Strong Inference: The WAY of SCIENCE

Thomas B. Kinraide R. Ford Denison

Valentine: *It may all prove to be true.*

Hannah: *It can't prove to be true, it can only not prove to be false yet.*

Valentine: (Pleased) *Just like science.*

– from "Arcadia," a play by Tom Stoppard

Science teachers and science textbooks commonly introduce students to the scientific method in elementary and junior high school, but the study of scientific method and philosophy can be a life-long endeavor. Our essay concentrates on a particular aspect of the scientific method – the testing of hypotheses. Concepts of hypothesis testing have changed even within the relatively short period of modern science. Specifically, the concept of proof has been abandoned for reasons we shall describe. Although we can not *prove* hypotheses, we can almost certainly *disprove* some hypotheses, if they are false.

Dr. Kinraide is a plant physiologist with a federal research agency. Dr. Denison is with the Department of Agronomy and Range Science, University of California-Davis, Davis, CA 95616, e-mail: rfdenison@ucdavis.edu. Both welcome comments and inquiries from students and others sent to the e-mail address.

To describe the modern method of hypothesis testing, we borrow the term "strong inference" from John R. Platt's *Science* (1964) essay by the same name. In brief, strong inference is the method of testing a hypothesis by deliberately attempting to demonstrate the falsity of the hypothesis. A hypothesis that repeatedly withstands attempts to demonstrate its falsity gains credibility, but remains unproven. We are confident that our essay reflects the thinking of most scientists that hypotheses are potentially disprovable but not provable. Nevertheless, we qualify these views somewhat, arguing that neither proof nor disproof is certain.

Strong inference is an avenue to knowledge that is systematically applied in science, but some practice of strong inference has occurred in human endeavors for thousands of years. For example, courts of law in ancient civilizations occasionally used elements of strong inference – facts were assembled from physical evidence and the testimony of witnesses; hypotheses were developed (only the grand vizier could have stolen the documents); and impossible or illogical consequences of the hypotheses were grounds for rejecting the hypotheses (an alibi would establish the grand vizier's innocence). Nevertheless, former and present methods of inference sometimes differ significantly – an ancient magistrate may have awaited a ghostly visitation during which the truth of a case would be revealed; the body of an accused witch may have been examined for incriminating marks; and confessions may have been extracted by torture. [This mixture of strong inference and alternative methods is described in tales of the historical Chinese magistrate, Judge Dee, by the Dutch diplomat and scholar Robert Van Gulik (1976).]

Even today, people rely upon alternative avenues to knowledge that may include intuition, revelation, and adherence to authority. We are reluctant still to use strong inference outside of enterprises that are recognizably scientific, and the application of strong inference to some beliefs may be impossible. Even when strong inference is possible, its application may be uncomfortable, and its application to the beliefs of others may be considered hostile. Challenges to authority and received wisdom may seem disloyal or arrogant. This reluctance to use strong inference follows understandably from the requirement that beliefs (or hypotheses) be subjected to deliberate attempts to demonstrate the falsity of the beliefs and by formulating and testing competing beliefs. Nevertheless, strong inference can be practiced with civility and can do much to offset our prejudices and natural gullibility.

A Definition of Hypothesis

Because the formulation and testing of hypotheses are at the heart of strong inference, we will present a definition of hypothesis here, however a detailed discussion of hypotheses will be delayed until some other terms, incorporated in the definition, are considered. For the definition of hypothesis, and most other terms, we have consulted *Webster's Third New International Dictionary, Unabridged* (Gove, 1976).

> **Hypothesis:** [An explanatory] *proposition tentatively assumed in order to draw out its logical or empirical consequences and so test its accord with facts that are known or may be determined.*

Inevitably, the burden of definition is shifted to other words. In the present case, "fact" is one of those words. Strong inference ultimately rests upon facts, and facts and hypotheses are sometimes confused with each other. Therefore, we shall consider first the concept of fact.

The Concept of Fact

> **Fact:** *An occurrence, quality, or relation the reality of which is manifest in experience or may be inferred with certainty.*

Here, too, the burden of definition is shifted to other words, among them, "experience" and "reality." To deal with these terms we must concede that science rests upon a few basic assumptions. Science assumes that nature has a reality independent of the human mind, and science assumes that the human mind can grasp the reality of nature. These epistemological issues are rarely considered in the ordinary practice of science.

