

# UC Irvine

## UC Irvine Previously Published Works

**Title**

ON THE SCIENTIFIC METHOD, ITS PRACTICE AND PITFALLS

**Permalink**

<https://escholarship.org/uc/item/0bk6922j>

**Journal**

HISTORY AND PHILOSOPHY OF THE LIFE SCIENCES, 16(2)

**ISSN**

0391-9714

**Author**

AYALA, FJ

**Publication Date**

1994

**Copyright Information**

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



---

## On the Scientific Method, Its Practice and Pitfalls

Author(s): Francisco J. Ayala

Source: *History and Philosophy of the Life Sciences*, Vol. 16, No. 2 (1994), pp. 205-240

Published by: Stazione Zoologica Anton Dohrn - Napoli

Stable URL: <https://www.jstor.org/stable/23331738>

Accessed: 24-07-2018 19:06 UTC

## REFERENCES

Linked references are available on JSTOR for this article:

[https://www.jstor.org/stable/23331738?seq=1&cid=pdf-reference#references\\_tab\\_contents](https://www.jstor.org/stable/23331738?seq=1&cid=pdf-reference#references_tab_contents)

You may need to log in to JSTOR to access the linked references.

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

*Stazione Zoologica Anton Dohrn - Napoli* is collaborating with JSTOR to digitize, preserve and extend access to *History and Philosophy of the Life Sciences*

## On the Scientific Method, Its Practice and Pitfalls

Francisco J. Ayala

*Department of Ecology and Evolutionary Biology  
University of California  
Irvine, CA 92717, USA*

**ABSTRACT** – This paper sets forth a familiar theme, that science essentially consists of two interdependent episodes, one imaginative, the other critical. Hypotheses and other imaginative conjectures are the initial stage of scientific inquiry because they provide the incentive to seek the truth and a clue as to where to find it. But scientific conjectures must be subject to critical examination and empirical testing. There is a dialogue between the two episodes; observations made to test a hypothesis are the inspiration for new conjectures. Inductive generalizations may also inspire hypotheses, but cannot validate them.

A hypothesis is empirically tested by ascertaining whether or not predictions about the world of experience deduced from the hypothesis agree with what is actually observed. This has been appropriately considered the 'criterion of demarcation' that distinguishes science from other knowledge. But scientific hypotheses must satisfy other tests as well, e.g., whether they have explanatory value and further understanding. I briefly explore such issues as verifiability and falsifiability, empirical content and truthfulness, contingency and certainty, fact and theory, error and fraud.

Science like any human activity is subject to error and to the foibles and other failings of human beings. But severe attempts of empirical falsification and other trials yield knowledge that stands the test of time and provides a foothold for further knowledge. Moreover, scientists have developed social mechanisms, such as peer review and publication, to evaluate their work. Because the research of scientists depends on the validity of previous knowledge, it is of great consequence that they discern valid from invalid knowledge and thus scientists are inclined to transcend ideology, nationality, friendship, monetary interest and other prejudices when the mettle of scientific knowledge is at stake.

I use historical examples to illustrate some relevant aspects of scientific practice: its success (Mendel), misrepresentation (Darwin), ideological abuse (Lysenko), arrogant violation of the requirement of testing (Koch), theory replacement (Priestly and Lavoisier, Newton and Einstein), and the indispensability of context (Oswald Avery and Alfred Wegener).

### *Introduction*

Knowledge derives from many sources. Examples include common sense experience, artistic expression, and philosophical reflection. Scientific knowledge, however, stands apart as special. The tremendous success of science as a mode of inquiry into the nature of the universe is a matter of wonderment. The technology derived from scientific knowledge is equally wondrous: the high-rise buildings of our cities, thruways and long span-bridges, rockets that bring men to the moon, telephones that provide instant communication across continents, computers that perform

complex calculations in millionths of a second, vaccines and drugs that keep bacterial parasites at bay, gene therapies that replace DNA in defective cells. All these remarkable achievements bear witness to the validity of the scientific knowledge from which they originated. No other kind of knowledge affects human life so pervasively and drastically.

Scientific knowledge is also remarkable in the way it emerges by way of consensus and agreement among scientists, and in the way new knowledge builds upon past accomplishment rather than starting anew with each generation or each new practitioner. Surely scientists disagree with each other on many matters; but these are issues not yet settled, and the points of disagreement generally do not bring into question previous knowledge. Modern scientists do not challenge that atoms exist, or that there is a universe with a myriad stars, or that heredity is encased in the DNA. Scientists differ from philosophers, who interminably debate the questions they seek to answer. Philosophers today focus on the same questions that were debated in Antiquity or in the Middle Ages or two decades ago, without ever coming to any definitive agreement. Not so with scientists, who build upon matters resolved in the past in order to formulate new questions and resolve them. Nor is there among scientists anything like the radically disparate and irreconcilable views held by different religions; or the ever-changing means of artistic expression.

What is it, then, that makes scientific knowledge different from all other activities by which we learn about the universe and about ourselves? In this paper, I approach the matter by first identifying some distinguishing traits of scientific knowledge. I then explain why science involves much more than simple inductive reasoning, and discuss the hypothetico-deductive method as a paradigm for understanding some distinctive features of the way in which scientists proceed in order to understand the world. I next consider the demarcation question, or how to distinguish valid from invalid scientific claims, and the social mechanisms by which scientific practice weeds out valid from invalid science. I use historical examples to illustrate relevant aspects of how scientific knowledge develops and how demarcation works in practice.

### *Science's Distinctive Traits*

Three characteristic traits jointly distinguish scientific knowledge from other forms of knowledge.<sup>1</sup> First, science seeks the systematic

<sup>1</sup> F.J. Ayala, 'Biology as an Autonomous Science', *American Scientist* 56 (1968), p. 207-221. See E. Nagel, *The Structure of Science*, New York: Harcourt, Brace, and World, 1961: 388-345.

organization of knowledge about the world. Common sense, like science, provides knowledge about natural phenomena, and this knowledge is often correct. For example, common sense tells one that children resemble their parents and that good seeds produce good crops. Common sense, however, shows little interest in systematically establishing connections between phenomena that do not appear to be obviously related. By contrast, science is concerned with formulating general laws and theories that manifest patterns of relations between very different kinds of phenomena. Science develops by discovering new relationships, and particularly by integrating statements, laws, and theories, which previously seemed to be unrelated, into more comprehensive laws and theories.

Second, science strives to explain *why* observed events do in fact occur. Although knowledge acquired in the course of ordinary experience is frequently accurate, it seldom provides explanations of why phenomena occur as they do. Practical experience tells us that children resemble one parent in some traits and the other parent in other traits, or that manure increases crop yield. But it does not provide explanations for these phenomena. Science, on the other hand, seeks to formulate explanations for natural phenomena by identifying the conditions that account for their occurrence.

Seeking the systematic organization of knowledge and trying to explain why events are as observed are two characteristics that distinguish science from common-sense knowledge. But these characteristics are also shared by other forms of systematic knowledge, such as mathematics and philosophy. A third characteristic of science, and the one that distinguishes the empirical sciences from other systematic forms of knowledge, is that scientific explanations must be formulated in such a way that they can be subjected to empirical testing, a process that must include the possibility of *empirical falsification*. Falsifiability has been proposed as the criterion of demarcation that sets science apart from other forms of knowledge.<sup>2</sup>

New ideas in science are advanced in the form of hypotheses. The tests to which scientific ideas are subjected include contrasting hypotheses with the world of experience in a manner that must leave open the possibility that anyone might reject any particular hypothesis if it leads to wrong predictions about the world of experience. The possibility of empirical falsification of a hypothesis is carried out by ascertaining whether or not precise predictions derived as logical con-

<sup>2</sup> K.R. Popper, *The Logic of Scientific Discovery*, London: Hutchinson, 1959: 40-42.

sequences from the hypothesis agree with the state of affairs found in the empirical world. A hypothesis that cannot be subject to the possibility of rejection by observation and experiment cannot be regarded as scientific.

I shall return below to this matter of 'empirical falsifiability', as the criterion of demarcation that sets apart science from other forms of knowledge. For now, I'll summarize my discussion of the nature of science by defining science as 'knowledge about the universe in the form of explanatory principles supported by empirical observation and subject to the possibility of empirical falsification'. Another definition would be the following: 'Science is an exploration of the material universe that seeks natural, orderly relationships among observed phenomena and that is self-testing'.<sup>3</sup> Many other definitions can be proposed, but seeking a 'perfect' definition is a futile endeavor. Science is a complex enterprise that cannot be adequately captured in a compact statement. In any case, my goal here is not so much to provide an adequate definition as it is to identify the traits that distinguish scientific knowledge. I will proceed by discussing first 'the method of induction', which is sometimes said to be the method followed by scientists. I will explain that induction is not a method by which we may establish the *validity* of scientific knowledge (although it is often a process by which we come upon new ideas, but this is an altogether different matter).

### *Induction in Science*

It is a common misconception that science advances by 'accumulating experimental facts and drawing up a theory from them'.<sup>4</sup> This misconception is encased in the much repeated assertion that science is inductive, a notion which can be traced to the English statesman and essayist Francis Bacon (1561-1626). Bacon had an important and influential role in shaping modern science by his criticism of the prevailing metaphysical speculations of medieval scholastic philosophers. In the nineteenth century the most ardent and articulate proponent of inductivism was John Stuart Mill (1806-1873), an English philosopher and economist.

<sup>3</sup> G.G. Simpson, *This View of Life*, New York: Harcourt, Brace, and World, 1964: p. 91.

<sup>4</sup> F. Jacob, *The Statue Within: An Autobiography*, New York: Basic Books, 1988: 224-225. Jacob writes that he started his scientific research under the naive misconception that science proceeds by induction, but soon realized that what was going on in a laboratory was quite different. See below.

