



Sir Francis Bacon

# The Scientific Method

by DAVID and JUDITH GOODSTEIN

How did the scientific method develop and do practicing scientists really use it?

*Judith Goodstein:* Two major philosophical schools of thought about the nature of the scientific enterprise are, first, that of the 17th-century philosopher of science, Sir Francis Bacon, and, second, that of the 20th century's Sir Karl Popper. Bacon's ideas on this subject have, of course, dominated Western scientific thinking for more than 300 years. In fact, it is Bacon to whom we owe the idea that there *is* a proper way to approach the study of science.

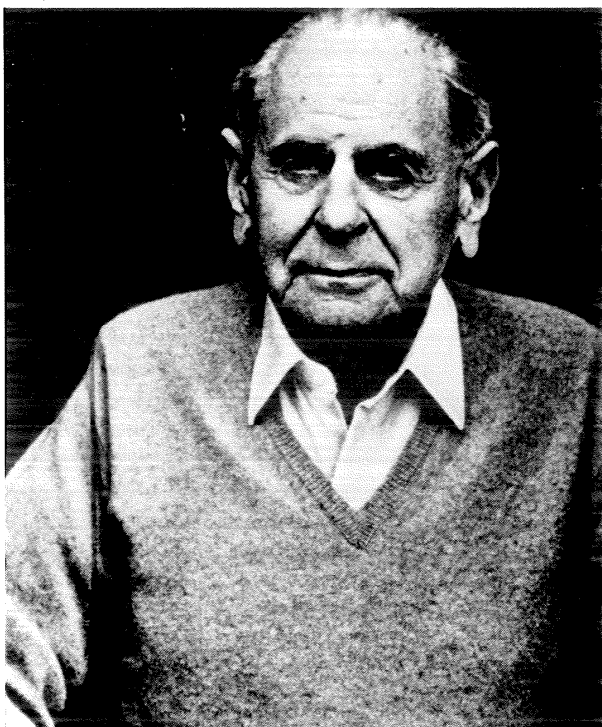
Francis Bacon was born in England in 1561. He was

educated at Trinity College and then entered Gray's Inn, where he studied law. He was also something of a politician, and he became chancellor under James I in 1618. Three years later he was dismissed after being convicted of taking bribes. His scientific contemporaries included Gilbert, Galileo, and Kepler, but he remained isolated from the scientific developments associated with them. He attacked both Copernicus and Ptolemy for producing only "calculations and predictions" instead of "philosophy, what is found in nature herself, and is actually and really

true.” His knowledge of the sciences, it turns out, was largely based on literary sources, and on this solid foundation he built his famous theory of the scientific method.

Bacon held the view that the scientist starts his research by recording observations. If these observations were correct, he believed, they would lead to equally correct judgments, or generalizations, about nature. In the application of this inductive process, Bacon outlined a necessary sequence of steps to be followed. To make true inductions, one must begin by purging the intellect of “idols” that obstruct man’s unprejudiced understanding of the world. If this is accomplished, the mind becomes, in Bacon’s phrase, “a clean slate,” on which true notions can be imprinted by nature itself.

Bacon’s inductive method started with observations that would lead to the construction of systematic tables of the presence, absence, and comparisons of properties. From these, inferences would be made that could then be “put to the question” by artificial experiment. While the ascent from particular observations to generalizations is a very complicated process, Bacon felt that done properly it would result in a number of inferences whose conclusions would be infallible. Furthermore, he was aware that infallibility would depend on there being only a finite number of properties — and on the scientist’s ability to list all of them in any given instance. Since some properties are “hidden,” he was amenable to the use of “aids to the senses,” which included the telescope, for example, and



Sir Karl Popper

many other kinds of laboratories and instruments.

Bacon invented a scientific Utopia in which there is a division of scientific labor. Those who do experiments and collect information form the first group; a second determines the significance of the information and experiments and carries out new ones; and a small third group, known as the interpreters, “raises the former discoveries into greater observations and axioms.” Bacon assigned 33 experimentalists to the first two tasks. He didn’t see the need for more than 3 interpreters.