Manifest Fact & Inferential Fact

The definition of fact indicates the existence of two kinds of fact – *manifest fact* and *inferential fact*. Again, some definitions may be helpful.

> **Manifest:** *Capable of being easily understood or recognized at once by the mind: not obscure: obvious.*

> **Inference:** *The act of passing from one or more propositions ... considered as true to another the truth of which is believed to follow from that of the former.*

Manifest facts are not highly dependent upon inference. We will call a fact that is highly dependent upon inference an *inferential fact*. To illustrate inferential and manifest facts, consider the case of a forest fire. If the fire occurred recently, then its occurrence is likely to be a manifest fact. It may have been observed by hundreds of people, and newspaper readers and television viewers are certainly being reasonable in accepting the occurrence of the fire as a manifest fact.

What if the fire had occurred 200 years ago? Most scientists would accept as fact (inferential fact) that a fire had occurred in an area if several observations pointed, convergently, toward a fire. These observations might include the absence of any trees in the area older than 200 years (despite the presence of older trees in surrounding areas), the scarcity or absence of old wood on the forest floor, and the presence of an ash layer beneath the recent leaf and twig litter. Perhaps none of these observations was convincing by itself (the ash may have been blown in from another fire some distance away). Convergence of evidence is the clincher.

In some cases, facts and hypotheses may be confused, but confusion may be avoided by remembering that a hypothesis is a candidate explanation, not a candidate fact. The statement "The Earth is spherical" in ancient times was a candidate fact, and in the present age of satellite photographs, and other evidence, the statement may be regarded as a manifest fact. The statement was also a hypothesis in ancient times, but only when used as an explanation for some other observation. Thus the statement "Vertical objects cast shadows of different length at different latitudes because the Earth is spherical" is a hypothesis (a candidate explanation) and not merely a candidate fact. If we confuse a candidate fact for a hypothesis, then we may conclude mistakenly that hypotheses are provable.

Scientific Facts Are Public

Another feature of scientific facts is that they are *public;* that is, a fact (especially a manifest fact) is accessible to all competent observers. The issue of competence is sometimes problematical. In science, public accessibility to facts is crucial even though comprehension of the facts is not always easy. The devotees of mystery cults may be entitled to *both* their own private opinions *and* their own private facts, but *science disallows private facts.*

The Concept of Hypothesis

"Science" and "strong inference" are not synonymous. Science is both a method and a body of knowledge. Facts can be compiled and many questions can be answered without the formulation and testing of hypotheses. Natural history inventories (lists of birds, plants, minerals, and other items) play a role in science and in society. The answer to some questions (What is the speed of light?) may require high technical skill but can be answered without the formulation of hypotheses. In some cases, laws of nature may be formulated without the explicit testing of hypotheses. (Laws are descriptive, often quantitative, but not explanatory, statements having a value intermediate between fact and hypothesis. Examples are *Ohm's law* [I = V/R], *Newton's law of motion* [e.g., F = ma], and the *law of conservation of charge.*)

Despite the possibility of some success in science without the testing of hypotheses, science attempts to do more than just compile and describe. Science attempts to explain. This requires the formulation of hypotheses in a creative process that may require the investigator to think beyond readily available explanations. A good hypothesis must be explanatory, but it must have another feature too: It must be testable by strong inference. If it is false, it must be possible to show that it is false.

A Case Study of Hypothesis Testing

A textbook that one of us (T. B. K.) assigned years ago as a college professor was *The Study of Biology,* 3rd Edition (Baker & Allen, 1977). The first two chapters of that book, *The Nature and Logic of Science* and *Testing Hypotheses and Predictions,* are excellent. The following case study is taken from that book.

