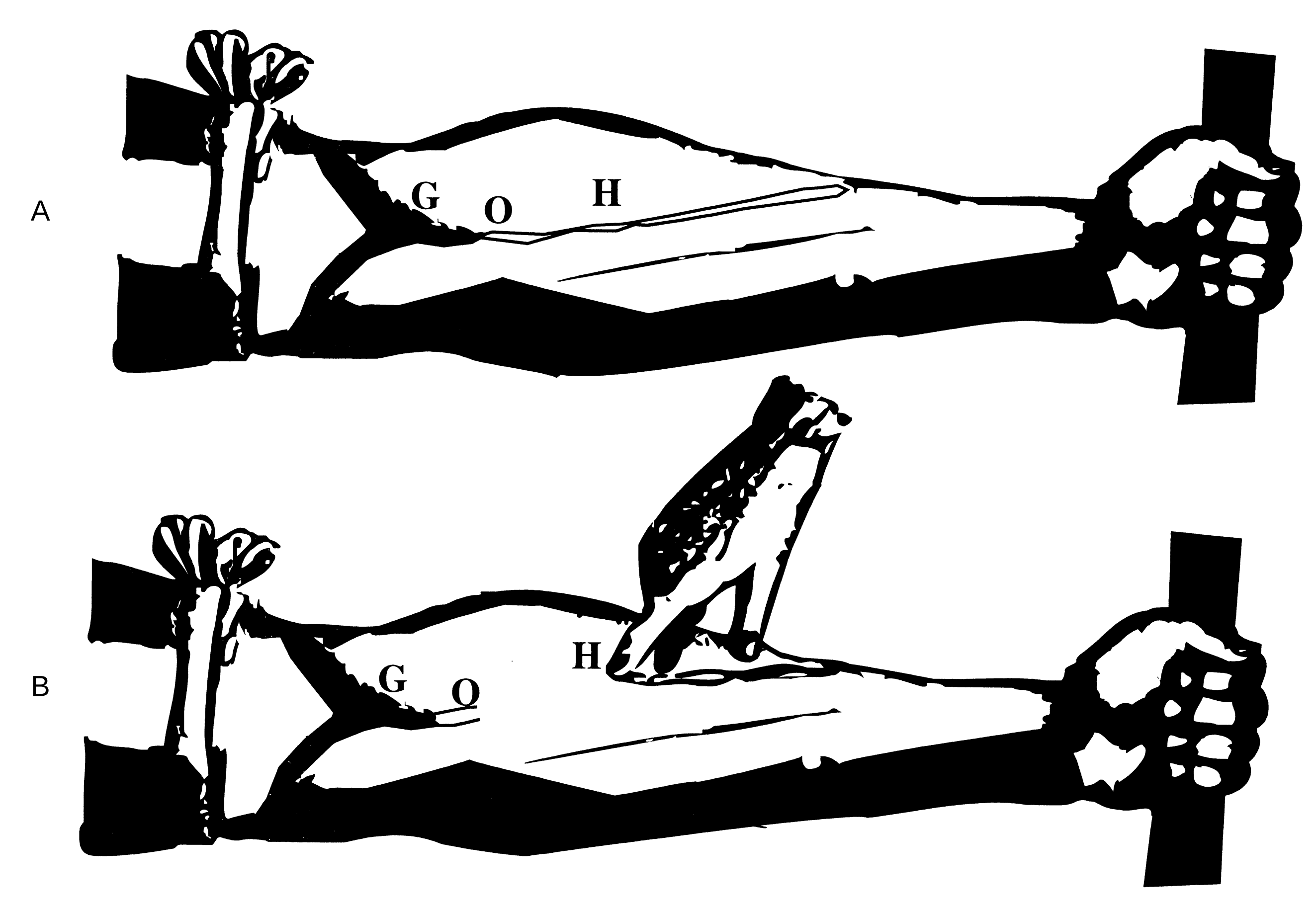

Figure 1. William Harvey's test of his one-way blood flow postulate.

What Finally Caused the Rejection of Galen's Theory?

Regardless of an impressive amount of both qualitative and quantitative evidence in favor of circulation theory and against Galen's alternative, many of Galen's supporters still held fast to prior beliefs. After all, they could still point to the fact that Harvey's postulated capillaries had not been seen. But finally in 1661, 14 years after Harvey's death, the Italian Marcello Malpighi observed capillary vessels in lungs. Malpighi's observation provided a most convincing piece of evidence in support of Harvey's theory.

Malpighi's discovery of capillaries is an excellent example of how theory generation and test directs observation—rather than the other way around. Clearly Malpighi was not merely looking at lungs to see what he might ''discover.'' Rather, he was specifically looking at areas between arteries and veins because these were areas where Harvey's theory claimed that connecting vessels should be found. In other words,

If . . . blood passes from arteries to veins through tiny vessels (Postulate 4 of Harvey's circulation theory),
and . . . the area between the arteries and veins is examined very closely (planned test),
then . . . tiny connecting vessels—the postulated capillaries—should be observed (expected result).

Consequently, Malpighi's ''discovery'' was very impressive circumstantial evidence in support of Harvey's theory *because* the theory led to the *prediction* that capillaries should eventually be seen. Thanks to the use of hypothetico-deductive reasoning, when capillaries were finally seen, the theory got a big boost and Galen's theory was finally dealt a fatal blow. Can we also find use of hypothetico-deductive reasoning in other biological research? Let's consider the research of Otto Loewi.