Bacon’s scientific methodology can be summarized as follows: 1. The scientist must start with a set of unprejudiced observations; 2. these observations lead infallibly to correct generalizations or axioms; and 3. the test of a correct axiom is that it leads to new discoveries. Three hundred years later, Sir Karl Popper arrived at a different view.

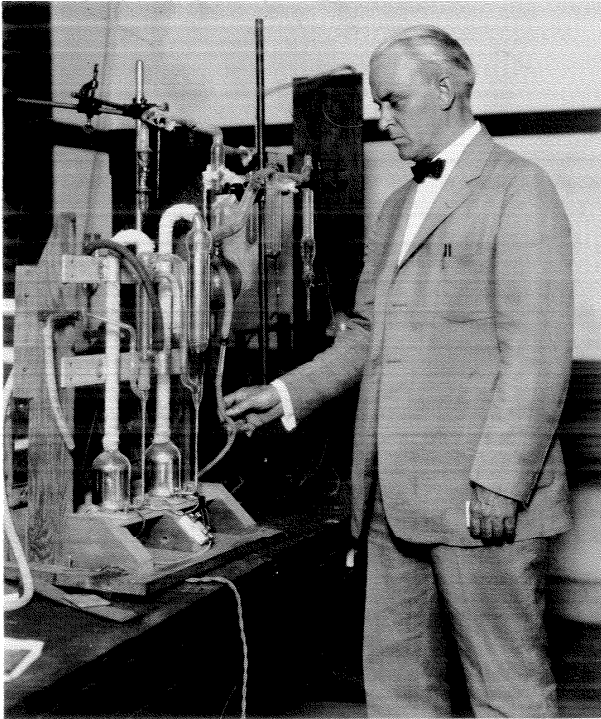
Popper is a contemporary philosopher of science. He was born in Vienna, where he received his university and graduate training and published his major work, *The Logic of Scientific Discovery*, in 1934. Popper’s scientific contemporaries include Einstein, Heisenberg, Bohr, and Born.

No one has ever accused Popper of learning his science through novels and plays. It was Einstein’s 1905 paper on special relativity that prompted him to begin studying the philosophy of science — because of the implications in that paper of what it means to say that two events at different points in space occur simultaneously. He was curious as to how one “verifies” this. As it turns out, he made “falsifiability,” rather than “verifiability,” the cornerstone of his ideas about how science operates.

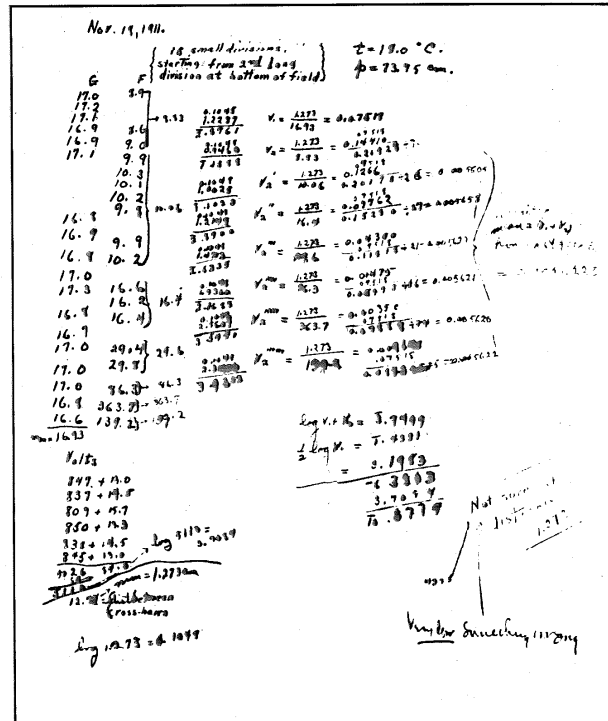
According to Popper, a scientist, whether theorist or experimentalist, puts forward statements (or systems of statements) and tests them step by step. The initial stage — the act of conceiving or inventing a theory — doesn’t interest him because that is a creative act that cannot be analyzed logically. Popper focuses his analysis on the next step, which consists of showing the proposal to be wrong. He says, in effect, that all scientific discoveries are refutations of past theories.