The Pacific salmon *Oncorhyncus kisutch* hatches in streams in the Northwest, swims to the sea, then, eventually, returns to streams to spawn. We may ask, and answer, the question "Do individual fish return to the streams of their birth?" without formulating an explanatory hypothesis. Tagging experiments have confirmed the fact that the fish predominantly do return to their natal streams. In order to determine how the fish do this, we can proceed in one of two ways. We can continue to study the fish, compiling facts in the hope that an answer may emerge. Sometimes "fishing expeditions" such as these can lead to serendipitous results, but eventually strong inference (hypothesis formulation and testing) is usually needed.

Platt, in the *Science* article cited above, makes an important suggestion: Formulate more than a single hypothesis. With more than one hypothesis, the investigator is less likely to adopt a "pet" hypothesis to which he/she becomes emotionally attached, and the necessary attempt to demonstrate the falsity of the hypotheses is less worrying – perhaps one will survive. Incidentally, *the negation of a significant hypothesis is a significant contribution to science.*

In our case study, two hypotheses as to how salmon find their way back to their natal streams might be these:

1. *Salmon find their way back using their sense of sight.*
2. *Salmon find their way back using their sense of smell* (detecting dissolved substances from their birth streams).

Hypotheses are formulated on the basis of prior knowledge, and we know that fish both see and smell. The hypotheses just stated were rather obvious possibilities, but the formulation of hypotheses may be very difficult. The observations for which an explanation is sought may be very strange (divorced from ordinary experience). Sometimes a hypothesis may be formulated that seems very good because it is compatible with almost all of existing knowledge, but not all of it. In that case, we must consider that the hypothesis, however attractive, may be wrong or that some of the accepted knowledge is wrong.

The next step in strong inference is to test the hypotheses. That is done by deliberately subjecting them to jeopardy, that is, by attempting to demonstrate their falsity. In our fish story, each of the two hypotheses has logical consequences that give rise to predictions as to the outcomes of certain experiments. The hypotheses and the predictions are often stated together in *if ... then ...* statements. It is very important to make these statements explicit. Such a formulation applied to our example may be "*If* salmon find their way back using their sense of sight, *then* salmon with shielded eyes (black plastic discs were used in an actual experiment) will predominantly fail to find their birth streams." The salmon did, in fact, find their way back in the experiment, and the hypothesis was thus considered to be false. The alternative was tested after formulating the statement "*If* salmon find their way back using their sense of smell, *then* salmon with a blocked sense of smell (benzocaine ointment was used) will predominantly fail to find their birth streams." This prediction came true, and the second hypothesis was regarded as supported, but not proved.

The Impossibility of Proof

The problem is that even false hypotheses may sometimes give rise to correct predictions. For example, consider the false hypothesis that salmon find their way back to their birth streams by the sense of sight. This gave rise to the prediction that sightless salmon will predominantly fail to find their birth streams. This prediction turned out to be incorrect in the experiment cited earlier, but conceivably the prediction could have been correct. Suppose the blindfolded salmon were so traumatized by the blindfolding operation that they did not try to return or that they became so confused without their sight that they ignored their sense of smell and swam off randomly from their release site. In such cases the prediction would have been correctly fulfilled. Is the hypothesis in that case "proved?" Certainly not, though the investigators may claim support for the sight hypothesis if they failed to observe the trauma or the confusion.

A logical truth table presented by Baker and Allen, and others, shows the relationships.

HYPOTHESIS	PREDICTION
True	Correct
False	Correct or Incorrect

According to the table, an incorrect prediction *always* corresponds to a false hypothesis, but a correct prediction can come from *either* a true *or* a false hypothesis. Because of these relationships, hypotheses are often regarded as potentially disprovable (falsifiable) but rarely provable. How then do some hypotheses come to be regarded as true?

> *A hypothesis is supported, but not proved, when repeated attempts to negate the hypothesis fail, when competing hypotheses are discredited, and when additional facts (not used in the initial development of the hypothesis) are successfully embraced by the hypothesis.*

In the case of the fish, the smell hypothesis withstood an opportunity for disproof, and the competing sight hypothesis was disproved. Still, the smell hypothesis is not proved. Perhaps smell plays no role, and a third sense is the key. Perhaps the benzocaine treatment so traumatized the fish that they could not function properly, or perhaps the benzocaine knocked out the third sense. These worries lead to additional hypotheses, predictions, experiments, and facts.