The Research of Otto Loewi

During the late 1800s most physiologists suspected that nerve signals were electrical in nature. Diffusing chemicals just seemed too slow to account for the apparent speed of nerve transmission. But in 1921, chemical transmission theory got a big boost thanks to a most improbable experiment conducted by German physiologist Otto Loewi. Prior to conducting his experiment, Loewi dissected out a frog's heart and the nerve connecting the heart to its spinal cord. When he electrically stimulated the nerve, the heartbeat slowed. So apparently the nerve helps regulate heart rate. Using this frog heart and nerve preparation, Loewi then designed a test of chemical transmission theory. In fact, he literally dreamed up the test in a dream! He was so excited by his dream that he awoke and immediately wrote down his plan. But in the morning when he tried to read what he had written, he found it unintelligible. Fortunately, a few nights later, the dream recurred. This time, taking no chances, Loewi awoke and immediately went to his lab to conduct the test. Basically the test amounted to stimulating the frog's nerve several times to slow its heart rate while the heart was bathed in a fluid. Then Loewi collected the fluid and applied it to another frog's heart to see what would happen.

So once again the key question is raised: Why did Loewi do this and what does its result imply? Consider the following reconstructed hypothetico-deductive argument:

If . . . the transmission of impulses between neurons and from neurons to muscle cells involves the flow of molecules across synapses (chemical transmission theory);
and . . . the frog's nerve (mentioned above) is stimulated several times to slow its heart rate while the heart is bathed in a fluid (planned test);
then . . . when that fluid is collected and applied to another frog's heart, its heartbeat should also slow (expected result). This result is expected because the imagined molecules produced by the stimulated nerve in the nerve-heart preparation should pass through the synapses separating the neurons and heart muscle cells and collect in the fluid. So when the fluid is applied to the second heart, the molecules in the fluid should produce the same effect, that is slow the second heart (theoretical rationale).
And . . . when Loewi conducted the experiment and applied the fluid to the second heart, he found that, as expected, the second heart slowed. The fluid had the same effect on several other tissues as well (observed result).
Therefore . . . Loewi's experimental evidence provides support for chemical transmission theory (conclusion).

The chemical that was presumably diffusing across synapses between the nerve and other tissues was later identified as acetylcholine. Subsequent research showed that acetylcholine is involved in the transmission of neural impulses from neurons to virtually all muscles in the body. But the important point in terms of the present argument is that the pattern of hypothetico-deductive reasoning is precisely the same as was used by Harvey. Let's now turn to geology to see if the pattern can also be found there.

Geology & the Research of Charles Lyell

Embedded in sedimentary rock layers are fossils of organisms that lived when the sediments were being deposited. As time passes, the fossilization process is repeated over and over again so more and more fossils form in layer after layer of sediments. Thus, the fossils become a chronological record of the history of life on Earth, with the most ancient organisms at the bottom and progressively more recent organisms toward the top. The data shown in Table 2 are of fossil sea shells collected from four sedimentary rock layers in England and were first published in 1854 by Charles Lyell. Notice that the table contains three columns. The first column lists the total number of fossil species found in each layer. The second lists the number of those fossil species that are still alive today. And the third column lists the percentages of fossil species still alive today. Why were these data collected and published and what do they imply?

As you know, Charles Darwin's evolution theory claims that present-day species arose from one simple type of organism many years ago through many

Table 2. Percentage of fossil species still alive today from four sedimentary rock layers that vary in age (from Lyell 1854).

Rock Layer	*Fossil Species*	*Alive Today*	*Percent of Fossil Species Still Alive*
Youngest	226	216	96%
Next Youngest	569	238	42%
Next Oldest	1,021	176	17%
Oldest	1,238	42	3%

gradual changes. Alternatively, special creation theory, at least the version that appears in *Genesis*, claims that present-day life forms were created by God during a span of a few days in the forms in which they appear today. Thus,

If . . . organisms changed over time (evolution theory),
and . . . a record of organisms living in the past is examined in the fossil record (planned test),
then . . . the younger, higher rock layers should contain more fossils of present-day species than the older, lower rock layers (expected result).

On the other hand:

If . . . organisms were created by an act of God and have not changed since creation (special creation theory),
then . . . the younger, higher rock layers should contain the same number of present-day species as the older, lower rock layers (expected result).
And . . . as can be seen in the third column, as we move up from the older, lower rock layers to the younger, higher layers, the percentages increase from 3%, to 17%, to 42% and finally to 96% (observed result). This is the increasing trend expected based on evolution theory.
Therefore . . . Lyell's correlational evidence (i.e. an inverse correlation between the age of the sediments and the percentages of present-day species) provides support for evolution theory (conclusion).

Additional hypothetico-deductive arguments that can be generated to test evolution theory and special creation theory using correlational and circumstantial evidence appear in Table 3.

Thus far we have argued that hypothetico-deductive thinking has guided the research in both biology and geology. But how about research in physics and chemistry? Can the use of hypothetico-deductive thinking also be found there? Or is research in these sciences based on other sorts of reasoning patterns? Let's consider the physics of light and the research of Thomas Young.