Philosophers say science is something that follows the scientific method. Yet philosophers Bacon and Popper present two opposing views of what the scientific method is. For Bacon, the unprejudiced, systematic observer is led infallibly to generalizations that in turn produce new discoveries. Popper says instead that theories are the product of inspiration, and progress consists in falsifying them by showing their predictions to be wrong. To find out which one of them is right, we turn to a practicing scientist.

*David Goodstein:* Sir Francis and Sir Karl certainly have very different views of how science — and scientists — operate, but we can compare them at two points. We can compare the two views of the process of creating theories, and we can ask what happens once the theory has been created. What does the experimentalist do? I’ll leave the



The instrument Robert Millikan is inspecting in this photograph taken in the 1920s at Caltech was used for spectroscopy and cosmic ray research. The famous Millikan oil drop experiment was performed



much earlier at the University of Chicago. Above is a page from Millikan's lab notebook dated November 18, 1911, and with a comment on the results in the lower righthand corner.

question of how theories are created for Madame, the archivist, to take up later. Right now my job is to discuss what experimentalists do once they have a theory to test.

As my first example, let me tell you about one of our own patron saints, Robert A. Millikan, and his oil drop experiment. Millikan had two things to find out with that experiment — whether the electric charge came in quantized units, and if so what the size of the unit was. And so, good, Baconian, dispassionate observer that he was, he had to go into the laboratory with no preconceived notions, look at his oil drop, make his measurements, and report *all* of the results, which — he says in the *Physical Review* — he did.

Now I've looked at some of Millikan's laboratory notebooks (written only for himself) to see how he worked. On December 20, 1911, he shows his readings — the voltages, the rate at which the drop is falling in gravity, measurements of what happens when the drop is in the field — and then he does his calculations. And what do we find at the bottom? His comments. "This is almost exactly right," he says. "This is the best I ever had." On another day his comment was, "Very low. Something is wrong." Another one says, "Beauty! Publish!"

If this seems shocking, let me assure you I am not trying to tell you that Millikan was being a bad scientist. He was one of the very best scientists. But he was doing what scientists always do when they're in the laboratory,

which is to look for the result that they want. To tell you about that, I'm going to analyze in a very general way a hypothetical, but realistic, experiment.

If you have a liquid like water, it can exist in equilibrium with its own vapor, and the curve along which it exists in that state is called the vapor pressure curve. If you warm the water and steam up, they can still be in equilibrium, but at a higher pressure. At higher pressure the vapor is more dense, and because the temperature is higher, the water is less dense. As you warm the system, the densities of these two fluids get closer and closer, and finally they become equal to each other — at what is called the critical point.

About 20 years ago it was discovered that at the critical point of any substance its heat capacity (the amount of heat put into the system divided by the change in temperature) becomes infinite. Various other properties were also found to behave peculiarly at the critical point, and a theory grew up to explain these so-called critical phenomena.

The theory makes a definite prediction about how the heat capacity becomes infinite. What is important is how far we are from the exact critical temperature (that is, the temperature at the critical point). Suppose we make a measurement near the critical temperature, but at a small temperature difference away from it that we can call  $\Delta T$ . The theory says that, if  $\Delta T$  is sufficiently small and we

plot the log of the heat capacity versus the log of  $\Delta T$ , we will find a straight line. Moreover, if we make these measurements both just above and just below the critical temperature, we should get two straight lines, and they should have the same slope. That is the prediction we are going to set out to prove.

Now, no matter how carefully we do our work, we immediately run into a severe problem. In order to find  $\Delta T$  at each point, we must know not only the temperature at which we are working, but also the critical temperature, to a very high precision. It is not good enough to look the critical temperature up in a book — probably we will need to know it more accurately than it has ever been measured. We must deduce it from our own experiment.