Another way of considering the general unprovability of hypotheses is that *no hypothesis can be considered proved if an alternative hypothesis, that excludes the possibility of the first hypothesis and is equally compatible with the facts, is possible.* Since we can never be sure that we have considered all possible hypotheses, proof remains unattainable.

Earlier, we stated that a hypothesis is a candidate explanation, not a candidate fact. The case of the salmon provides an illustration of the difference. Early on, people may have observed that the salmon in a particular stream were physically similar to each other and different from salmon in another, distant stream. A couple of hypotheses may be stated:

1. *Only salmon of a particular body type are able to navigate a particular stream and that is why they look alike.*
2. *Salmon return to their natal streams to spawn and look alike because they are genetically similar.*

The "fact" that salmon do return to their natal streams establishes the truth of the statement "Salmon return to their natal streams," but this statement was a candidate fact, not a hypothesis, and the second hypothesis remains unproved.

The Uncertainty of Disproof

Although scientists often refer to the disprovability of hypotheses (as we have), we contend that disproof is uncertain also. The reason for this is the requirement for the prediction of *logical* consequences in the testing process, but we can never be certain that our predicted consequences are logical. As an example let's return to one of our *if ... then ...* statements. "*If* salmon find their way back using their sense of smell, *then* salmon with a blocked sense of smell will predominantly fail to find their birth streams." Suppose that we had unwittingly made the illogical statement "*If* salmon find their way back using their sense of smell, *then* the Red Sox will win the World Series." If the Red Sox failed to win, we would have concluded, falsely, that the hypothesis was false.

The Red Sox example used a preposterously illogical prediction, but some illogical predictions are not so obviously illogical, and the problem is not trivial in some cases. Sometimes scientists disagree over the cogency of a predicted outcome, especially in complex situations where variables are hard to control (see *The Triumph of Sociobiology* by John Alcock [2001] for interesting discussions of some uncertainties and controversies). An outcome that constitutes adequate grounds for the rejection of a hypothesis for one investigator may be viewed as inadequate by another investigator. The problem of the illogical prediction can be alleviated by testing additional predictions and by the public critique of the methods and conclusions. (The initial stage of public critique is the expert "peer review" of scientific manuscripts prior to publication. See the Acknowledgment in this essay.) Despite the uncertainty of disproof, scientists accept the qualified use of terms such as "disproof," "falsification," and "negation," but not the term "proof."

The Concept of Theory

When a hypothesis has undergone very extensive testing, especially if the testing attacked the hypothesis from many different angles using independent lines of evidence, then the hypothesis may graduate to the status of theory or, together with other hypotheses and principles, become incorporated into a theory. A dictionary definition of theory is this:

> **Theory:** *The coherent set of hypothetical, conceptual, and pragmatic principles forming the general frame of reference for a field of inquiry.*

The term theory implies that the component hypotheses are very likely to be true and that together are important and comprehensive. Theories, like well-supported hypotheses, give rise to predictions that are consistently correct, but in the case of theories the range of predictions is often wider than the range of predictions for hypotheses. Theories come to provide a conceptual framework for scientific thought. Some examples include *The Atomic Theory, The Theory of Evolution, The Germ Theory of Disease, The Theory of Relativity,* and *The Quantum Theory.* Despite their high status, theories are still hypothesis-like (perhaps we could call them metahypotheses), and as such they are necessarily vulnerable. That is, they must be testable, and potentially falsifiable.

Will Strong Inference Always Work?

Some issues that would seem to be accessible by strong inference remain controversial because of emotional involvement, inadequacy of definitions, or a variety of technical difficulties. For example, a few scientists and public policy makers refuse to acknowledge that HIV is the causative agent in AIDS, and the causes, and even the occurrence, of global warming remain controversial.

For many people, science is not the only pathway to knowledge. For them, propositions may rest upon personal revelation or upon religious authority, to cite just two additional pathways to knowledge. For the faithful, faith propositions are considered to be truths, not hypotheses. With regard to the term hypothesis, believers and scientists are in agreement. In most cases, neither scientists (many of whom are religious) nor religious believers (some of whom are scientists) consider religious beliefs to be hypotheses; believers because they consider applying the term to religious teachings to be belittling, and scientists because the term hypothesis can be applied only to statements that their adherents are willing to subject to possible disproof.