Physics & the Research of Thomas Young

In 1807 Thomas Young, an English physicist, conducted an experiment on the nature of light. Young's experiment involved shining light through a slit in a black screen onto a second black screen with two slits, and then onto a third slit-less, white screen. When Young conducted the experiment, he found that light passed through the slits in both the first and second screens and showed up on the third screen as a pattern of bands. Clearly, we cannot learn much about Young's thinking from this experiment other than the fact that, at least sometimes, scientists conduct experiments. To go further, we will need to try to reconstruct Young's thinking. In other words, we need to know why Young conducted his experiment and what he learned from its results.

Prior to conducting the experiment, we can be confident that Young had seen sunlight stream past clouds in apparent straight lines. We can also be confident that he had seen waves on water that travel outward in concentric circles after a pebble or stone has been tossed into a pond. Also, based on these sorts of analogies, and on the ideas of his contemporaries, it seems reasonable to suspect that Young's purpose in conducting the experiment was to find out whether light travels in straight lines or in concentric waves. But was Young using hypothetico-deductive reasoning?

Suppose that light does in fact travel in straight lines and we conduct Young's experiment. What

Table 3. Additional hypothetico-deductive arguments to test evolution theory and special creation theory.

If . . . evolution theory is correct,
and . . . we compare fossils from the oldest layers at the bottom of the canyon to fossils from the younger/higher layers and to present-day species (test conditions),
then . . . 1) species that lived in the remote past (lower layers) should be different from those living today (expected result);
2) the older (lower) the layer, the less likely that it should contain fossils similar to present-day species (expected result);
3) only the simplest organisms should be found in the oldest layers containing fossils and the more complex ones should be only in the more recent layers (expected result);
4) comparison of fossils from layer to layer should show gradual changes in fossil forms (expected result); and
5) fossils of intermediate forms (i.e. between major groups) should be found (expected result).

If . . . special creation theory is correct,
and . . . we compare fossils from the oldest layers at the bottom of the canyon to fossils from the younger/higher layers and to present-day species (test conditions),
then . . . 1) species that lived in the remote past (lower layers) should be similar to those living today (expected result);
2) the older/lower layers should be just as likely to contain fossils similar to present-day species as the younger/higher layers (expected result);
3) the simplest as well as the most complex organisms should be found in the oldest layers containing fossils as well as in the more recent layers (expected result);
4) comparison of fossils from layer to layer should not show gradual changes in fossil forms (expected result);
5) fossils of intermediate forms (i.e. between major groups) should not be found (expected result); and
6) fossils of land plants should be found in lower/older layers than fossils of sea creatures because land plants presumably were created before sea creatures (expected result).

should happen? Consider the following argument: If light travels in straight lines and shines through the first slit, then it should pass through the slit, but should be blocked by the second screen, thus should not reach the third screen. On the other hand, suppose light travels in concentric waves. Then what should happen? If light travels in waves—like concentric waves on water—and light shines through the first slit, then it should pass through the slits in both the first and second screens and should show up on the third screen as a pattern of bands (see Figure 2). And, as mentioned, when Young conducted the experiment, he found that the light passed through the slits in both the first and second screens and showed up on the third screen as a pattern of bands. What does this tell us? The result seems to be telling us that light travels in concentric waves.

Thus, Young's reasoning can be reconstructed as follows:

If . . . light travels in straight lines (straight-line hypothesis),
and . . . light shines through the first slit as described above (planned test),
then . . . the light should pass through the slit, but should be blocked by the second screen, thus should not reach the third screen (expected result).

On the other hand:

If . . . light travels in concentric waves (concentric-waves alternative),
then . . . the light should pass through the slits in both the first and second screens and should show up on the third screen as a pattern of bands (expected result).
And . . . when the experiment was conducted (actual test) the light passed through the slits in both the first and second screens and showed up on the third screen as a pattern of bands (observed result).
Therefore . . . the experimental evidence does not provide support for the straight-line hypothesis. Rather, the evidence supports the alternative that light travels in concentric waves (conclusion).