The critical temperature is just the temperature at which the heat capacity we are measuring becomes infinite. Of course, we cannot really measure an infinity in the laboratory, but we can accomplish the same purpose by assembling our data and using them to choose a critical temperature that makes those two curves into parallel straight lines. If we do that, to be sure, we are not really testing the theory, but there may be no other way to do it.

There are other problems as well. The theory only applies, strictly speaking, to a sample of infinite size, in the absence of gravity. We can use the theory itself to correct our data for the size of our sample and the presence of gravity, but those corrections, in more subtle ways, amount to doing the same thing we did with the critical temperature: They make it easier to make the experiment agree with the theory.

There is more we could say, but the point should now be obvious: Experiments don't give clear answers; they are ambiguous, and the art of collecting and interpreting experimental data is subtle and complex.

What happens after we've done our experiment and evaluated our results? Seemingly, the first possibility is that the theory will turn out to be all wrong. We've made our measurements, and it's clear that nothing will make them come out to be two straight parallel lines. Now that's not going to be the outcome, because some of the data were available before the theory was formulated. In fact, the theory grew out of that data, so you know it is approximately true.

The next possibility is that we make our measurements, plot our data, get two straight parallel lines, write the paper, and send it off to be published.

The third possibility is that we go through all of this action — and we don't get two straight parallel lines. At that point we start examining the experiment to find out what went wrong — just as Robert Millikan did. Note that if we do get those two straight parallel lines we don't examine the data to find out what went right. That effect alone builds in a strong bias for the experimenter to get the results he wants.

You may say, "That's ridiculous. We're good Baconians; we go into the laboratory with a clean slate in our

heads. We have unprejudiced minds. We just make our measurements, and nature tells us the results — and that's all we want. Isn't that true?" And the answer is, "Of course not."

To do this experiment, to make these measurements, some guy worked 70 hours a week for a solid year. Did a passion for the dispassionate collection of data drive him to work that hard? It's nonsense to think the human animal works that way. Whatever the motives that drive us to do science, they are not the dispassionate collection of data. It follows from that that the experimenters always want something from their results, and we have to know what that is if we are going to analyze the scientific process in a reasonable way.

I think this is the way it works: When a theory first comes out, the experimenter prefers to confirm that it is correct. The reason for that is very simple. Suppose the theory comes out and you do a brilliant experiment that shows clearly, unambiguously, once and forever that the theory is wrong. Well, that's the end of the story. The theory is gone forever, and so is your experiment. On the other hand, if you show that the theory is right, you've made a contribution to the growth of knowledge, and it will be remembered and be important. So which do you want — to show that it's right or that it's wrong?

What happens over a period of time is that a number of experiments are done showing that the theory is right. After a while, the theory becomes a law of nature, a part of the received wisdom. Now, it would be really exciting to be able to show that it was wrong, to tear down a crusty prejudice standing in the way of new knowledge. If you can do that, you've made a contribution. Furthermore, all of the things that made it possible to show that it was right now make it possible to show that it was wrong.

This is the Popper stage — the stage of falsifiability — when a theory is tested and ultimately found false. Sometimes, however, all attempts to disprove the result fail, and the theory stands — as in the case of the critical point theory I have described.

That's the way I see the scientific method operating experimentally. The other part of the story is how the theories arise. Is it by the Baconian inevitable generalizations from dispassionately gathered facts or by some sort of a mystical act of creation, as Popper thinks? For an answer to that, we turn to the archivist, who will give us a historical example.

*Judith Goodstein:* We've talked about what two philosophers say the scientific method is, and Monsieur, the physicist, has told us how it works today — which seems to indicate that neither Bacon nor Popper is all-powerful in the minds of 20th-century scientists. How did it work in a simpler, more classical past, around the turn of the 19th century, for example?

