A Brief History of the Hypothesis

David J. Glass[1],* and Ned Hall[2],*
[1]Novartis Institutes for Biomedical Research, Cambridge, MA 02139, USA
[2]Department of Philosophy, Harvard University, Cambridge, MA 02138, USA
*Correspondence: david.glass@novartis.com (D.J.G.), ehall@fas.harvard.edu (N.H.)

Two terms used as frameworks for scientific experimentation—the "hypothesis" and the "model"—carry distinct philosophical assumptions, with important consequences for the practicing scientist.

Scientists are commonly taught to frame their experiments with a "hypothesis"—an idea or postulate that must be phrased as a statement of fact, so that it can be subjected to falsification. The hypothesis is constructed in advance of the experiment; it is therefore unproven in its original form. The very idea of "proof" of a hypothesis is problematic on philosophical grounds because the hypothesis is established to be *falsified*, not verified. The second framework for experimental design involves building a model as an explanation for a data set. A model is distinct from a hypothesis in that it is constructed after data are derived. In contrast to the hypothesis, the model must be held up for verification—its success is determined by its ability to predict a particular outcome. Furthermore, an unsuccessful, or not fully successful, model need not be scrapped in the way that the alternative framework urges the rejection of a falsified hypothesis, but it may instead serve as the starting point for a suitably refined successor. The concept of a model's "verification" requires an acceptance of "inductive reasoning"—a form of logic that allows the scientist to both generalize a particular result and say that the same result will occur in the future—which itself has been criticized. Although many scientists use the term "hypothesis" when they mean "model," we will maintain the distinction that the hypothesis is an unproven premise whereas the model is data derived, to discriminate between "top-down premise/deduction" and "bottom-up data/induction." A recent article proposes that the availability of large amounts of scientific data renders the need for a pre-existing hypothesis obsolete (Anderson, 2008). Given the ability of scientists to gather vast amounts of scientific data—by sequencing genomes, surveying changes in the expression of every gene, or analyzing proteomic changes in response to a stimulus—is a hypothesis the most appropriate way to frame such experiments (Glass, 2006)? Here, we discuss the philosophical reasoning that motivated original and current notions of the hypothesis and its implications for scientific experimental design.

The *Novum Organum*

Galileo helped to initiate the renaissance in science that occurred in Europe in the 16th and 17th centuries, by upending religious dogma as to the centrality of Earth's position in the universe, through mathematical reasoning and observation. He was described by a Paduan contemporary as "the father of experiments and all their exactness." Galileo's approach was in line with Greek antecedents, especially Aristotle and Archimedes, in his reliance on "deductive" reasoning stemming from hypotheses. The hypothesis as it was used in the 1500s was a premise—a starting point based on unproven assumptions. From the initial premise, deductions would be made, and their success or failure was determined by subjective assessments as to whether they were satisfactory in their explanations of the premise. Although this method resulted in "satisfactory" conclusions as applied by Galileo, the same cannot be said of some other 16th century physicists, who applied fictitious and nonrealistic theories to physics as well as to astronomy (Blake et al., 1966). The lack of a foundation for the hypothesis characterized the 1500s as a "century of confusion" (Hall, 1962), in which there was clear excitement over new developments but no programmatic march forward due to a lack of an accepted method to distinguish between various claims of discovery. Galileo represented a move to a hypothesis more grounded in realism, and an increased emphasis on the experiment as the basis for conclusions, but he did not systematize his approach into a methodology that could be clearly followed by others.

It was Francis Bacon who in 1620 wrote an approach to scientific methodology—his *Novum Organum* or "new instrument"—new because Bacon took issue with Aristotle's *Organon* (the term given by Aristotle's followers to his system of logic) in several respects (Bacon, 1620). Bacon noted that deductive reasoning by itself is not sufficient because if the premise is set in advance of the experiment, for example by a hypothesis, the reasoning would be twisted to meet that premise. Bacon therefore argued that a purely experimentation-based methodology was necessary, and that to solve problems with pre-existing bias, "the only hope is true *induction*."

With "inductive reasoning" a data set is taken and used to infer that under similar circumstances the result will be repeated, and that the finding as applied to a specific case may be generalized to other cases of like kind. In the case of gravity, one observes first that a mass falls toward earth at a particular rate, and later that it does so in a predictable and "verifiable" manner, meaning that after having established the rule for how quickly objects descend, one can predict that descending bodies will continue to follow

the rule. The prediction that a falling apple will behave in the future the same way it behaved in the past is inductive reasoning, as is extending the findings to other objects, such as an orange or a meteor. Bacon was careful to distinguish the type of induction that is adequate, and the process that he described sounds like scientific method as it is currently practiced, that is, a series of experiments that allow the scientist to make claims as to how things work, based on the process of refining a model by the gathering of "negatives" and "affirmatives." This "bottom-up" approach was required to escape preconceived notions, including dogma. The difference between prior method and the methodology that launched a revolution in European science was a strict process of questioning in which experimental data, not pre-existing ideologies, were the basis for knowledge.