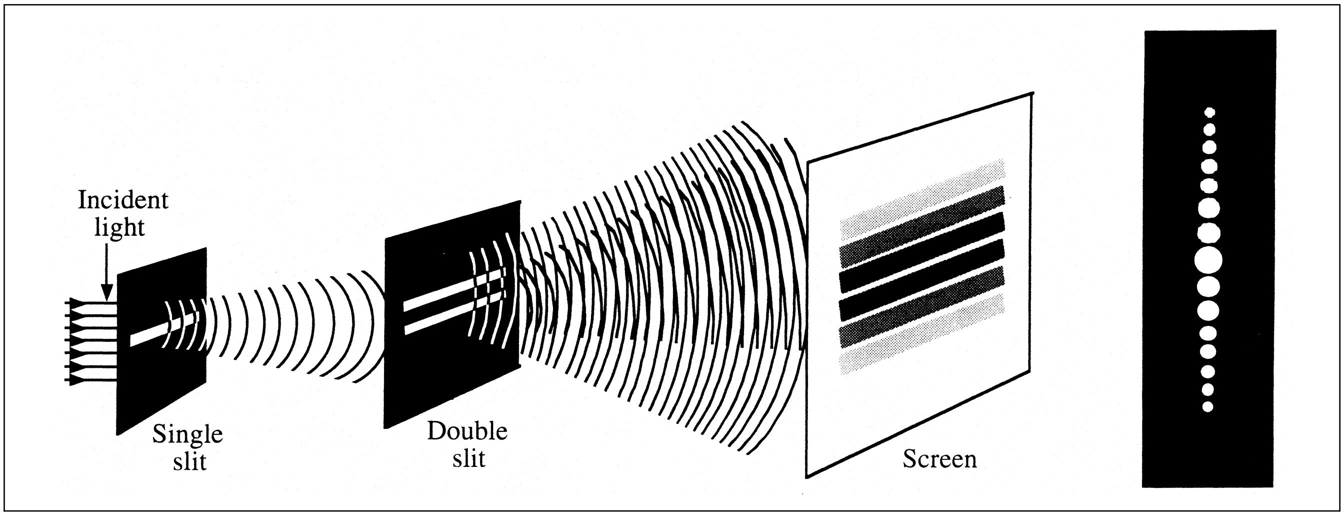

Figure 2. Thomas Young's experiment to find out if light travels in straight lines or in concentric waves.

Can the use of hypothetico-deductive thinking also be found in chemical research? To find out, let's turn to atomic theory and the research of John Dalton.

Atomic Theory & the Chemical Research of John Dalton

In 450 B.C., the Greek philosopher Democritus theorized that all matter is composed of tiny, invisible and indivisible ball-like particles. He named the imaginary ball-like particles "atoms" from the Greek word *atomos*, which means indivisible. (Note Democritus' use of analogical reasoning in the generation of scientific ideas, i.e. the tiny particles are imagined to be ball-like).

Although Democritus' idea seemed like a good one to many through the years, it was not until the 19th century and the research of John Dalton that convincing evidence was obtained that atoms actually exist. In 1810, Dalton published the data shown in Table 4 in his two-part treatise, *A New System of Chemical Philosophy*. The data are from five gases that Dalton suspected were made of different combinations of only two kinds of atoms, nitrogen (N) and oxygen (O). As was the case with Young's experiment, these data alone tell us little about why Dalton gathered the data and why they imply. For that, we will have to try to reconstruct Dalton's reasoning.

Like Democritus, Dalton suspected that all matter is composed of indivisible atoms that weigh different amounts. He also suspected that forces of attraction cause atoms to bond with others to form what he called "compound atoms" (today's molecules). For example, one atom of A might bond with one atom of B to form a molecule of A-B. He imagined other

Table 4. Dalton's data.

Gas Name	Relative Weight	Percentage by Weight	
		Nitrogen	*Oxygen*
"Nitrous"	12	42	58
Nitrous oxide	17	59	41
"Nitric acid"	19	27	73
"Oxynitric acid"	26	20	80
"Nitrous acid"	31	33	67

(Note that the names of the gases in quotations are not the names used today).

ways that atoms of A and B might form molecules such as: A-B-B, A-B-A, B-A-B, etc. The important point in terms of Dalton's conceptions (i.e. his atomic theory) is this: If matter is composed of indivisible atoms, then combinations of atoms should always exist in whole-number ratios because atoms presumably cannot be broken apart. In other words, if fractions of atoms do not exist, then fractions of atoms in combination with others also do not exist (i.e. one atom of A might combine with one atom of B, but one atom of A *cannot* combine with, say, 1/2 an atom of B).

Dalton's conceptions about atoms led him to expect that the weights of various combinations of two kinds of atoms would be in simple whole-number ratios. Because nitrous gas was lightest with a relative weight of 12 units, Dalton suspected that it consisted of molecules of one nitrogen atom bonded to one oxygen atom (i.e. NO). When Dalton heated nitrous gas, it separated into nitrogen, which remained as a gas, and oxygen, which was absorbed by a chemical. Dalton then weighed the nitrogen gas and the weight increase of the chemical. As shown in Table 4, the weights, in terms of percentages, were 42% for nitrogen and 58% for oxygen. Because Dalton assumed that nitrous gas contained an equal number of nitrogen and oxygen atoms, he reasoned that one nitrogen atom must weigh about 4 units to about 6 units for one oxygen atom (i.e. 42 to 58 is close to 40 to 60, or 4 to 6).

Dalton then hypothesized that the next gas, nitrous oxide, consists either of two nitrogen atoms for every one oxygen atom (i.e. N-O-N or N_2O), or one nitrogen atom for every two oxygen atoms (i.e. O-N-O or NO_2). The two nitrogen atoms (4 + 4 = 8 weight units) and one oxygen atom (6 weight units) N_2O hypothesis led Dalton to predict that the weight ratio of nitrogen to oxygen should be 8 nitrogen to 6 oxygen or 8/14 = 57% nitrogen to 6/14 = 43% oxygen. The one nitrogen atom (4 units) and two oxygen atoms (6 + 6 = 12 units) NO_2 hypothesis led Dalton to predict that the weight ratio of nitrogen to oxygen should be 4 nitrogen to 12 oxygen or 4/16 = 25% nitrogen to 12/16 = 75% oxygen.

Thus, Dalton's reconstructed reasoning can be summarized as follows:

If . . . nitrous oxide gas consists of N_2O molecules (N_2O hypothesis),
and . . . the gas is heated and separated into nitrogen and oxygen containing components and then weighed (planned test),
then . . . the weight ratio should be 57% nitrogen to 43% oxygen (expected result).

On the other hand,

if . . . nitrous oxide gas consists of NO_2 molecules (NO_2 hypothesis),
then . . . the weight ratio should be 25% nitrogen to 75% oxygen (expected result).
And . . . the actual percentages listed in Table 4 are 59% for nitrogen and 41% for oxygen (observed result). These percentages are much closer to those predicted by the N_2O hypothesis than to those predicted by the NO_2 hypothesis.
Therefore . . . the N_2O hypothesis is supported and the NO_2 hypothesis is contradicted (conclusion).