Induction was proposed by Bacon and Mill as a method of achieving *objectivity* while avoiding subjective preconceptions, and of obtaining *empirical* rather than abstract or metaphysical knowledge. In its extreme form this proposal would hold that a scientist should observe any phenomena that he encounters in his experience, and record them without any preconceptions as to what to observe or what the truth about them might be. Truths of universal validity are expected eventually to emerge, as a result of the relentless accumulation of unprejudged observations. The methodology proposed may be exemplified as follows. A scientist measuring and recording everything that confronts him observes a tree with leaves. A second tree, and a third, and many others, are all observed to have leaves. Eventually, he formulates a universal statement, 'All trees have leaves'.

The inductive method fails to account for the actual process of science. First of all, no scientist works without any preconceived plan as to what kind of phenomena to observe. Scientists choose for study objects or events that, in their opinion, are likely to provide answers to questions that interest them. Otherwise, as Darwin (1903) wrote, 'one might as well go into a gravel-pit and count the pebbles and describe the colours'.<sup>5</sup> A scientist whose goal was to record carefully every event observed in all waking moments of his life would not contribute much to the advance of science; more likely than not, he would be considered mad by his colleagues.

Moreover, induction fails to arrive at universal truths. No matter how many singular statements may be accumulated, no universal statement can be logically derived from such an accumulation of observations. Even if all trees so far observed have leaves, or all swans observed are white, it remains a logical possibility that the next tree will not have leaves, or the next swan will not be white. The step from numerous singular statements to a universal one involves logical amplification. The universal statement has greater logical content – it says more – than the sum of all singular statements.

Another serious logical difficulty with the proposal that induction is 'the' method of science, is that scientific hypotheses and theories are formulated in abstract terms that do not occur at all in the description of empirical events. Mendel, the founder of genetics, observed in the progeny of hybrid plants that alternative traits segregated according to certain proportions. Repeated observations of these proportions could never have led inductively to the formulation of his hypothesis

<sup>5</sup> F. Darwin, *More Letters of Charles Darwin*, 2 volumes, London: Murray, 1903, vol. 1, p. 195.

that 'factors' (genes) exist in the sex cells and are rearranged in the progeny according to certain rules. The genes were not observed, and thus could not be included in statements reflecting what Mendel observed. The most interesting and fruitful scientific hypotheses are not simple generalizations. Instead, scientific hypotheses are creations of the mind, imaginative suggestions as to what might be true.

Induction fails in all three counts pointed out. It is not a method that insures objectivity and avoids preconceptions, it is not a method to reach universal truth, and it is not a good description of the process by which scientists formulate hypotheses and other forms of scientific knowledge. It is a different matter that a scientist may come upon a new idea or develop a hypothesis as a consequence of repeated observation of phenomena that might be similar or share certain traits. But how we come upon a new idea is quite a different matter from how is it that we come to accept something as established scientific knowledge. I shall come back to this point in a moment.

### *The Hypothetico-Deductive Method*

I have already stated that the validity of scientific ideas ('hypotheses') is established by deriving ('deduction') their consequences as to what should be the case in the real world, and then proceeding to ascertain whether or not the derived prediction is correct. (It is of the essence of the process, as I shall explain, that whether such consequences are the case not be already known if the observation of such consequences is to serve the purpose of validating the idea; it is also required that the consequences be unlikely). The scientific method is, accordingly, said to be *hypothetico-deductive*.

The analysis of the hypothetico-deductive method may be traced to William Whewell (1794-1866) and William Stanley Jevons (1835-1882) in Great Britain, and to Charles S. Peirce (1838-1914) in the United States. The key features of the hypothetico-deductive method have been well characterized by Karl R. Popper<sup>6</sup> and C.G. Hempel.<sup>7</sup> Scientists, of course, practiced the hypothetico-deductive method long before it was adequately defined by philosophers. Eminent practitioners of the method include Blaise Pascal (1623-1662) and Isaac New-

<sup>6</sup> *The Logic of Scientific Discovery* (footnote 2); also, K.R. Popper, *Conjectures and Refutations: The Growth of Scientific Knowledge* (Routledge and Kegan Paul, London, 1963).

<sup>7</sup> C.G. Hempel, *Aspects of Scientific Explanation*, New York: Free Press, 1965.



ton (1624-1727) in the seventeenth century and, among nineteenth-century biologists, Claude Bernard (1813-1878) and Louis Pasteur (1822-1895) in France, Charles Darwin (1809-1882) in England, and Gregor Mendel (1822-1884) in Austria. These and other successful scientists practiced the hypothetico-deductive method even if some of them, Darwin for example, claimed to be inductivists in order to conform to the claims of contemporary philosophers.

Here is how the Nobel Laureate François Jacob, in his autobiography, describes research at the Pasteur Institute in Paris that led in the 1950's to one of the fundamental discoveries of molecular biology:

What had made possible analysis of bacteriophage multiplication, and understanding of its different stages, was above all the play of hypotheses and experiments, constructs of the imagination and inferences that could be drawn from them. Starting with a certain conception of the system, one designed an experiment to test one or another aspect of this conception. Depending on the results, one modified the conception to design another experiment. And so on and so forth. That is how research in biology worked. Contrary to what once I thought, scientific progress did not consist simply in observing, in accumulating experimental facts and drawing up a theory from them. It began with the invention of a possible world, or a fragment thereof, which was then compared by experimentation with the real world. And it was this constant dialogue between imagination and experiment that allowed one to form an increasingly fine-grained conception of what is called reality.<sup>8</sup>

Science is a complex enterprise that essentially consists of two interdependent episodes, one imaginative or creative, the other critical. To have an idea, advance a hypothesis, or suggest what might be true is a creative exercise. But scientific conjectures or hypotheses must also be subject to critical examination and empirical testing. Scientific thinking may be characterized as a process of invention or discovery followed by validation or confirmation. One process concerns the formulation of new ideas ('acquisition of knowledge'), the other concerns their validation ('justification of knowledge').

Scientists like other people come upon new ideas, *acquire* knowledge, in all sorts of ways: from conversation with other people, from reading books and newspapers, from inductive generalizations, and even from dreams and mistaken observations. Newton is said to have been inspired by a falling apple. Kekulé had been unsuccessfully attempting to devise a model for the molecular structure of benzene. One evening he was dozing in front of the fire. The flames

<sup>8</sup> F. Jacob, *The Statue Within* (footnote 4), pp. 224-225.

appeared to Kekulé as snake-like arrays of atoms. Suddenly one snake appeared to bite its own tail and then whirled mockingly in front of him. The circular appearance of the image inspired in him the model of benzene as a hexagonal ring. The model to explain the evolutionary diversification of species came to Darwin while riding in his coach and observing the countryside.

I can remember the very spot in the road ... when to my joy the solution came to me ... The solution, as I believe, is that the modified offspring ... tend to become adapted to many and highly diversified places in the economy of nature.<sup>9</sup>

Hypotheses and other imaginative conjectures are the initial stage of scientific inquiry. It is the imaginative conjecture of what might be true that provides the incentive to seek the truth and a clue as to where we might find it.<sup>10</sup> Hypotheses guide observation and experiment because they suggest what to observe. The empirical work of scientists is guided by hypotheses, whether explicitly formulated or simply in the form of vague conjectures or hunches about what the truth might be. But imaginative conjecture and empirical observation are mutually interdependent episodes. Observations made to test a hypothesis are often the inspiring source of new conjectures or hypotheses. As described by Jacob, the results of an experiment often inspire the modification of a hypothesis and the design of new experiments to test it.<sup>11</sup>

The starting point of scientific inquiry is the conception of an idea, a process that is, however, not a subject of investigation for logic or epistemology. The complex conscious and unconscious events underlying the creative mind are properly the interest of empirical psychology. The creative process is not unique to scientists. Philosophers as well as novelists, poets, and painters are also creative; they too advance models of experience and they also generalize by induction. What distinguishes science from other forms of knowledge is the process by which this knowledge is justified or validated.

<sup>9</sup> Charles Darwin, *The Autobiography of Charles Darwin (1809-1882)*, Nora Barlow (ed.), London: Collins, 1958: 120-121.

<sup>10</sup> See P.B. Medawar, *The Art of the Soluble*, London: Methuen, 1967. This small book provides a very eloquent, yet profound, discussion of the scientific method as a dialogue between the two essential episodes of science: conjectures and refutations. My discussion of this subject is importantly derived from Medawar's.

<sup>11</sup> Footnote 8.

*The Criterion of Demarcation*

Testing a hypothesis (or theory) involves at least four different activities. First, the hypothesis must be examined for internal consistency. A hypothesis that is self-contradictory or not logically well-formed in some other way should be rejected.

Second, the logical structure of the hypothesis must be examined to ascertain whether it has explanatory value, i.e., whether it makes the observed phenomena intelligible in some sense, whether it provides an understanding of why the phenomena do in fact occur as observed. A hypothesis that is purely tautological should be rejected because it has no explanatory value. A scientific hypothesis identifies the conditions, processes, or mechanisms that account for the phenomena it purports to explain. Thus, hypotheses establish general relationships between certain conditions and their consequences or between certain causes and their effects. For example, the motions of the planets around the sun are explained as a consequence of gravity, and respiration as an effect of red blood cells that carry oxygen from the lungs to various parts of the body.