Newton's Rejection of the Hypothesis

Starting with the second edition of his seminal work, the *Principia*, Isaac Newton included a now-famous philosophical section, in which appeared the phrase "*Hypotheses non fingo*." Andrew Motte, who did the first English translation in 1729, rendered the term as, "I frame no hypotheses." Newton stated his views on hypotheses even more explicitly in his *Opticks*, noting that "Hypotheses are not to be regarded in experimental Philosophy" (Newton, 1721).

Newton was consistent with Bacon in the primacy of the experiment or proof to construct a rule as to how reality operated and in his willingness to use the rule inductively. He also came to reject the hypothesis as being inconsistent with this bottom-up approach as it would frame the project with an unproven premise—in line with the criticism made earlier by Bacon. As for how his data-based rules were to be used, Newton wrote that inductions "should be considered either exactly or very nearly true until new phenomena may make them either more exact or liable to exceptions" (Newton, 1729); his laws were accepted because they succeeded in describing how the physical world worked. Newton thus saw a distinction between some claims about the world, hypotheses, that, due to a lack of

experimental proof, should be avoided and others, inductions, that, thanks to their grounding in experiment, deserved to be supported. Newton was similar to Bacon in his willingness to amend models based on data. Therefore, the way in which Newton established inductions by empirical evidence was not the way that a mathematical theorem could be established—by deduction from some axioms.

Hume's Rejection of Inductive Reasoning

The idea that past experience can be used as "proof" of future outcomes was rejected by the 18th century Scottish philosopher David Hume. Hume introduced a "radical skepticism"—the idea that one could not use past experience to predict the future. He applied the rejection of experience even to gravity, noting that the prior experience that an object may be wed to certain qualities is no guarantee that this will be the case in the future. Even the idea of identity—for example, whether one could say that a chair left unobserved in a room upon exiting continues to exist in the absence of verification, or whether it would be the same chair upon re-entering the room—is not verifiable (Hume, 1749).

Hume's most notable contribution to the philosophy of science was this "problem of induction"… that one cannot claim that a past result predicts the future because such a claim is based on the unprovable premise that a thing and its attributes will remain bound to each other, that nature's laws are stable. Hume's clever objection came by establishing first that all inductive reasoning is based on the assumption that nature is uniform (throughout space and time), second by inquiring how we justify this assumption, and then by pointing out that it looks circular to say "We are justified in believing that nature is uniform because it has always been uniform in the past."

Critical Rationalism

In the 20th century, the Austrian philosopher Karl Popper sought to produce a philosophy of science that was consistent with Hume's critique of induction (Popper, 1959). Popper's work is the basis for much of the framing terminol-

ogy used by practicing scientists today, particularly his use of the word "hypothesis," which was distinct from the way the term was used by earlier philosophers (Popper, 1959). Popper's solution to the "problem" of inductive reasoning was to suggest a methodology where concepts are subjected to falsification, as opposed to verification. In this way, one could avoid the circumstance where an idea is stated to be "true," implying that it will hold to be accurate in the future, and rather focus on whether the idea could be proven false. Such an approach seems attractive because it establishes a framework where a single piece of contrary evidence would be deemed sufficient to claim that a hypothesis had been proven wrong. However, a corollary to the framework is that no amount of supporting, or nonfalsifying, evidence would be sufficient for verification, where the term "verification" is used to claim that a rule could be said to be predictive or to make use of induction. Popper's philosophical approach was termed "Critical Rationalism," continuing Hume's theme that inductive reasoning was not rational.

In Defense of Induction

Critical Rationalism has not been without detractors, and inductive reasoning does not lack for defenders. However, the most common argument leveled against Popper's framework is that Critical Rationalism fails to avoid inductive reasoning (Kuhn, 1977). For example, one might ask first what motivated the construction of a hypothesis; the motivation is demonstrably the scientist's inference based on past experience. Popper responded that anything could be used to motivate a hypothesis, and that the type of motivation was not relevant (Popper, 1959). This answer may sound disingenuous and inconsistent with how science actually operates. As the mathematician Henri Poincaré noted, "It is often said that experiments should be made without preconceived ideas. That is impossible. Not only would it make every experiment fruitless, but even if we wished to do so, it could not be done" (Poincaré, 1952).