Another look at Table 4 shows that the percentages by weight of nitrogen and oxygen for Dalton's third gas, nitric acid, are 27% for nitrogen to 73% for oxygen. Because these percentages are so close to those predicted by the NO_2 hypothesis, it seems likely that nitric acid is made of NO_2 molecules. Based upon this sort of reasoning, one can also generate and test hypotheses about the molecular nature of the remaining gases, oxynitric acid and nitrous acid.

Thus, Dalton's general argument in favor of atomic theory and for why chemists believe that atoms exist can be reconstructed as follows:

If . . . matter consists of indivisible atoms that have specific weights and combine with one another in specific ways to form molecules (atomic theory);
and . . . combinations of atoms (molecules) are separated into their parts and weighed (planned test);

then . . . the ratios of weights of those parts should be in simple whole-number ratios (expected result).
And . . . when the planned tests are conducted, the ratios of weights of those parts are in simple whole-number ratios (observed result).
Therefore . . . atomic theory has been supported (conclusion).

Importantly, Dalton's reasoning follows the same hypothetico-deductive pattern as seen in the research discussed in the previous sections.

Elements of a General Hypothetico-Deductive Method

Regardless of which hypothesis or theory comes out the winner or loser, all of the forgoing examples suggest that at the core of doing science lies the ''*If . . . and . . . then . . . And/But . . . Therefore . . .*'' pattern of hypothetico-deductive reasoning. Therefore these examples suggest that, as Lewis claimed, scientific research is guided by a general hypothetico-deductive method. That method can be characterized as one that involves the six elements listed in Table 5.

Can the Hypothetico-Deductive Method Be Used for Practical Purposes?

Why Did TWA Flight 800 Explode?

Is hypothetico-deductive thinking useful only in science, or can it be employed to solve practical problems? Consider the job of investigators from the National Transportation Safety Board. In July 1996, TWA Flight 800 had just taken off when it suddenly and mysteriously exploded and fell into the Atlantic Ocean, killing all 230 people on board. After several months of recovery and painstaking reassembly, the assembled wreckage provided crucial circumstantial evidence allowing alternative hypotheses to be tested.

For example, gaping holes in the plane's side look like possible missile strikes:

If . . . the gaping holes were caused by missile strikes (missile-strike hypothesis),
and . . . the metal pieces around the holes are carefully examined (planned test),
then . . . traces of explosives should be found (expected result).
But . . . when the metal pieces were carefully examined, no traces of explosives were found (observed result).
Therefore . . . most likely, the holes were not caused by missile strikes (conclusion).

Also investigators noticed twisted metal pieces around one of the gaping holes that suggested damage from a bomb that exploded from the inside.

If . . . the hole and twisted metal pieces were caused by an exploding bomb from the inside (inside-bomb hypothesis),

Table 5. Elements of a general hypothetico-deductive method of doing science.

Deriving explanations for natural phenomena involves:

1. . . . raising causal questions about nature (e.g. How does light, which originates at one place, travel to another? What gives matter its distinctive properties? Does matter consist of tiny indivisible particles? How does blood circulate? What caused present-day species diversity?).
2. . . . creatively inventing / generating tentative hypotheses and theories by borrowing ideas from related contexts by using analogies (i.e. a creative process sometimes referred to as analogical reasoning or abduction). Here hypotheses and theories are understood to differ in terms of their complexity, abstractness and generality, rather than in terms of their degree of certainty. Note that the abductive process is seldom, if ever, a conscious one. In fact, as the Loewi example demonstrates, as well as several others described by McKeller (1957), abduction can even take place while one is asleep. Also note that the process is not inductive as often claimed by textbook authors (for a review see Gibbs & Lawson 1992).
3. . . . assuming for the time being that the tentative hypothesis or theory is correct. This allows one to plan some test that when conducted will provide evidence (either circumstantial, correlational or experimental) that in turn will allow the tentative hypothesis or theory to be tested. But, just as there is a creative element to hypothesis / theory generation, there is also a creative element to test design. Thus, ''cold'' logic alone will not suffice.
4. . . . using *If . . . and . . . then . . .* reasoning (the process of deduction) to generate an expected result(s), of the planned test. Expected results are sometimes referred to as predicted results or simply as predictions.
5. . . . conducting the planned test and gathering and analyzing the evidence (i.e. the observed results / data). Once the results have been obtained, they are then compared with the expected results to determine the extent to which the expected and observed results match.
6. . . . stating and communicating conclusions regarding the relative support or lack of support obtained for the tentative hypothesis or theory based on the extent of the above match. But keep in mind that a successful match between expected and observed results does not *prove* a hypothesis or theory correct in any absolute sense because two or more hypotheses / theories may lead to the same prediction. In the same vein, a poor match does not *disprove* a causal claim because one can never be certain that the cause of the mismatch does not lie in an uncontrolled test rather than in an incorrect claim. In other words, absolute proof and disproof are not possible.

and . . . the metal pieces around the hole are carefully examined (planned test),
then . . . telltale pits should be found in them (expected result).
But . . . no telltale pits were found (observed result).
Therefore . . . most likely, the hole was not caused by a bomb (conclusion).