Third, the hypothesis must be examined for its consistency with hypotheses and theories commonly accepted in the particular field of science, or to see whether it represents any advance with respect to well-established alternative hypotheses. Lack of consistency with other theories is not always ground for rejection of a hypothesis, although it will often be. Some of the greatest scientific advances occur precisely when it is shown that a widely-held and well supported hypothesis is replaced by a new one that accounts for the same phenomena that were explained by the preexisting hypothesis, as well as other phenomena it could not account for. One example is the replacement of Newtonian mechanics by the theory of relativity, which rejects the conservation of matter and the simultaneity of events that occur at a distance – two fundamental tenets of Newton's theory.<sup>12</sup>

Examples of this kind are pervasive in rapidly advancing disciplines, such as molecular biology at present. The so-called 'central dogma' holds that molecular information flows only in one direction, from DNA to RNA to protein. The DNA contains the genetic information that determines what the organism is, but that information has to be expressed in enzymes (a particular class of proteins) that guide all chemical processes in cells. The information contained in the DNA

<sup>12</sup> Joseph Schwartz, *The Creative Moment*, New York: Harper Collins, 1992, chapters 1 and 2.

molecules is conveyed to proteins by means of intermediate molecules, called messenger RNA. David Baltimore and Howard Temin were awarded the Nobel Prize for discovering that information could flow in the opposite direction, from RNA to DNA, by means of the enzyme reverse transcriptase. They showed that some viruses, as they infect cells, are able to copy their RNA into DNA, which then becomes integrated into the DNA of the infected cell, where it is used as if it were the cell's own DNA.<sup>13</sup>

Other examples are the following. Until very recently, it was universally thought that only the proteins known as enzymes could mediate (technically 'catalyze') the chemical reactions in cells. However, Thomas Cech and Sidney Altman received in 1989 the Nobel Prize for showing that certain RNA molecules act as enzymes and catalyze their own reactions.<sup>14</sup> One more example concerns the so-called 'co-linearity' between DNA and protein. It was generally thought that the sequence of nucleotides in the DNA of a gene is expressed consecutively in the sequence of aminoacids in the protein. This conception was shaken by the discovery that genes come in pieces, separated by intervening DNA segments that do not carry genetic information; Richard Roberts and Philip Sharp received the 1993 Nobel Prize for this discovery.<sup>15</sup>

These revolutionary hypotheses were published after their authors had subjected them to severe empirical tests. Theories that are inconsistent with well-accepted hypotheses in the relevant discipline are likely to be ignored when they are not availed by convincing empirical evidence. The microhistory of science is littered with farfetched or *ad hoc* hypotheses, often proposed by individuals with no previous or posterior scientific achievements. Theories of this sort usually fade away because they are ignored by most of the scientific community, although on occasion they engage their interest because the theory may have received attention from the media or even from political or religious bodies. The flop over 'cold fusion' is an example of an unlikely and poorly tested hypothesis that received some attention from the scientific community because its proponents were well-established scientists.<sup>16</sup>

<sup>13</sup> H.M. Temin, S. Mizutani, 'RNA-Dependent DNA Polymerase in Virions of Rous Sarcoma Virus', *Nature* 226 (1970), 1211; D. Baltimore, 'Viral RNA-Dependent DNA Polymerase in Virions of RNA Tumor Viruses', *Nature* 226 (1970), 1209.

<sup>14</sup> T.R. Cech, 'Self-splicing DNA: Implications for Evolution', *Inter. Rev. Cytol.* 93 (1985), 3.

<sup>15</sup> See, e.g., F. Crick, 'Split Genes and RNA Splicing', *Science* 204 (1979), 264-271; P. Chambon, 'Split Genes', *Scientific American* 244 (1981), 60-71.

<sup>16</sup> The hapless protagonists of the cold fusion fiasco are Martin Fleischmann and B. Stanley Pons. The tale is well told by Gary Taubes, *Bad Science: The Short Life and Weird Times of Cold Fusion*, New York: Random House, 1993.

The fourth and most distinctive test consists of putting on trial an empirically scientific hypothesis by ascertaining whether or not predictions about the world of experience derived as logical consequences from the hypothesis agree with what is actually observed. This is the critical element that distinguishes the empirical sciences from other forms of knowledge: the requirement that scientific hypotheses be empirically falsifiable. Scientific hypotheses cannot be consistent with all possible states of affairs in the empirical world. A hypothesis is scientific only if it is consistent with some but not with other possible states of affairs not yet observed in the world, so that it may be subject to the possibility of falsification by observation. The predictions derived from a scientific hypothesis must be sufficiently precise that they limit the range of possible observations with which they are compatible. If the results of an empirical test agree with the predictions derived from a hypothesis, the hypothesis is said to be provisionally corroborated; otherwise it is falsified.

The requirement that a scientific hypothesis be falsifiable has been appropriately called the *criterion of demarcation* of the empirical sciences because it sets apart the empirical sciences from other forms of knowledge. A hypothesis that is not subject to the possibility of empirical falsification does not belong in the realm of science.<sup>17</sup>

### *Verifiability and Falsifiability*

The requirement that scientific hypotheses be falsifiable rather than simply verifiable seems surprising at first. It might seem that the goal of science is to establish the 'truth' of hypotheses rather than attempt to falsify them, but it is not so. There is an asymmetry between the falsifiability and the verifiability of universal statements that derives from the logical nature of such statements. A universal statement can be shown to be false if it is found inconsistent with even one singular statement, i.e., a statement about a particular event. But, as I pointed out above in the discussion of induction, a universal statement can never be proven true by virtue of the truth of particular statements, no matter how numerous these may be.

Consider a particular hypothesis from which a certain consequence is logically derived. Consider the argument: If the hypothesis is true, then the specific consequence must also be true; it is the case that the

<sup>17</sup> K.R. Popper (footnotes 2 and 6).

consequence is true; therefore the hypothesis is true. This is an erroneous kind of inference called by logicians the 'fallacy of affirming the consequent'. The error of this kind of inference may be illustrated with the following trivial example: If apples are made of iron, they should fall on the ground when they are cut off a tree; apples fall when they are cut off; therefore, apples are made of iron. The conclusion is invalid even if both premises are true. The reason is that there may be some other explanation or hypothesis from which the same consequences or predictions are derived. The observed phenomena are true because they are consequences from this different hypothesis, rather than from the one used in the deduction.

The proper form of logical inference for conditional statements is what logicians call the *modus tollens* (= manner of taking away). It may be represented by the following argument. If a particular hypothesis is true, then a certain consequence must also be true; but evidence shows that the consequence is not true; therefore the hypothesis is false. By way of simple example, consider the following argument. If apples are made of iron, they will sink in water; they do not sink, therefore they are not made of iron. The *modus tollens* is a logically conclusive form of inference. If both premises are true, the conclusion falsifying the hypothesis necessarily follows.

It follows from this reasoning that it is possible to show the falsity of a universal statement concerning the empirical world; but it is never possible to demonstrate conclusively its truth. This asymmetry between verification and falsification is recognized in the statistical methodology of testing hypotheses. The hypothesis subject to test, the *null hypothesis*, may be rejected if the observations are inconsistent with it. If the observations are consistent with the predictions derived from the hypothesis, the proper conclusion is that the test has failed to falsify the null hypothesis, not that its truth has been established.

The requirement that scientific hypotheses be falsifiable also has a parallel in statistical inference, namely in the demand that the power of the test be greater than zero. Statisticians recognize two kinds of errors: a Type I error, the probability of rejecting the null hypothesis when it is true, usually represented as  $\alpha$  and a Type II error, the probability of not rejecting the hypothesis when it is false, symbolized as  $\beta$ . Scientists pay considerable attention to Type I errors and thus choose  $\alpha$  levels sufficiently low, but pay less attention to Type II errors. Yet the power of the test depends on the probability,  $1-\beta$ , of rejecting the null hypothesis when it is wrong. Thus, small levels for both  $\alpha$  and  $\beta$  are desirable. Although for any given test the magni-

tudes of  $\alpha$  and  $\beta$  are inversely related, the value of  $\beta$  may be reduced by increasing the sample size or the number of replications in a test.

### *Empirical Content or 'Truthfulness'*

Tests of a scientific hypothesis must have a positive probability of resulting in the rejection of the hypothesis if this is false. A scientific hypothesis divides all particular statements of fact into two subclasses. First, we have the class of all statements with which it is inconsistent; this is the class of the 'potential falsifiers' of the hypothesis. Second, there is the class of all statements that the hypothesis does not contradict, the class of 'permitted' statements. A hypothesis is scientific only if the class of its potential falsifiers is not empty, because the hypothesis makes empirically meaningful assertions only about its potential falsifiers – it asserts that they are false. 'Not for nothing do we call the laws of nature "laws": the more they prohibit the more they say'.<sup>18</sup>

The empirical or information content of a hypothesis (the 'truthfulness' conveyed by a scientific statement) is measured by the class of its potential falsifiers. The larger this class, the greater the information content of the hypothesis. A hypothesis asserts that its potential falsifiers are false; if any of these is true, the hypothesis is proven false. A hypothesis or theory consistent with all possible states of affairs in the natural world (e.g., 'birds have wings because God made them so; fish do not for the same reason') lacks empirical content and hence is not scientific.

### *Contingency and Certainty in Science*

Scientific hypotheses can only be accepted contingently, since their truth can never be conclusively established. This does not mean that we have the same degree of confidence in all hypotheses that have not yet been falsified. A hypothesis that has passed many empirical tests may be said to be 'proven' or 'corroborated'. The degree of corroboration is not simply a matter of the number of tests, but rather their severity. Severe tests are precisely those that are very likely to have outcomes incompatible with the hypothesis if the hypothesis is false.

<sup>18</sup> K.R. Popper (footnote 2), pp. 40-42, 91-92, 119-121.

The more precise the predictions being tested, the more severe the test. A so-called critical or crucial test is an experiment for which competing hypotheses predict alternative, mutually exclusive outcomes. A critical test thus will corroborate one hypothesis and falsify the others.