In 1807 the English chemist Sir Humphry Davy, who had impeccable scientific credentials, isolated the chemical



Sir Humphry Davy

elements potassium and sodium and, in 1808, the metals of the alkaline earths, barium and calcium. Two years later he established the elemental nature of chlorine and predicted the existence of fluorine. Soon after, he and Gay-Lussac established iodine as a third halogen. He also did extensive and diverse other research, but did he have a scientific method?

To his public he certainly preached the methods of Bacon. "The legitimate practice, that sanctioned by the precepts of Bacon," he said, "is to proceed from particular instances to general ones, and to found hypotheses upon facts to be rejected or adopted as they are contradictory or conformable to new discoveries."

What about the genesis of Davy's ideas? Did he follow his own advice? The answer is, almost never. He held many unorthodox ideas, and he clung to them tenaciously. He embraced theories not in vogue. He speculated, for example, about the composition of ammonia and water, two compounds considered by his fellow chemists to be well established experimentally. He also argued that the chemical elements were not the simplest obtainable units of matter, and he spent many research hours trying to decompose nitrogen into a metallic base and oxygen.

This was a period of time in which the ideas of Lavoisier and Dalton dominated chemistry. Lavoisier believed that he could systematize chemistry around a simple

principle, namely oxygen, which would explain all chemical phenomena. He divided substances into elements and compounds, defining an element as any body that could not be further decomposed. He assumed, further, that elements maintained their individuality when they combined to form compounds.

Davy, on the contrary, believed that a few fundamental particles of matter composed all the simple substances that were commonly called elements. To him, Lavoisier's definition of an element did not offer any clues as to the internal nature of matter. Davy was always careful to distinguish between Lavoisier's "simple bodies" and the "true elements of bodies" — the fundamental particles of matter composing all substances.

Davy also quarreled with the atomic theory of matter proposed by John Dalton. Dalton's theory of matter incorporated the historical idea that the "ultimate particles of matter" were best expressed by the word "atom" because it signified indivisibility. His theory offered an explanation of what was going on when chemical combinations occurred. Although it was an internal explanation of the behavior of matter, it put great stress on the individuality of the elements, whose relative atomic weights were tabulated. The theory, in Davy's opinion, sacrificed the idea of a unity of matter. If there were discrete atoms for each element without any possibility of their further reduction, then the explanation of the properties of substances required as many different kinds of atoms as the number of known elements. And that number kept growing. Between 1800 and 1812 chemists added 15 new elements to the list of 18 previously known. Davy viewed this trend with alarm. Dalton's use of the term "element" in his theory precluded Davy's dream of a "real indestructible principle" of matter ever being realized.

Davy's speculations turned on his assumption that the elements of Lavoisier and Dalton were complex bodies. His announcement of the discovery of potassium and sodium was coupled with what he called a "phlogistic" theory of matter because he thought this theory better expressed his belief that the metals all contained a common substance and because he sought a simple system of chemistry. Davy's assumptions belie all of his Baconian admonitions about the role that facts play in advancing the progress of a science. The adoption of a theory which assumes that the elements are *not* simple is Davy's first requirement for chemistry's inclusion as a "true science."

Many of the experiments Davy performed after 1806 bore the mark of his search for the few fundamental particles that compose all matter. His researches were not based on random analogies. The analogies were inspired by the idea that speculations about the unity of matter must be translated into laboratory experiments.

The fact is that Lavoisier and Dalton as well as Davy indulged in speculation. All three chemists paid lip service to the Baconian idea that the scientist does not start his research by speculating and forming hypotheses. Yet, judg-

ing from the historical record, each of them allowed his work to be guided by unsupported philosophical assumptions about nature.

Well, where does that leave us? Is there a scientific method? For an answer to that question, we turn again to the physicist.

*David Goodstein:* It seems to me that a number of questions arise from this discussion. The first is, what would be the purpose of a theory of the scientific method that we really believed in? What would we use it for?

It certainly is not needed by the scientists. Nobody needs to tell them what to do when they go into the laboratory. They may give lip service to Bacon or even to Popper, but they don't really pay any attention to them because they know exactly what they want to do.