Although not scientific, faith propositions are not necessarily in conflict with science, but they may be. *A tenet of faith that cannot be accessed by strong inference because it is beyond the technical or epistemological scope of science is not in conflict with science.* Examples include doctrines that claim consciousness in inanimate objects, a purpose to life, or rewards or punishments after death. Science cannot now address these propositions, although it may be able to do so in the future (formerly, only faith, not science, could address such issues as the cause of disease, the change of seasons, and the formation of stars).

Some faith propositions are clearly in conflict with science. *A tenet of faith that can be accessed by strong inference may be, but is not necessarily, in conflict with science.* The indigenous religion of Hawaii provides a fascinating case study. At the time of European discovery, Hawaiian society was encumbered by hundreds of taboos whose violation was thought to ensure calamity for individuals and society (Malo, 1959). This religion disintegrated quickly as Hawaiians observed that Europeans (and Hawaiians influenced by Europeans) could violate the taboos and live to tell about it. The Hawaiian nobility quickly embraced the religion of the Europeans and ordered the destruction of idols and the abandonment of many taboos. The causes of this religious transition are complex, but the obvious conflict between reality and some of the faith propositions surely played a role.

A Summary of Strong Inference

1. Observed and inferred facts inspire a question.

2. The question inspires one (or preferably more) hypotheses. This is a creative process. Several hypotheses may be proposed, and they need not have a high likelihood of being supported, but a good hypothesis must be an explanatory statement that is testable.

3. The hypotheses are *deliberately* subjected to jeopardy (falsification) by, first, stating the logical consequences of the hypotheses. Statements in the form "*if* (the hypothesis), *then* (the consequences)" are useful.

4. Next, the accuracy of the predicted consequences are tested by the acquisition of new facts from experimentation, or observation, or from the body of known facts not already used to formulate the hypotheses.

5. Incompatibility between prediction and outcome leads to the rejection of hypotheses, and compatibility leads to tentative acceptance. In all cases, *repeated* incompatibility or compatibility from separate lines of testing is desirable.

6. The hypotheses, together with the facts and the record of the inferential process, are submitted to public scrutiny and may become accepted into the body of public knowledge.

7. An accepted hypothesis typically spawns the acquisition of more facts and the formulation of new hypotheses (perhaps by the critics of the old hypothesis). These ongoing exercises in strong inference may cause the revision or rejection of the accepted hypothesis.

8. A hypothesis, or more often a collection of complementary hypotheses, may become incorporated into a theory.

Acknowledgment

We are grateful for an insightful review that helped to clarify our thinking on at least one important issue.

References

Alcock, J. (2001). *The Triumph of Sociobiology.* Oxford: Oxford University Press.

Baker, J.J. & Allen, G.E. (1977). *The Study of Biology,* 3rd Edition. Menlo Park: Addison-Wesley Publishing Company.

Malo, D. (1959). The Alii and the common people. In A.G. Day & C. Stroven (Eds.), *A Hawaiian Reader* (pp. 337-345). Honolulu: Mutual Publishing. *Note:* David Malo was a native Hawaiian and lived from 1793 to 1853.

Platt, J.R. (1964). Strong Inference. *Science, 146,* 347-353.

Stoppard, T. (1993). *Arcadia.* London: Faber & Faber, Ltd.

Van Gulik, R. (1976). *Celebrated Cases of Judge Dee.* New York: Dover Publications.

Webster's Third New International Dictionary of the English Language, Unabridged. (1976). P.B. Gove (Ed.). Springfield: G. & C. Merriam Company.

Note: The items listed above are publications cited in our essay. They do not constitute a *Bibliography* for the topics of scientific method or philosophy. Better libraries and bookstores stock many books suitable for lay people. Some magazines and newspapers include book reviews, and some introductory science textbooks include bibliographies that may be helpful. Searching the Web on the topic of philosophy of science, or using key words in combination with "science and hypothesis," will lead to many interesting sites, some of which include discussions of Sir Karl Popper's contributions to the concepts presented in our essay.