One example is the experiment by Matthew Meselson and Franklin Stahl<sup>19</sup> testing the double helix model of DNA proposed by James Watson and Francis Crick<sup>20</sup> that marks the beginning of molecular biology, one of the great scientific revolutions of all times. The double-helix model predicts that the replication of DNA is 'semiconservative', that is that each daughter DNA molecule will consist of one parental strand (the conserved strand) and a newly synthesized strand. Two other possible models of DNA replication are (1) the *conserved* model, according to which the parental DNA molecule is fully conserved and the daughter molecule consists wholly of newly synthesized DNA; and (2) the *dispersive* model, according to which both daughter DNA molecules are newly synthesized and the parental molecule becomes degraded into its component fragments (nucleotides), which are then used, together with additional nucleotides, in the synthesis of the daughter DNA molecules.

Meselson and Stahl produced bacteria with heavy nitrogen (the isotope  $^{15}\text{N}$ ) in their DNA; then transferred these bacteria to a medium containing light ( $^{14}\text{N}$ ) nitrogen. They also had a method to determine precisely the density of the DNA in the bacteria. The double-helix model predicted that after one generation of replication all the DNA will be intermediate in density (because one strand of each molecule would have heavy nitrogen and the other strand light nitrogen). This was also predicted by the dispersive model (because each molecule would have about equal number of heavy and light nucleotide components); but not by the conserved model (which predicted that half the DNA molecules would be heavy and half light). The double-helix model predicted that after a second round of replication, half the DNA molecules would be intermediate in density and half would be light. The other two models made different predictions for the second generation molecules. In particular, the dispersive model predicted that all DNA molecules would be identical to one another, with density one quarter of the way between the light and the heavy molecules. (The predictions of the three models were also different for the third

<sup>19</sup> M. Meselson and F. Stahl, 'The Replication of DNA in *Escherichia coli*', *Proc. Natl. Acad. Sci. USA* 44 (1958), 671-682.

<sup>20</sup> J.D. Watson and F.H.C. Crick, 'A Structure for Deoxyribose Nucleic Acid', *Nature* 171 (1953), 964.



and later rounds of DNA replication.) Meselson and Stahl carried out this critical experiment and corroborated the double-helix model and rejected the other two.

The larger the variety of severe tests withstood by a hypothesis, the greater its degree of corroboration. Hypotheses or theories may thus become established beyond reasonable doubt. The double-helix model of DNA, for example, was also corroborated by an experiment performed by J. Herbert Taylor and his colleagues using autoradiographically labelled DNA from plant roots,<sup>21</sup> and by direct microscopic observation of replicating chromosomes (the cell bodies containing the DNA).<sup>22</sup> Since the 1960s the observations and experiments corroborating the double-helix model (and falsifying alternative models) of the DNA are so numerous and consistent as to defy summary even in a book-length discussion.

### *'Fact' and 'Theory' in Scientific Use*

Scientific hypotheses or models that have become established beyond reasonable doubt are sometimes referred to by scientists as 'facts'. For example, the molecular composition of matter, the DNA double-helix, and the evolution of organisms are said to be facts. The theoretical possibility that these and other hypotheses or explanations might be wrong remains as an abstraction, but they have been confirmed in so many ways, and so much knowledge has been built upon hypotheses such as these, that it would be totally unreasonable to expect they will be proved wrong at some future time. We simply do not expect that the sun will stop rising or that snow will melt into something other than water.

Scientists, however, sometimes refer to a well established hypothesis or explanation by calling it a 'theory' or a 'model'. Scientists, for example, speak of the 'molecular theory of matter' or of the 'theory of evolution'. These expressions do not challenge that the knowledge in question is well corroborated. Rather, in scientific usage, the term 'theory' often implies a body of knowledge, a set of interrelated principles and explanations and the facts that support them. Scientific usage differs in this, as in many other cases, from common usage. In

<sup>21</sup> J.H. Taylor, P.S. Woods, and W.L. Hughes, 'The Organization and Duplication of Chromosomes as Revealed by Autoradiographic Studies Using Tritium-Labelled Thymidine', *Proc. Natl. Acad. Sci. USA* 43 (1957), 122.

<sup>22</sup> J. Cairns, 'The Chromosome of *Escherichia coli*', *Cold Spring Harbor Symposia on Quantitative Biology* 28 (1963), 43.

common language, a 'theory' is an imperfect fact, an explanation for which there is little or no evidence – as in 'I have my own theory as to who assassinated President Kennedy'.

### *Error and Fraud in Science*

The procedure by which scientific hypotheses are empirically tested and rejected (the *modus tollens*) is a logically conclusive method – if a necessary consequence of a premise is false, then the premise must also be false. Nevertheless, the process of falsification is subject to human error. It is possible, for example, that an observation or experiment contradicting a hypothesis may have been erroneously performed or erroneously interpreted. Thus, it is often required, particularly in the case of important or well-corroborated hypotheses, that the falsifying observation be repeatable or that other falsifying tests be performed.

The *modus tollens* may also lead to an erroneous conclusion if the prediction tested is not a necessary logical consequence from the hypothesis. The connection between a hypothesis and specific predictions derived from it is often not a simple matter. The logical validity of an inference may depend not only on the hypothesis being tested, but also on other hypotheses, whether explicitly stated or not, as well as on assumptions concerning the particular conditions under which the deduced inferences obtain (boundary conditions). If a particular prediction is falsified, it follows that the hypothesis tested as well as other hypotheses necessarily implied and the boundary conditions cannot all jointly be correct. The possibility exists that one of the subsidiary hypotheses or some assumed condition may be false. Thus, a proper test of a hypothesis assumes (and, in some cases, it tests) the validity of all other hypotheses and conditions involved in the design and performance of the experiment or observation by which the hypothesis is tested.

Erroneous conclusions in science are frequently a consequence of erroneous assumptions in the design or performance of experiments. The erroneous assumptions may be erroneous hypotheses assumed to be correct, or mistakes in the materials or conditions used. One reason why scientists invest so much of their time and effort in the peer review process (see below) is that they want to weed out erroneous hypotheses as well as erroneous procedures.

An experiment (as it might be performed in the laboratory investigating, say, issues on population genetics) may take several months

and require the investment of tens of thousands of dollars in materials, labor, and equipment costs. It is for this reason also that scientists must specify in full detail the materials, conditions, and procedures used in their experiments. In the standard format of a scientific paper, there is a detailed section, often entitled 'Materials and Methods', that follows the introduction setting up the problem, but precedes the presentation of results. Because a scientist's work depends on the validity of the work of others, the scientific profession is self-policing. Surely abuses occur, but usually scientists are the ones who discover the violations of scientific mores. Their stakes are high.

Failure of adequate testing is usually the most flagrant violation, accounting for erroneous scientific conclusions. But whenever these conclusions are of theoretical or practical import, other scientists will perform additional tests and uncover the error. Improper or inadequate testing is sometimes accompanied by other violations of the canons of science. As we shall see below, Robert Koch, the discoverer of the tuberculosis bacillum, took advantage of his considerable prestige to avoid submitting his claim of having found a cure to proper peer review. The proponents of cold fusion made the same error of inadequate testing, but also sought extensive publicity and financial backing by communicating their claims to the media, instead of submitting them to peer review and publication in scientific journals.

Errors in science are not always due to mistaken assumptions, nor are they often fraudulent. There are four stages in what is a continuous progression from unavoidable error to fraud.<sup>23</sup> First, there are 'quirks of nature', events that may happen because of unknown laws of nature, or that come to be, although quite improbable. This situation may be illustrated with an example, which is only a caricature. Assume that a scientist is asked to find out whether heads and tails are equally probable for a particular coin. The scientist throws the coin 20 times, obtains heads every time, and concludes that the coin is biased. Yet this outcome is compatible with a fair coin: the proba-

<sup>23</sup> See National Academy of Sciences, *On Being a Scientist*, by the Committee on the Conduct of Science (National Academy Press, Washington, DC, 1989). According to this document, 'Instances of scientific fraud have received a great deal of public attention in recent years, which may have exaggerated perceptions of its apparent frequency. Over the past few decades, several dozen cases of fraud have come to light in science. These cases represent a tiny fraction of the total output of the large and expanding research community. Of course, instances of scientific fraud may go undetected, or detected cases of fraud may be handled privately within research institutions. But there is a good reason for believing the incidence of fraud in science to be quite low. Because science is a cumulative enterprise, in which investigators test and build on the work of their predecessors, fraudulent observations and hypotheses tend eventually to be uncovered. Science could not be the successful institution it is if fraud were common. The social mechanisms of science, and in particular the skeptical review and verification of published work, act to minimize the occurrence of fraud'.

bility of all 20 throws yielding heads is only one in a million, but it may indeed happen. The example is a caricature, because an experiment so simple should be repeated many more times before reaching any conclusion. The possibility that quirks of nature may occur is one reason why experiments are replicated by scientists.

Errors may also be due to 'honest' mistakes. A scientist may have mistakenly used the wrong material, made the wrong measurement, or assumed the wrong conditions. These errors are usually discovered by repetition. But a scientist does not have unlimited resources or time, so that even the most conscientious scientist can make a mistake. Errors of this kind are corrected when other scientists reproduce the experiments or test the same hypothesis in some other way.

A third source of error is negligence. A scientist may reach the wrong conclusion because of haste, inattention or sloppiness. These and similar faults are violations of the standards expected in science and they are condemned by scientists, even though the erroneous results are not intentional.

Finally, there is outright fraud, when a scientist conceals, modifies, or fabricates the results. This is an even more grievous violation of scientific standards than carelessness and is accordingly penalized when discovered. Sloppiness and fraud can both do countless harm to the scientific enterprise. However, the conclusions based on them are unlikely to persist if they are significant, because other scientists will seek to corroborate or falsify any results of interest. Sometimes the errors will be discovered, at great personal cost, by other scientists who had assumed their validity in performing their own experiments.