What then caused the explosion? A massive hole in the plane's belly suggested that an internal explosion of the center fuel tank may have taken place. But at the time of this writing, no convincing evidence has been found for or against this hypothesis. Further, it is unclear what could have set off the fuel vapors. Investigators think it might have been faulty wiring, or perhaps static electricity. Lacking evidence to support this freak-accident hypothesis, the question of what caused TWA Flight 800 to explode remains unanswered. Clearly this is a serious problem. If the cause remains unknown, there is nothing that anyone can do to prevent a similar disaster in the future.

Do Breast Implants Cause Disease?

Can hypothetico-deductive thinking also be used to discover if breast implants cause connective-tissue disease? When silicone-gel-filled breast implants were first marketed in the United States during the 1960s, there was no governmental agency regulating their safety. Not until 1976 did the Food and Drug Administration (FDA) expand its safety testing to include breast implants. But because they had been marketed a decade prior, and had no reports of complications, they were considered safe. Then during the 1980s, stories began to spread about a possible link between breast augmentation and connective-tissue disease. Concern was fueled by a 1982 report that three Australian women with silicone-gel-filled implants developed connective-tissue disease. Later that year in San Francisco, the first of several multimillion dollar lawsuits was filed against breast implant manufacturers.

The media took over from there. In 1990, Connie Chung, a nationally known television news reporter, sensationalized the issue on a program in which she interviewed women who blamed their implants for autoimmune disease. Without questioning the evidence, Chung blamed the FDA for promoting ''these dangerous devices.'' Shortly after, politicians and advocacy groups took stands against breast implants, even though convincing evidence that the implants caused disease was still lacking. The FDA gave in to the pressure and banned silicone-gel implants from further use. The ban lead to a flood of women suing the manufacturers, which quickly accumulated into awards of billions of dollars. In 1995, Dow Corning, one of the manufacturers, was forced to file for bankruptcy.

But do breast implants really cause disease? Clearly experimental evidence would be a great help in answering this question. However, human experiments are unethical and cannot be conducted. But several thousand healthy women have voluntarily chosen to have breast augmentation. In a sense, they have become a voluntary ''experimental'' group. And several thousand similar women have not had implants, thus can serve as the comparison ''control'' group. Using such women as a source of correlational evidence, we get the following argument:

If . . . silicone-gel-filled breast implants cause connective-tissue disease (breast-implants-cause-disease hypothesis),
and . . . the incidence of connective tissue disease, as well as other diseases (e.g. autoimmune diseases), in women with silicone-gel implants is compared with disease incidence in women without implants (planned test),
then . . . the disease incidence should be significantly higher for the women with implants than for the women without implants (expected results).

Data for the first such correlational study were collected by researchers at the Mayo Clinic in early 1990s. Their results were reported in *The New England Journal of Medicine* in 1994 (Gabriel et al. 1994). The study included 749 women who lived in Olmsted County, Minnesota, and had received implants between 1964 and 1991. An age-matched group of 1498 women without implants was also selected from the same county to serve as the ‘‘control.’’ Researchers culled each women’s medical records for a period of approximately eight years to look for diagnoses of connective-tissue disease. Five of the 749 implant women (0.6%) had been given a diagnosis of connective-tissue disease during this time period compared to 10 of the 1498 ‘‘control’’ women (also 0.6%). In other words, the incidence of connective-tissue disease was not higher in the implant group than in the non-implant group. In fact, the incidence was exactly the same (observed results). Therefore, no evidence was found to support the breast-implants-cause-disease hypothesis (conclusion). Several subsequent studies have also failed to establish a link between implants and disease. Nevertheless, claims are still being filed and damages are still being awarded. Juries continue to ignore, or fail to comprehend, the arguments and the evidence!

Implications for Instruction

The forgoing analysis of both basic and applied science supports Lewis’s position that science is basically a hypothetico-deductive enterprise. As argued, the hypothetico-deductive method has at its core a pattern of ‘‘*If . . . and . . . then . . . And/But . . . Therefore* . . .’’ argumentation. This pattern allows hypotheses and theories, once generated, to be tested based on the degree of match or mismatch between deduced expectations and observed results gathered by either circumstantial, correlational or experimental means.

The fact that science has at its core a general hypothetico-deductive research method suggests that a greater awareness of that method and a greater adherence to its tenants would improve the quality and generalizability of science instruction (see the Appendix for an interview with a practicing scientist who uses the hypothetico-deductive method to guide not only his research but also his teaching). Figure 3 shows one important way in which we encourage students to focus on the method. As you can see, the figure consists of six empty boxes connected by the words ‘‘*If . . . and . . . then . . . therefore . . .*’’ Following each lab activity in which students have generated and tested hypotheses or theories, students are asked to fill in the boxes to construct one complete hypothetico-deductive argument based on their lab work. At first this proves to be an extremely difficult task as students have seldom, if ever, been asked to reflect on their reasoning in this way. But with experience, practice, and much instructor patience, students show considerable improvement. We continue to search for better ways to help students gain competence in the use of hypothetico-deductive reasoning and would welcome suggestions from others who share this goal.

Note

This material is based upon research partially supported by the National Science Foundation under grant No. DUE 9453610. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Science Foundation.