As to a defense of induction, one might ask how it is possible to distinguish between a scientific fact and

science fiction if one is not allowed to say that a particular model predicts the future better than an alternative. If one cannot say that the experience that an apple falls toward Earth predicts that the apple will fall toward Earth, then science has no greater claim than religion or fantasy as to the fate of the apple when it is released tomorrow. Poincaré argued that if a thing and its aspect are separable, this should be seen with a sufficient number of tests, and thus one will eventually succeed in either falsifying an incorrect hypothesis or in using the lack of falsifying data as proof of verification. Poincaré's reasoning is both inductive and probabilistic, as it is based on doing a sufficient number of experiments to obtain data that can be used to predict future outcomes.

Even Hume invoked probability to reject claims of miracles—he wrote that nothing is credible which is contradictory to experience, or at variance with the laws of nature (Hume, 1749)—a position that seems to be a defense of induction, or pretty close to it. Thus, despite Popper's blanket rejection of probability as justifying induction, it is probability (which necessarily implies a data set based on experimentation) that may answer the "problem of induction" (Carnap, 1980; Hall and Hájek, 2001). If the reader finds the reasoning to be circular, the response is that probability shifts the burden to the critic. If one has shown in a sufficient number of instances that result A is achieved, then at some point it becomes the critic's responsibility to show that B, or at least "not A," may be achieved. If there is no evidence for "B" (or if the result "B" is improbable) then one may say that predicting result A is rational. This is tricky, though, because of the Humean claim that prior probabilities are irrelevant to future instances. Thus, for a system based on probabilities to get a proper footing, one must recognize that there are built-in preferences for particular types of models and be satisfied with a system that is shown to be "workable," if not ultimately "provable" (Russell, 1912). The experience that the law of gravity holds "true" in thousands of attempts to show it not to be true allows the scientist to ask the skeptic to demonstrate how it could be

untrue in the future. As the British philosopher Bertrand Russell explained, the relation between a thing and the rule that controls that thing may be shown to be nonseparable by experience, thus establishing the force of inductive reasoning (Russell, 1912).

As for claims that falsification is more "scientific" than verification, there are problems with falsification as well; for example, it might be impossible to conclusively "disprove" a hypothesis because despite the claim about a single counterexample being sufficient, there remain semantic maneuvers that might be used to reinstate a hypothesis, by ignoring or defining away negative data. Although Thomas Kuhn was not an inductivist, he quoted Popper in this way: "In point of fact, no conclusive disproof of a theory can ever be produced; for it is always possible to say that the experimental results are not reliable or that the discrepancies which are asserted to exist between the experimental results and the theory are only apparent and that they will disappear with the advance of our understanding." Kuhn continues, "For Sir Karl [Popper], they are an essential qualification which threatens the integrity of his basic position. Having barred conclusive disproof, he has provided no substitute for it, and the relation he does employ remains that of logical falsification. Though he is not a naïve falsificationist, Sir Karl, may, I suggest, legitimately be treated as one" (Kuhn, 1977). Robert Nozick was even more dismissive, making note of the inductive properties of Popper's anti-inductive philosophy and finally labeling Popper as "incoherent" (Nozick, 2001).

Hypotheses and Experiments
If the hypotheses used in current scientific practice are in fact held up for verification, this would seem to reinvigorate the same concerns about hypotheses that Bacon and Newton delineated. The hypothesis may be "dangerous"—it may be used to filter data and induce bias. Therefore, scientists operating in an inductive framework might be insulated from an impulse to defend an unproven premise by adhering to a bottom-up method, producing experimentally derived data in order to build

a model, and then subjecting the model to tests for its ability to predict the future. If the model passes such a test, its inductive power is demonstrated.

How then might one frame the first experiment, before sufficient data are gathered to produce a model? Absent information as to how a process works, it would seem that a question is the appropriate tool because the question, as opposed to a hypothesis, properly identifies the scientist as being in a state of ignorance when data are absent (Glass, 2006). The question is then used as a basis to accumulate data. From the data one then builds a model, which can be subjected to tests for its inductive ability. As Newton noted, one gathers the "negatives and affirmatives" to refine the model, until a predictive construct is derived. Such a methodology would eliminate the "hypothesis" term and substitute the "question" for settings where experiments are performed before sufficient data exist and the "model" for situations where the scientist is working with sufficient data to produce a construct that can be tested for inductive power.