### *The Scientific Method in Practice*

The model of scientific practice that I have sketched can be exemplified *ad infinitum* in the history of science. Generally known examples are Galileo's and Newton's experiments demonstrating the laws of motion, Blaise Pascal's measurements of atmospheric pressure, William Harvey's demonstration of the circulation of the blood, Antoine Lavoisier's rejection of the phlogiston theory and demonstration of the existence of oxygen, Louis Pasteur's experiments on fermentation and putrefaction showing that they are caused by living organisms, and many others.<sup>24</sup> I have outlined earlier the experiment of Meselson and Stahl demonstrating that DNA replicates as predict-

<sup>24</sup> See, e.g., M. Goldstein and I.F. Goldstein, *How We Know*, New York: Plenum Press, 1978; and J. Schwartz, *The Creative Moment* (footnote 12).

ed by the double-helix model. The two episodes that characterize scientific knowledge can be seen in every case. The formulation of a daring hypothesis is associated with experiments cleverly designed to falsify the hypothesis if it were not true.

I shall now describe in somewhat more detail another example: Mendel's discovery of the laws of heredity and his formulation of a theory that remains the core of the science of genetics. Mendel's example is telling because it shows the dialogue between hypothesis and experiment. Initial experiments designed to test simple hypotheses (e.g., whether both the maternal and paternal trait are passed on to the progeny), lead to the formulation of new hypotheses (the first and second law of heredity), which are further tested and stimulate a general theory of heredity, which is then subject to critical experiments. It is notable that all of this is accomplished in a single scientific paper, the author of which was an obscure school teacher.

### *A Historical Paradigm: Mendel's Discovery of the Laws of Heredity*

Gregor Mendel (1822-1884) was an Augustinian monk living in the Austrian city of Brunn (now Brno, Czech Republic). He had studied under distinguished scientists at the University of Vienna and become a high school science teacher. Mendel succeeded where better known contemporary scientists and distinguished predecessors had failed: he discovered the laws of inheritance and formulated the theory upon which all of modern genetics is built.

Mendel performed experiments with pea plants and reported his discoveries in a paper published in 1866, 'Experiments in Plant Hybridization', remarkable for his lucid awareness of the requirements of the scientific method.<sup>25</sup> Mendel formulated hypotheses; examined their consistency with previous results; then submitted the hypotheses to severe critical tests and suggested additional tests that might be performed.

Mendel's genius is evident in his recognition of the conditions required to formulate and test a theory of inheritance: different traits in a plant (such as flower color or seed shape) should be considered individually; alternative states of the traits should differ in clear-cut

<sup>25</sup> Mendel's paper has been reprinted in English translation in numerous publications. The one herein used is from E.W. Sinnott, L.C. Dunn, and T. Dobzhansky, *Principles of Genetics*, New York: McGraw Hill, 1958, Appendix, pp. 419-443. A short biography of Mendel, as well as an annotated edition of his classic paper, can be found in Alain F. Coreos and Floyd V. Monaghan, *Gregor Mendel's Experiments on Plant Hybrids. A Guided Study*, New Jersey: Rutgers University Press, New Brunswick, 1993.

ways (such as white and purple flower color); and ancestry of the plants should be precisely known by using only true-breeding lines in the experiments. (In modern jargon, these are 'boundary conditions' that must obtain in order to ascertain the patterns by which parental traits are inherited by their offspring.) Mendel's hypotheses were formulated in probabilistic terms; accordingly, he obtained large samples and subjected them to statistical analysis.

Mendel studied the transmission of seven different traits in the garden pea, *Pisum sativum*, including the color of the seed (yellow versus green), the configuration of the seed (round versus wrinkled), and the height of the plant (tall versus dwarf). The results of Mendel's experiments are too well known to need detailed presentation here, but it is worth analyzing the various stages of his methodology. His first series of experiments was with plants that differ in a single trait. The regularities observed led to certain generalizations having the form of law-like statements: only one of the two traits (the *dominant* trait) appears in the first generation progenies; after self-fertilization, three-fourths of the second-generation progenies exhibit the dominant trait, and one-fourth exhibit the other (*recessive*) trait; the second-generation plants exhibiting the recessive trait breed true in the following generations, but the plants exhibiting the dominant trait are of two kinds, one-third breed true, the other two-thirds are hybrids. Mendel tested these generalizations by repeating his experiments for each of the seven characters. These generalizations were summarized in a law, later called the Principle of Segregation: hybrid plants produce seeds that are one-half hybrid, one-fourth pure breeding for the dominant trait, and one-fourth pure breeding for the recessive trait.

Mendel tested the hypothesis of segregation by deriving and verifying additional predictions. For example, he predicted that after  $n$  generations of self-fertilization the ratio of true-breeding to hybrid plants in the progeny of a hybrid should be  $2^n - 1$  to 1. He explicitly stated that this prediction would obtain only if the following condition obtained, that all plants have 'equal average fertility ... in all generations' (which is an interesting insight on the consequences of natural selection, a notion that was elucidated by his 13 years-older contemporary, Charles Darwin).

The study of the offspring of crosses between plants differing in two traits (e.g., round and yellow seeds in one parent, wrinkled and green seeds in the other parent) led him to formulate a second law, later called the Principle of Independent Assortment: 'The principle applies that in the offspring of the hybrids in which several essentially different characters are combined, ... the relation of each pair of different characters in hybrid union is independent of the other differ-

ences in the two original parental stocks'. He corroborated this principle by examining progenies of plants differing in three and four traits. He correctly predicted and corroborated experimentally that in the progenies of plants hybrid for  $n$  characters there will be  $3^n$  different classes of plants.

The formulation and experimental testing of the two principles stated (also known as the First and Second Law of Inheritance) take up only approximately the first half of Mendel's paper. Midway through the paper Mendel advanced what he properly called a 'hypothesis' or theory to account for his previous results and for the two laws. The second half of the paper is dedicated to the derivation of predictions from the theory and to test them.

Mendel's theory of inheritance contains the following elements: (1) for each character in any plant, whether hybrid or not, there is a pair of hereditary factors ('genes'); (2) these two factors are inherited one from each parent; (3) the two factors of each pair segregate during the formation of the sex cells, so that each sex cell receives only one factor; (4) each sex cell receives one or the other factor of a pair with a probability of one-half; (5) alternative factors for different characters associate at random in the formation of the sex cells.

Mendel's well-deserved eminence as one of the great scientists of all times rests particularly on the formulation of this theory of heredity. Mendel was also quite aware of the logical status of his proposal, namely that it was a hypothesis that required experimental corroboration. Just after formulating the theory that I have summarized in the previous paragraph, Mendel wrote that 'this hypothesis would fully suffice to account for the development of the hybrids in the separate generations', i.e., the hypothesis is consistent with his previous experiments. But that was not enough, as he recognized, since the theory had been designed to fit the results. New tests would be required. He writes: 'In order to bring these assumptions to an experimental proof the following experiments were designed'. The experiments are two series of back-crosses that confirm segregation and independent assortment in the egg cells, and then in the pollen cells.

### *The Destruction of Knowledge by Ideology: Lysenko and Genetics in the Soviet Union*

An egregious example of scientific fraud is the case of Lysenkoism, which violated virtually every canon of scientific practice and that

stands as dramatic counterexample to Mendel's achievements.<sup>26</sup> In February 1935, the agronomist Trofim Denisovich Lysenko – an opportunist charlatan with pretensions of being a great revolutionary scientist – addressed the Second Soviet Congress of Collective Farms on the shameful status of Soviet agriculture. Lysenko castigated Soviet geneticists, accusing them of being enemies of the people who were destroying Soviet agriculture by relying on abstract theories imported from the capitalistic West. Stalin, presiding over the event, expressed his approval: 'Bravo, comrade Lysenko, bravo!'

Stalin's public approval consummated Lysenko's meteoric rise to fame and power. For three long decades, until the fall of Krushchev in October 1964, Lysenko and his partisans presided over Soviet agriculture, imposed their ideas on biology, and completed the elimination of Soviet genetics (and of numerous Soviet geneticists, who were sentenced to death, sent to concentration camps, or at best removed from their research and teaching positions).<sup>27</sup> The Soviet Union, a country with enormous agricultural potential, would as a consequence become, for many years extending into the present, agriculturally insufficient and backward in biology (contrary to its successes in other disciplines, like physics and mathematics).

Lysenko denounced genetics as a capitalistic science that perpetuated the notion that there are qualitative differences – claimed to be rooted in the genes – in plants, animals, or people. Such immutable differences do not exist, according to Lysenko; rather, differences between individuals are due to environmental effects and can be radically modified by exposing organisms to appropriate environmental challenges. Therefore, the production of new crops, or their adaptation to new habitats, need not be the long process of selection of suitable genotypes claimed by the capitalists, but can be simply and rapidly accomplished by exposing seeds or young plants to suitable conditions. At the height of his power, under Stalin's protecting approval, Lysenko's absurd utterances included the claim that in the appropriate environment wheat plants produce rye seeds.

Lysenko promised rapid increases in crop yields and the transformation of barren or poor lands into agricultural windfalls. He intro-

<sup>26</sup> An authoritative version of the Lysenko affair can be found in Z.A. Medvedev, *The Rise and Fall of T.D. Lysenko*, edited and preface by I.M. Lerner, New York: Columbia University Press, 1969.

<sup>27</sup> The distinguished Russian evolutionist Nikolai N. Vorontsov has written of Lysenko's impact on Soviet biology: 'I remember 1948 very well, that fall, in all universities, in all institutions three thousand biologists lost their jobs and all possibility of research – three thousand'. See 'Current State of Evolutionary Theory in the USSR', in L. Warren and H. Kopyowski (eds.), *New Perspectives in Evolution*, New York: John Wiley, 1991: 68.



duced practices such as the 'vernalization' method of seed adaptation to harsh climates or the grassland system of crop rotation, which proved to be gigantic agricultural catastrophes. He suppressed genetics research and eliminated the teaching of genetics from universities and agricultural institutes.