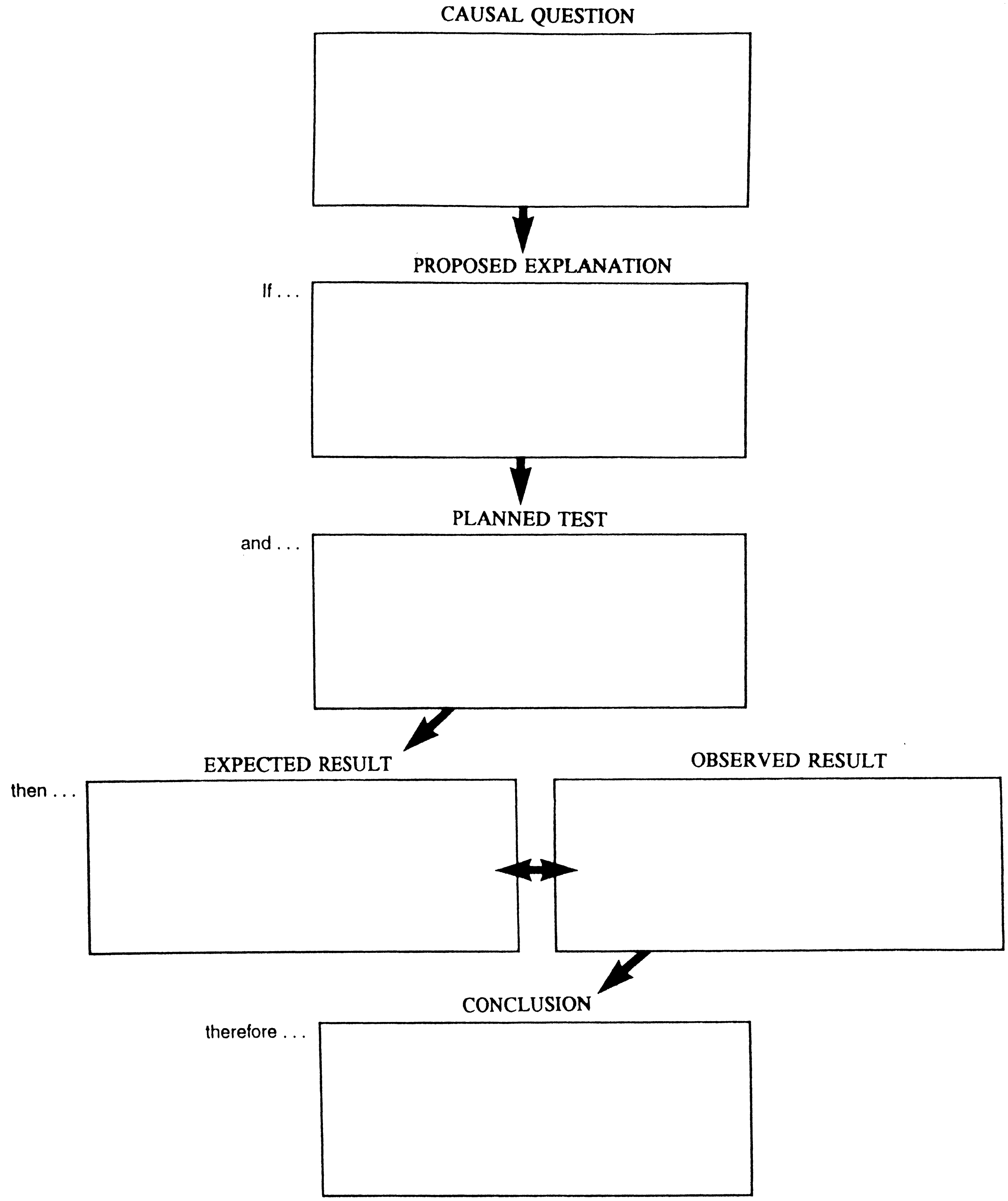

Figure 3. Students fill in the boxes to construct one hypothetico-deductive argument based on their previous lab work.

References

Alcock, J. (1996). Provisional rejection of three alternative hypotheses on the maintenance of a size dichotomy in males of Dawson's burrowing bee, *Amegilla dawsoni* (Apidae, Apinae, Anthphorine). *Behavioral Biology and Sociobiology, 39,* 181–188.

Gabriel, S.E., O'Fallon, W.M., Kurland, L.T., Beard, C.M., Woods, J.E. & Melton, J. (1994). Risk of connective-tissue diseases and other disorders after breast implantation. *The New England Journal of Medicine, 330*(24), 1696–1702.

Gibbs, A. & Lawson, A.E. (1992). The nature of science as reflected by the work of biologists & biology textbooks. *The American Biology Teacher, 54*(3), 137–152.

Lewis, R.W. (1988). Biology: A hypothetico-deductive science. *The American Biology Teacher, 50*(6), 362–366.
Lyell, C. (1854). *A Manual of Elementary Geology*. New York: Appleton.
McKeller, P. (1957). *Imagination and Thinking: A Psychological Analysis*. New York: Basic Books.
Wilson, E.O. (1998). *Consilience: The Unity of Knowledge.* New York: Knopf.

Appendix

An Interview With A Practicing Biologist About the Hypothetico-Deductive Method and Science Teaching

What follows is a transcription of an interview with John Alcock. Alcock is a well known behavioral ecologist, author of over 100 research papers and a top selling behavioral ecology textbook. Alcock clearly lays out alternative hypotheses, expected / predicted results and their tests in a recent paper entitled ''Provisional rejection of three alternative hypotheses on the maintenance of a size dichotomy in males of Dawson's burrowing bee, *Amegilla dawsoni*'' (Alcock, 1996). Alcock was interviewed to learn about his views on the hypothetico-deductive method and how it affects his approach to research and to science teaching (Alcock's responses appear in italics).