The Critical Rationalist in Medicine
For those working in biology, and particularly in medicine, it may be worth noting that there are ethical problems that might accrue if inductive reasoning is abandoned. Clinical trials are conducted not simply to determine if a particular treatment works in an isolated instance but more specifically to determine whether that treatment will be effective when generalized to other patients with that illness. Conducting clinical trials constitutes an explicit effort to capture enough experience to accurately predict whether a treatment will be beneficial to the general public. If it is said that a clinical trial is established as an example of hypothesis falsification, one might fairly ask how an inductive conclusion could be made. In particular, it is not clear how one could ethically espouse that a treatment be advocated for a patient if one adheres to the idea that inductive conclusions are not rational. There would be no basis to say that the prior experience with the drug in a clinical trial setting might be applied to a future patient's case.

Of course, such an ethical concern would not be a reason to embrace inductive reasoning were inductive reasoning not workable. One accepts inductive reasoning because it can be demonstrated that past experience is predictive of future outcomes, within a range of probability. It is this issue of probability that has been so troublesome to many philosophers. Biological induction necessarily contrasts with the absolute predictability one might find in physical laws, properly circumscribed. However, others have accepted probability as sufficient for claims of causality. Therefore, as long as one gains a sufficiently large data set that is representative of the variations observed in the clinic, and achieves this by demonstrating the reproducibility of the data, one can join in the inductive project with a clinical trial outcome.

The relevance of a philosophical posture for a physician or scientist is made clearer by considering the patient who seeks treatment for his disease. The patient who receives a new treatment based on the results of an experimental trial, and as a result survives (as revealed by the experience of those patients who were not similarly treated), must continue to be offered the therapy and would not be well served by a philosophical program that claims this past experience is irrelevant. Therefore, the clinical trial should not be framed with a hypothesis aimed at falsification but rather should be explicitly inductive because the project of the physician is explicitly inductive.

When one turns to basic biology, the issue is straightforward. Here, the scientist must ask whether verification or falsification is being sought. If the scientist concludes from a result that a similar result will occur, or that the data will establish a rule for how things work, then the project is inductive. If the experimental design is inductive, then the Critical Rationalist framework is inconsistent with the project that is being performed, and the hypothesis should be abandoned in favor of either a question or, if sufficient data are available, a model. We propose that building hypotheses should be abandoned in favor of posing a straightforward question of a system and then receiving an answer, using that answer to model reality, and then testing the reproducibility and predictive power of the model, modifying it as necessary.

Finally, when confronting a project where comprehensive data sets are accumulated, such as genome sequences, a hypothesis may not even be feasible. What would such a hypothesis be? Only a question is required: "What is the sequence of genome X?". This would be followed by an answer, the genome sequence, that can be tested for reproducibility by further sequencing, allowing for an increasingly improved model for the genome. Thus, although a hypothesis might have been thought to be necessary in the past, it no longer seems to be so. It is better to see science as a quest for good questions to try to answer, rather than a quest for bold hypotheses to try to refute.

REFERENCES

Anderson, C. (2008). The End of Science. Wired, July 16, 2008. http://www.wired.com/science/discoveries/magazine/16-07/pb_intro.

Bacon, F. (1620). Translation used: The Novum Organum, or a True Guide to the Interpretation of Nature, 2005 (Cambridge: Cambridge University Press).

Blake, R.M., Ducasse, C.J., and Madden, E.H. (1966). Theories of Scientific Method (Seattle, WA: Washington University Press).

Carnap, R. (1980). Studies in Inductive Logic and Probability (Berkeley, CA: University of California Press).

Glass, D. (2006). Experimental Design for Biologists (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Hall, A.R. (1962). The Revolution in Science 1500–1750 (London: Longman).

Hall, N., and Hájek, A. (2001). Induction and Probability. In Blackwell's Companion to Philosophy of Science, P. Machamer and M. Silberstein, eds. (Blackwell), pp. 149–172.

Hume, D. (1749). Text used: An Enquiry Concerning Human Understanding, 1999 (Oxford: Oxford University Press).

Kuhn, T.S. (1977). The Essential Tension; Selected Studies in Scientific Tradition and Change (Chicago, IL: University of Chicago Press).

Newton, I. (1721). Text used: Opticks, or A Treatise of the Reflections, Refractions, Inflections, and Colours of Light, 1979 (New York: Dover Publications).

Newton, I. (1729). Text used: The Principia, A New Translation by I. Bernard Cohen and Anne Whitman, Preceded by A Guide to Newton's Principia by I. Bernard Cohen, 1999 (Berkeley, CA: University of California Press).

Nozick, R. (2001). Invariances, the Structure of the Objective World (Cambridge, MA: Belknap Press).

Poincaré, H. (1952). Science and Hypotheses (New York: Dover Publications).

Popper, K. (1959). Text used: The Logic of Scientific Discovery, 2002 (London: Routledge).

Russell, B. (1912). Text used: The Problems of Philosophy, 2007 (New York: Cosimo Classics).