How could absurd claims of such enormous magnitude and economic consequence persist for decades? Social, political, and other factors came, of course, into play. The relevance to my present purposes is that Lysenko completely abrogated the traditional practices of science. He avoided properly designed tests that could falsify his theories and, instead, supported his claims with crude experiments that could be interpreted at will. Contrary evidence was denied or denounced on the grounds that nothing could possibly be right that contradicted the superior ideology of Marxism-Leninism. The large scale failure of Lysenko's agricultural practices was attributed to subversion by the farmers and enemies of the people. Any evidence, any practice, any theory was measured by its congruence with Marxist ideology; all, and only those, actions and results were acceptable that served the cause of the Soviet State.

The extent to which political considerations rather than scientific practice dominated the Lysenko affair is apparent in the stenographic record of the session of the Lenin Academy of Agricultural Sciences of the USSR (July 31 - August 7, 1948). On this occasion Lysenko routed the remnants of genetics (and the geneticists) in the Soviet Union. In the opening address, Lysenko stated:

The party, the Government and J.V. Stalin personally, have taken an unflagging interest in the further development of the Michurin teaching. There is no more honorable task for us as Soviet biologists than creatively to develop Michurin's teachings. [Ivan Vladimirovich Michurin (1855-1935) was the Russian horticulturist, whose ideas concerning the inheritance of acquired characteristics Lysenko was consecrating.]

The transcript of the concluding meeting of the Academy's session includes Lysenko's concluding remarks:

Comrades, before I pass to my concluding remarks, I consider it my duty to make the following statement. The question is asked in one of the notes handed to me, 'What is the attitude of the Central Committee of the Party to my report?' I answer: The Central Committee of the Party examined my report and approved it. (Stormy applause. Ovation. All rise.)

Long live the Michurin teaching, which shows how to transform living nature for the benefit of the Soviet people! (Applause.)

Long live the Party of Lenin and Stalin which discovered Michurin for the world (Applause) and created all the conditions for the progress of advanced materialist biology in our country. (Applause.)

Glory to the great friend of science, our leader and teacher, Comrade Stalin! (All rise, prolonged applause.)<sup>28</sup>

*The Curious Case of Darwin, or the Discrepancy Between What Scientists Say and What They Do*

Few scientists in the nineteenth century or at any earlier time equal Mendel's clear delineation of the scientific method he was pursuing. In the English-speaking countries, scientists advanced hypotheses and tested them in their work, but often claimed in their writings to be following the orthodoxy of inductionism, proclaimed by philosophers as the method of good science. Charles Robert Darwin (1809-1882) is a remarkable example of this discrepancy.

In his *Autobiography* Darwin says that he proceeded 'on true Baconian principles and without any theory collected facts on a wholesale scale'.<sup>29</sup> The opening paragraph of *Origin of Species* conveys the same impression:

When on board H.M.S. *Beagle*, as naturalist, I was much struck with certain facts in the distribution of the inhabitants of South America, and in the geological relations of the present to the past inhabitants of that continent. These facts seemed to me to throw some light on the origin of species – that mystery of mysteries, as it has been called by one of our greatest philosophers. On my return home, it occurred to me, in 1837, that something might perhaps be made out on this question by *patiently accumulating and reflecting on all sorts of facts which could possibly have any bearing on it*. After five years' work I allowed myself to speculate on the subject, and drew up some short notes; these I enlarged in 1844 into a sketch of the conclusions, which then seemed to me probable: from that period to the present day I have steadily pursued the same object. [Emphasis added]

Darwin claims also in other writings to have followed the inductivist canons. The facts are very different from these claims, however. Darwin's notebooks and private correspondence show that he entertained the hypothesis of the evolutionary transmutation of species

<sup>28</sup> The complete transcript of the session has been published in translation in *The Situation in Biological Sciences* (International Publishers, New York, 1949). The excerpts can be found on pp. 49, 605, and 617 and are cited in L. Warren and H. Koprowski (eds.), *New Perspectives in Evolution*, (footnote 27), p. 76.

<sup>29</sup> C. Darwin, *Autobiography*, (footnote 9), p. 119.

shortly after returning from the voyage of the *Beagle*, and that the hypothesis of natural selection occurred to him in 1838 – several years before he claims to have allowed himself for the first time ‘to speculate on the subject’. Between the return of the *Beagle* on October 2, 1836, and publication of *Origin of Species* in 1859 (and, indeed, until the end of his life), Darwin relentlessly pursued empirical evidence to corroborate the evolutionary origin of organisms, and to test his theory of natural selection.

Why this disparity between what Darwin was doing and what he claimed? There are at least two reasons. First, in the temper of the times, ‘hypothesis’ was a term often reserved for metaphysical speculations without empirical substance. This is the reason why Newton, the greatest ever theorist among scientists, also claimed, *hypotheses non fingo* (‘I fabricate no hypotheses’). Darwin expressed distaste and even contempt for empirically untestable hypotheses. He wrote of Herbert Spencer: ‘His deductive manner of treating any subject is wholly opposed to my frame of mind. His conclusions never convince me ... His fundamental generalizations (which have been compared in importance by some persons with Newton’s Laws!) which I dare say may be very valuable under a philosophical point of view, are of such a nature that they do not seem to me to be of any strictly scientific use. They partake more of the nature of definitions than of laws of nature. They do not aid me in predicting what will happen in any particular case’.<sup>30</sup>

There is another reason, a tactical one, why Darwin claimed to proceed according to inductive canons: he did not want to be accused of subjective bias in the evaluation of empirical evidence. Darwin’s true colors are shown in a letter to a young scientist written in 1963: ‘I would suggest to you the advantage, at present, of being very sparing in introducing theory in your papers (I formerly erred much in Geology in that way); *let theory guide your observations*, but till your reputation is well established, be sparing of publishing theory. It makes persons doubt your observations’.<sup>31</sup> Nowadays also scientists, young or not, often report their work so as to make their hypotheses appear as afterthoughts, conclusions derived from the evidence at hand, rather than as preconceptions tested by empirical observations.

Darwin rejected the inductivist claim that observations should not be guided by hypotheses. The statement quoted earlier, ‘A man might

<sup>30</sup> C. Darwin, *Ibid.*, p. 109.

<sup>31</sup> F. Darwin (footnote 5), volume 2, p. 323 (emphasis added). See David Hull, *Darwin and His Critics*, Cambridge, Massachusetts: Harvard University Press, 1973.

as well go into a gravel-pit and count the pebbles and describe the colours', is followed by this telling remark: 'How odd it is that any-one should not see that all observation must be for or against some view if it is to be of any service!'<sup>32</sup> He acknowledged the heuristic role of hypotheses, which guide empirical research by telling us what is worth observing, what evidence to seek. He confesses: 'I cannot avoid forming one [hypothesis] on every subject'.<sup>33</sup>

Darwin was an excellent practitioner of the hypothetico-deductive method of science, as modern students of Darwin have abundantly shown.<sup>34</sup> Darwin advanced hypotheses in multiple fields, including geology, plant morphology and physiology, psychology, and evolution, and subjected his hypotheses to empirical test. 'The line of argument often pursued throughout my theory is to establish a point as a probability by induction and to apply it as a hypothesis to other parts and see whether it will solve them'.<sup>35</sup> Popper<sup>36</sup> has not only made clear that falsifiability is the criterion of demarcation of the empirical sciences from other forms of knowledge, but also that falsification of seemingly true hypotheses contributes to the advance of science. Darwin recognized the same: 'False facts are highly injurious to the progress of science, for they often endure long; but false views, if supported by some evidence, do little harm, for every one takes a salutary pleasure in proving their falseness; and when this is done, one path towards error is closed and the road to truth is often as the same time opened'.<sup>37</sup>

Some philosophers of science have claimed that evolutionary biology is a historical science that does not need to satisfy the requirements of the hypothetico-deductive method. The evolution of organisms, it is argued, is a historical process that depends on unique and unpredictable events, and thus is not subject to the formulation of testable hypotheses and theories. Such claims emanate from a monumental misunderstanding. There are two kinds of questions in the study of biological evolution.<sup>38</sup> One concerns history: the study of phylogeny,

<sup>32</sup> F. Darwin (footnote 5), volume 1, p. 195.

<sup>33</sup> C. Darwin, *Autobiography* (footnote 9), p. 141.

<sup>34</sup> See, e.g., Gavin De Beer, *Charles Darwin, A Scientific Biography*, New York: Doubleday, Garden City, 1964; M.T. Ghiselin, *The Triumph of the Darwinian Method*, Berkeley: University of California Press, 1969; D. Hull, *Darwin and His Critics* (footnote 31); E. Mayr, 'Introduction' In: Charles Darwin, *On The Origin of Species*, Cambridge, Massachusetts: Harvard University Press, 1964.

<sup>35</sup> Charles Darwin, 'Darwin's Notebooks on Transportation of Species', G. De Beer (ed.), *Bulletin British Museum* 2 (1960), 23-200.

<sup>36</sup> K.R. Popper (footnote 2).

<sup>37</sup> Charles Darwin, *The Descent of Man and Selection in Relation to Sex*, London: Murray, 1871; 2nd edition, 1889.

<sup>38</sup> See Theodosius Dobzhansky, *Genetics and the Origin of Species*, 3rd edition, New York: Columbia University Press, 1951: 11-12.

the unravelling and description of the actual course of evolution on earth that has led to the present state of the biological world. The scientific disciplines contributing to the study of phylogeny include systematics, paleontology, biogeography, comparative anatomy, comparative embryology, and comparative molecular biology. The second kind of question concerns the elucidation of the mechanisms or processes that bring about evolutionary change. These questions deal with causal, rather than historical, relationships. Population genetics, population ecology, paleobiology, and many other branches of biology are the relevant disciplines.