How do you do science? In other words, do you have a general plan of attack, a general set of strategies, a general method, that you use from one study to the next?

Yes, in terms of selection of topics I am committed to studies of insect mating behavior. The basic technique is the standard one. I am using evolutionary theory to come up with questions. Once I have questions, I am developing hypotheses that are consistent with selection theory and testing them the old fashioned way.

What is the old-fashioned way?

By using them to generate predictions for which it is possible to collect data so that we can examine the validity of the predictions.

Once you have data, how do you examine their validity?

Well, by matching the expected results against the actual ones.

How do you draw conclusions from that? Or do you?

Yes, in my case the conclusions are invariably in the form of the data support or invalidate the particular hypothesis.

How generalizable is this technique of generating and testing hypotheses? For example, is it limited to your field of research?

I believe it is fundamental to all science. It is the essence of what is called the scientific method.

The scientific method? Is there only one?

I think, well, there is descriptive science, which is the foundation for asking causal questions. And the kind of science which has the greatest significance for everybody—the causal question answering science for which this hypothesis-testing technique is, I believe, fundamental. I have never seen any study, never had anyone explain to me how any study did not use this particular approach, even if they claimed that there are multiple scientific approaches.

Does this method, this thinking process, actually guide your research?

Very self consciously, yes.

Do you think the method applies to other professional fields, even to non-professional aspects of one's life?

I certainly do. I think you could actually have, as E. O. Wilson has argued (Wilson, 1998), a superior economics, a far superior sociology, a far superior women's studies, were this technique applied vigorously throughout. Laying out the hypotheses, thinking through the predictions in advance, is hugely helpful.

How did you come to use this method?

At this stage, I cannot recreate the steps that led to my current firm views. But it did have something to do with thinking through teaching biology to undergraduates.

How about when you were a graduate student at Harvard? Did you use the method then?

I was definitely unaware of what I was doing, just following through. Well, the scientific method is common sense, logic I'd say, and not that obscure. But I wasn't self conscious. I was intuitive and intuitive throughout much of my early career. I only became aware of it in the past 10 to 15 years, perhaps in conjunction with teaching undergraduates. I do not know.

Do you think that other people's research would be improved by explicit use of the method?

I think it would be improved in two ways. First, it would help the researcher be more systematic in thinking through what he or she should be looking for. There is a tendency to think of alternative hypotheses after the fact and then try to scramble about and hand wave your way out of the problem. Were it actually applied rigorously in advance, it would save you a lot of heartache and wasted effort. Second, it would have a huge positive effect in the writing of the paper which would enable you to convince your colleagues that you had done what you set out to do. Papers that utilize something along the lines of the hypothesis, prediction, outcome, conclusion format, are papers that are readily understood.

How good are typical college undergraduates at using this reasoning process, this method?

I think that the typical undergraduate has an intuitive grasp of it because so much of life revolves around figuring out what caused this or that to happen. And people generally do a decent job at it, of course with all sorts of interesting exceptions. But obviously being self consciously aware of what they are doing and the nature of the logic, the average undergraduate doesn't have a clue.

Is trying to improve their understanding and use of the method important?

Sure, this is the fundamental goal of my undergraduate biology course.

How do you go about trying to improve their reasoning and the ability to use the method?

I would say the key weapons are exam questions, sample questions, and quiz questions so that students are forced to put a label on a hypothesis as opposed to a prediction, or forced to look at data and say no, that is not the conclusion, that is the actual result that was gathered to test the hypothesis with the conclusion being hypothesis rejected, hypothesis accepted.

Do your lectures contain examples of this sort of reasoning?

Yes. I write every lecture to revolve around hypothesis, prediction, test, conclusion—every single one.

How successful have you been?

I have a feeling that I reach about one third of the students at a level that ought to be there. The next one third, I reach with intermediate effectiveness. And the bottom one third, I definitely do not affect.

What needs to be done in your course or elsewhere to be more successful?

If high schools worked at conveying an enthusiasm for this issue, if that were to happen, then obviously we would be way ahead; and we could move to a farther point down the turnpike. Generously, there are maybe 10 major conceptual systems in biology that every student should know. And the foundation for those conceptual systems is always the scientific method. So that should be the premiere goal and understanding the 10 major conceptual systems is the secondary goal. All the other material is information that will be entering one ear and passing out the other.

So if you were ultimately successful in an introductory undergraduate course, does this mean that the teacher in subsequent courses could forget about this method?

No, I think it is entirely useful to keep going over it in each new context. At some future date when the scientific method is used in disciplines other than science, then the student could move from class to class with the beautiful result that what he/she learned in the science class is applicable in the humanities class and vice versa. Understanding how valid discoveries and conclusions are made ought to be of extreme interest to any educated person.

Is there anything else you would like to add?

Yes. It seems that the missing element in all of this is getting a social and emotional context in which a student can absorb this information. I do not know why it is compelling to me. But it is. I find it fascinating to look at a paper and figure out what the process was, how the results were constructed. But I know that this enthusiasm is not shared by most undergraduates. So there has to be another tack to get them involved.