One purpose for which a theory of the scientific method is used is as an objective test to distinguish between sciences and pseudosciences. It gives us a way of ruling out astrology, for example, as a candidate for being a science. Of course, we don't really need it to rule out astrology as a science, but such a test becomes significant when it is applied to marginal cases as, for example, psychoanalysis. There was a seminar here on campus a couple of years ago at which a philosopher of science discussed whether psychoanalysis is a science. His analysis was purely in terms of Popper; that is, it is a science if it makes falsifiable predictions and tests them experimentally. It is not a science if it doesn't do that.

By contrast, we have the case of physics, which is treated differently from the pretenders to science. An excellent example is the critical point theory we discussed earlier. It is intrinsically unfalsifiable because we are not told just how small  $\Delta T$  must be before the theory becomes valid. In a marginal science, the philosophers would rule out such a theory as being unscientific. But the physicists had no doubt they were dealing with good physics, and they proceeded to incorporate it into their body of knowledge.

I think this fact points out that the philosophers are really saying, "Science is what physics does, and other things are sciences to the extent that they do the same things that physics does. So we should figure out what physics does, and then present it so that other things can imitate it and thereby become sciences." It shouldn't need very much thought to see that it would be destructive to another field for it to force itself to follow a methodology that is nothing but a mistaken notion of what physics does.

The second question that arises is, if we don't believe Popper or Bacon gave us the true scientific method, does such a thing exist? I suppose the answer to that is yes. Furthermore, Bacon and Popper each have a piece of the truth, I think, but neither of them has cornered the market, as he thinks he has.

Something identifiable does go on in science by which some sort of empirical information gets put together; in some way theories or generalizations are formulated; and

they are then tested in some way by experiments that either prove them to be wrong — that is, falsify them — or lead to new discoveries and deeper and larger theories. That's the way science works, but that really describes everything from solid-state physics to French cooking. Is there a more precise, objective criterion that distinguishes between sciences and pseudosciences?

I think there is, and it is this: What characterizes a science is the absolute unswerving belief by its practitioners that they are dealing with laws of nature. I don't mean approximate truths or generally true things; I mean hard, real laws, whose consequences arise from direct causal links. The consequence of that belief is the conviction that if you did an experiment somebody else could repeat it and get the same result.

We've already seen that the result of an experiment is so subtle that its repeatability is in some doubt. Yet the scientist must believe in repeatability because that belief gives science its integrity. One reason a scientist doesn't cheat is that cheating obviously doesn't pay if somebody else can easily find him out by duplicating his experiment.

Let me repeat — the thing that keeps this whole balky, complicated machine on its tracks is the absolute unswerving belief by every practitioner that there are laws to be discovered. Now any philosopher can prove to you that there is no way of distinguishing between laws as constructs of the human intellect and laws that exist objectively in nature. Nevertheless, every scientist must believe in the depths of his soul that those laws exist and that the results of his experiment arise from those laws by direct causal chains that can't be broken.

I think that belief in those laws exists in physics and chemistry and in some other disciplines in which the objects of study are vastly more complex and less well understood than they are in physics and chemistry — biology, for example. Those are real sciences. On the other hand, I don't think those laws exist, for example, in psychoanalysis. Far more importantly, I don't think that the practitioners of psychoanalysis believe that they are dealing with hard laws that are connected by direct causal links to the results of what they do. And because they don't believe it, regardless of what methodology they use, what they do is not a science.

Much the same can be said for most of what we call the "social sciences," which does not mean that they are useless or unimportant. It just means they should not try to succeed by imitating physics. Real scientists sometimes pretend to have followed one or another "scientific method." In other fields, practitioners sometimes actually modify their methods in the hope of satisfying the philosophers and thereby being accepted as real scientists. All of this does little good and may sometimes do real harm. The huge success of the scientific enterprise is not due to its method but rather to the fact that its methods match its substance. There is no magical prescription for other fields of knowledge to be as successful. □