There can be little doubt that the causal study of evolution proceeds by the formulation and empirical testing of hypotheses, according to the same hypothetico-deductive methodology characteristic of the physicochemical sciences and other empirical disciplines concerned with causal processes. But even the study of evolutionary history is based on the formulation of empirically testable hypotheses. Consider a simple example. For many years specialists proposed that the evolutionary lineage leading to man separated from the lineage leading to the great apes (chimpanzee, gorilla, orangutan) before the lineages of the great apes separated from each other. Some recent authors have suggested instead that man, chimpanzees, and gorillas are more closely related to each other than the chimpanzee and the gorilla are to the orangutan and other Asian apes. A wealth of empirical predictions can be derived logically from these competing hypotheses. One prediction concerns the degree of similarity between enzymes and other proteins. It is known that the rate of amino acid substitutions is approximately constant when averaged over many proteins and long periods of time. If the older hypothesis is correct, the average amount of protein differentiation should be greater between man and the African apes than among these and orangutans. On the other hand, if the newer hypothesis is correct, man, gorilla and chimpanzee should have greater protein similarity than any of the three has with orangutans. These alternative predictions provide a critical empirical test of the hypotheses. The available data favor the second hypothesis. Man, chimpanzee, and gorilla appear to be phylogenetically more closely related to each other than any one of them is related to orangutans.

Certain biological disciplines relevant to the study of evolution are largely descriptive and classificatory. Description and classification are necessary activities in all branches of science, but play a greater role in certain biological disciplines, such as systematics and biogeography, than in other disciplines, such as population

genetics. Nevertheless, even systematics and biogeography use the hypothetico-deductive method and formulate empirically testable hypotheses.

*Theory Replacement: Phlogiston and Lavoisier; Newtonian Mechanics and Einstein*

Science is progressive. Theories that are accepted at one time may later be rejected because they are found to be wrong. More often, however, particularly in well-developed scientific disciplines, a theory that accounts at a time for much that is known, is only rejected when it becomes replaced by a different theory that accounts for the same phenomena, as well as others that the former theory left unexplained. Two examples illustrate these two situations: the phlogiston theory that was replaced by Lavoisier's discovery of oxygen and his theory of combustion, and Newton's theory of motion, which was replaced by the theory of relativity.

Johann Becher in 1669 proposed that matter consisted of three kinds of earth: the vitrifiable, the mercurial, and the combustible. A substance such as wood consisted of combustible earth plus ashes. When the wood was burned, combustible earth was liberated. The hypothesized combustible earth was named 'phlogiston' half a century later by Georg Stahl, who claimed that the corrosion of metals was also a form of combustion, and that phlogiston was lost in the process. The phlogiston theory was accepted by Joseph Priestly and other eminent eighteenth-century scientists.

The phlogiston theory was demolished by Antoine Lavoisier (1743-1794) in a series of experiments published in 1787.<sup>39</sup> This publication was followed in 1789 by his *Traité élémentaire de chimie*, which may very well be considered the treatise that initiates modern chemistry. Lavoisier rejected the phlogiston theory on the grounds that it led to erroneous predictions. He first noticed that the ash of wood (or other burned organic substances) weighed less than these substances did before burning, whereas sulfur and phosphorus weighed more, although phlogiston had been liberated in both cases according to the theory. Lavoisier then tested the phlogiston theory by systematically weighing all substances involved in the combustion or calcination of a

<sup>39</sup> See Henry Guerlac, 'Lavoisier, Antoine Laurent', *Dictionary of Scientific Biography*, volume VIII, New York: Scribner's Sons, 1973: 66-99.

variety of organic substances as well as metals. These experiments manifested the presence of two substances in air, one (which he named oxygen) was absorbed by burning, the other was the 'nonvital' air (nitrogen) that remained behind. He proposed that combustion was not the result of liberation of the hypothetical phlogiston, but the combination of the burning substance with oxygen. He tested this theory with carefully designed experiments in which all substances involved were weighed before and after burning; but also by extending the theory to other processes involving oxidation, such as rusting, and to a variety of natural phenomena. Thus, Lavoisier explained water as the product of the combination of oxygen and hydrogen. He applied this methodology, of testing theories by predicting events and precisely measuring their outcome, to the resolution of numerous matters of public interest. In a well known instance, he collaborated with Benjamin Franklin in debunking Franz Anton Mesmer's claim that he was able to cure by means of 'animal magnetism'.

The phlogiston story illustrates an important dimension of the scientific process: the reluctance of scientists to reject an accepted theory until another becomes formulated that accounts for the phenomena explained by the pre-existing theory. Joseph Priestley and other contemporary scientists had continued for a time to accept the phlogiston theory even in the face of falsifying experiments. The phlogiston theory became generally rejected only towards the end of the eighteenth century, after Lavoisier had developed and corroborated his own theory of combustion.

Scientific advance occurs not only, as with phlogiston, through the replacement of an erroneous theory by a correct one, but also by the replacement of a largely correct theory by a more precise or more inclusive theory. The examples are numerous. One famous instance is the replacement of Newtonian mechanics by Einstein's theory of relativity. As it is often the case in the progress of scientific knowledge, the predictions made by the earlier theory are largely correct, which is why the theory, in this case Newtonian mechanics, had passed numerous tests and become generally accepted. But the newer scientific theory is able to account for phenomena left unexplained by the previous theory. In some instances, this happens because the new theory has much greater generality and is able to account for phenomena explained by different theories or even different disciplines. One example is statistical mechanics, which became able to account for many conclusions of thermodynamics once it was discovered that the temperature of a gas reflects the mean kinetic energy of its molecules.

In the case of Einstein vis-à-vis Newton, it is particularly interesting that fundamental assumptions of the Newtonian theory, e.g., that mass is constant and that space and time are absolute realities, are rejected by the theory of relativity. Yet, with respect to bodies with intermediate mass and intermediate velocities (that is, the bodies and motions encountered in the course of ordinary experience), Newton's and Einstein's theories make practically identical predictions.<sup>40</sup>

Isaac Newton (1642-1727) is one of the greatest scientists of all times. He formulated the laws of motion and the law of gravity, developed a theory of light, invented the calculus, and much more. Newton's myriad discoveries include solutions of the so-called 'two-body problem', i.e., the shape and size of the planetary orbits; the mass of the moon (one-eightieth that of the earth), calculated from the heights of the tides; the tilt of the earth's axis (23.5 degrees) that accounts for the seasons; the size of the earth's bulge at the equator; and he showed that the periods of the orbits of the planets should be proportional to the square of their distance from the sun, rather than to three halves as predicted by Descartes' theory.

Albert Einstein (1879-1955) is another scientific giant who, like Newton, made discoveries of monumental importance. In 1905, he formulated the special theory of relativity, which sets that the mass of a body is not constant, as assumed by Newton's theory and common sense experience, but increases with the speed of the body and becomes nearly infinite as a body approaches the speed of light. (The equation is  $m = m_0 / \sqrt{1 - v^2 / c^2}$ , where  $m_0$  is the mass at rest,  $v$  is the speed of the body, and  $c$  is the speed of light in vacuum.) Einstein's general theory of relativity (1916) sets that mass is not constant, but rather can be converted into energy, as famously expressed by the equation  $E = mc^2$ ; that, also contrary to common sense experience and the Newtonian theory, space and time are not absolute; that the same two events may be simultaneous for one observer, but not so for a different observer; that the speed of light is the maximum possible velocity in the universe; that the rate of a moving clock decreases as its velocity increases (and thus that if a space traveller were to leave his twin brother on Earth while he travelled at great speed for a year, he would discover upon return that he was younger than his twin brother); and so on. The special theory of relativity is now well confirmed and the general theory has been shown consistent with some critical experiments designed to test it. Concerning phenomena of

<sup>40</sup> J. Schwartz (footnote 12).



ordinary experience, the results predicted by relativity and Newtonian mechanics are virtually identical, although the two theories greatly deviate in their predictions of phenomena occurring at velocities near the speed of light.

As in the case of relativity theory, scientific knowledge often advances by the substitution and supplementation of one theory by another more complete, more precise, or more inclusive. Thus, the modern theory of genetics, for example, has identified conditions that are exceptions to Mendel's second law; has defined the chemical composition of genes; has subsumed much that had been earlier formulated by the cell theory; and has integrated Darwin's theory of natural selection in the subdiscipline known as population genetics.

### *Hurried Science: Robert Koch's Failed Tuberculosis Vaccine*

Robert Koch (1843-1910) by his middle thirties was already considered a distinguished scientist. While a practicing physician and working in a modest laboratory that he had built in his own home (in the small northern German town of Wollstein), he developed methods to culture and photograph bacteria. These methods led to the discovery of the life cycle of anthrax (which made it possible to explain the recurrence of the disease in long unused pastures). He later acquired a scientific post in Berlin, where he started investigating tuberculosis, the major cause of mortality among young adults in nineteenth century Europe. On March 24, 1882, Koch announced that he had discovered the cause of tuberculosis, the tubercle bacillus, a discovery that brought him further fame and later the Nobel Prize.<sup>41</sup>

Koch isolated and cultured the tubercle bacillus and set to find a cure for tuberculosis. He would soon announce that he had discovered a substance that could protect against tuberculosis and even cure it. This announcement was received as a bombshell by the medical world. English journals like *The Lancet* and *The British Medical Journal* published complete translations of the article, and the *Review of Reviews* dedicated nearly a complete issue to the subject. Arthur Conan Doyle, who was still practicing medicine although already a well known writer, arrived in Berlin shortly after the announcement, and would soon publish an article on Koch and his discovery.

<sup>41</sup> A brief, readable, and well documented biography is C.E. Dolman, 'Koch, Heinrich Hermann Robert', *Dictionary of Scientific Biography*, volume VII, New York: Scribner's Sons, 1973: 420-435.

Two matters were troublesome with Koch's announcement. One was that he refused at first to reveal the nature of the curative substance, although he did so a year later under the pressure of public criticism. The second matter may have been related to the first: Koch's experimental testing of the vaccine was virtually lacking. It was eventually to prove ineffective as either prevention or cure. Koch had anticipated, on the basis of limited evidence, that an injection of dead bacilli to a person who would later be infected with living ones, would result in a local reaction that might protect the person. In any case the local reaction would serve for diagnostic purposes. Perhaps because of his early successes (which included the discovery of the agent of cholera and its mode of transmission), Koch had become persuaded that his hypothesis for diagnosis and cure would prove to be correct. Thus, he proceeded to announce it as a curative method without adequate testing. *The British Medical Journal*, which had earlier celebrated the original announcement, published a devastating article condemning Koch for having attempted to keep secret the composition of the substance and for having recommended it as a remedy without adequate testing.

*The Explanatory Context of Discovery, or Why Empirical Testing is Not Enough: Avery's DNA and Wegener's Continental Drift*

Empirical testing may be necessary, but is not sufficient for the scientific community to accept a new hypothesis. A hypothesis that has withstood even the most severe attempts to falsify it will not be accepted unless it has explanatory value; i.e., unless it can be understood within the contemporary scientific context and unless it makes the problem at hand intelligible. There are several notable cases of scientific discoveries that were not accepted at the time because they were 'premature', they were not contextually intelligible. Mendel's discovery of the laws of inheritance may fit this situation. Two more recent examples are the discovery by Oswald Avery (1877-1955) and his colleagues that DNA is the hereditary substance (rather than protein as was generally believed at the time); and (2) the theory of continental drift proposed by Alfred Wegener (1880-1930).

Avery was a distinguished scientist at a leading research institution, the Rockefeller Institute for Medical Research in New York. In 1944 he published a paper, with his colleagues C.M. MacLeod and M. McCarthy, showing that the 'transforming factor' responsible for the hereditary specificity of *Pneumococcus* bacteria (agents of severe

pneumonia) was deoxyribonucleic acid (DNA) and that protein was not involved at all.<sup>42</sup> Avery had performed a careful series of diverse and very specific tests that definitely identified DNA as the transforming factor and excluded other molecular species. There was no challenge to the experimental results, but the scientific community refused for several years to accept that DNA is the substance of heredity. This reluctance derived precisely from what was known about DNA, which 'knowledge' made it impossible for DNA to encode hereditary information. It turned out eventually that 'what was known' about DNA was wrong; at least one seemingly inconsequential fact was. DNA became accepted as the hereditary substance only after the erroneous 'detail' was corrected.

Nucleic acid was discovered in 1869 by Johann Friedrich Miescher, a 25 year-old Swiss. By the 1920s two kinds of nucleic acid (RNA, i.e., ribonucleic acid, and DNA) had become known, and their composition was soon thereafter elucidated. DNA was shown to be made up of four relatively simple components (nucleotides) similar to each other in all respects except their nitrogen base, which could be one of four: adenine, guanine, cytosine and thymine (usually represented by A, G, C, and T). Much of the relevant knowledge came from Phoebus Aaron Levene, an organic chemist of towering reputation also working at the Rockefeller Institute. Levene had proposed that DNA was made up of long repetitions of the four nucleotides following one another in an invariant fashion. This was called 'the tetranucleotide hypothesis', which was accepted without challenge – largely because accurate measurement of the proportions of the four nucleotides was not possible with the analytical methods of chemistry available at the time, but also because it was incorporated in the model for the composition of DNA elucidated by the highly reputed Levene.

The tetranucleotide hypothesis entailed that DNA could not be the carrier of hereditary information. An endless repetition of the same four components in the same order could not encode information of any kind, for the same reason that a repetition of the same four letters of the English alphabet cannot convey semantic information, no matter how long the sequence. Protein, on the contrary, was known to be made up of some twenty different amino acids, which varied in proportion from one to another protein. Protein, therefore, could be an informative molecule, whereas DNA was a 'stupid' molecule. Since

<sup>42</sup> O.T. Avery, C.M. MacLeod, and M. McCarthy, 'Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types. Induction of Transformation by a Deoxyribonucleic Fraction Isolated From *Pneumococcus* Type III', *J. Experimental Medicine* 79 (1944), 137-158.

both protein and DNA were present in the nucleus of the cell, it was generally assumed that protein would prove to be the carrier of hereditary information. In any case, DNA could not be, the experiments of Avery notwithstanding, because it could not convey information. Later on, after the chemist Erwin Chargaff at Columbia University showed that the proportions of the four bases, A, T, C, and G, vary from one to another organism and the tetranucleotide hypothesis was rejected, DNA became promptly accepted as the hereditary chemical. The race to determine its structure was on, a feat that was accomplished in 1953 by James Watson and Francis Crick.<sup>43</sup>

A somewhat different state of affairs, but grounded on the same need for explanatory value, is the case of Alfred Wegener, a respected meteorologist and geologist, who first proposed in 1912 and developed in 1915, the hypothesis of continental drift.<sup>44</sup> He noted the complementary shape of the coastlines on both sides of the Atlantic and reviewed geological and paleontological evidence scattered in the literature that led him to conclude that during the late Paleozoic (225 to 350 million years ago) all continents were assembled into a single supercontinent, which he named 'Pangea'.

Wegener tested his hypothesis that the continents had drifted by searching the literature for relevant geological, biogeographical, and paleoclimatological evidence. The evidence was striking, showing for example that strata and folds on opposite sides of the Atlantic fitted precisely with each other, and extended beyond the coastlines in complementary patterns. Wegener, however, was unable to produce a convincing explanation of how the continents could move. His hypothesis was rejected with disbelief and the evidence relegated to a curiosity. It was only three decades later that continental drift would become accepted, after the theory of plate tectonics provided a plausible mechanism for continental displacement.

### *Social Mechanisms: Peer Review and Publication*

The process of testing a scientific hypothesis may corroborate it or falsify it. Corroboration may later be overturned. Falsification is a logically conclusive method: if a necessary consequence of a premise is false, then the premise must also be false. But both falsification and

<sup>43</sup> Footnote 20.

<sup>44</sup> See K.E. Bullen, 'Wegener, Alfred Lothar', in *Dictionary of Scientific Biography*, volume XIV, New York: Scribner's Sons, 1976: 214-217.

corroboration are subject to human error. For example, the *modus tollens* may lead to erroneous conclusions if the prediction tested does not in fact logically follow from the hypothesis. Moreover, an observation or experiment contradicting a hypothesis may have been performed or interpreted erroneously. Thus, scientists require that their experiments be made public with sufficient detail so that they can be repeated.

The actual replication of experiments is, nevertheless, selective. It is usually reserved for experiments of unusual significance or for those that conflict with established knowledge. Confronted with a new result that impacts their own work, scientists usually do not proceed to check it by repeating it, but rather will build upon the result and modify their own hypotheses and design their own experiments accordingly. If something goes wrong with their work, they may turn back upon the original results and check them by repeating them. But time, resources, and prestige will have been lost along the way.<sup>45</sup>

To minimize such problems, review mechanisms have become an integral part of science. The scientific community simultaneously seeks to encourage innovative thinking and assure that new ideas are subjected to rigorous review. On the one hand, science is a creative process, in which advances occur only if researchers are encouraged to develop and test innovative ideas. On the other hand, because science is a cumulative subject in which each scientist must build on the work of others, the scientific community has great stakes in weeding out false ideas. Accordingly, creativity is tempered by the need for rigorous review of new results.

Peer review represents both an effort to police scientific claims and to assure their widest possible dissemination. The pressure on scientists to publish derives not only from narrow concerns for recognition and career advancement, but also from the desire of all scientists to learn of new developments that may guide their own work. Because submitting a paper for peer review is the best way to disseminate and to establish priority for a new discovery or idea, the process serves to get new information out fast as well as to control its quality. The comments of peer reviewers contribute to the advancement of science by helping proponents of new hypotheses to improve their research and interpretations.

Peer-review scrutiny of science takes place in a variety of contexts. Informal review can occur when scientists discuss their work with one

<sup>45</sup> See *On Being a Scientist* (footnote 23).

another at the laboratory bench, during conversations and seminars, and at scientific meetings. Formal peer review is generally an integral part of the scientific publication process and of the process by which funds are granted for the conduct of research. Any claim that would significantly add to, or change the body of scientific knowledge must be regarded skeptically if it has not been subject to some form of peer scrutiny, preferably by submission to a reputable journal. Publication in a peer-reviewed journal does not by itself guarantee the validity of the published results; nor is there reason for outright rejection of every work that has not been published in a reputable journal. But one should treat with utmost suspicion a proposition that has not been subjected to peer review.

Peer review delays somewhat the publication of results, but the delay and the large investment of time that reviewers and journal editors dedicate to the process are justified by the need to weed out erroneous results. The process of peer review is subject to human error and prejudices, but it is the most accessible and often most dependable element of the process of invention, validation, and refinement by which scientific knowledge advances.

Peer review does not thwart new ideas. Journal editors and the 'scientific establishment' are not hostile to new discoveries. Science thrives on discovery, and scientific journals compete to publish new breakthroughs. Indeed, the most prestigious prizes are awarded to those scientists who make the most daring and dramatic discoveries, even when these contradict revered theories. We referred earlier to the revolutionary character of Einstein's relativity theory and to the explosive advances of molecular biology, some of which were triggered by a sequence of unanticipated discoveries, many rewarded by the Nobel Prize and other awards, that contradicted previous assumptions.

Mistakes, errors, failures, and prejudices infect science as well as other human activities. But the large and ever expanding body of scientific knowledge and its useful applications attest to the success of the scientific enterprise. The distinctive methodology of science accounts for much of this success, but the institutional mechanisms that have been developed contribute to that success.