

A Critique of the Hypothesis, and a Defense of the Question, as a Framework for Experimentation

David J. Glass

Scientists are often steered by common convention, funding agencies, and journal guidelines into a hypothesis-driven experimental framework, despite Isaac Newton's dictum that hypotheses have no place in experimental science. Some may think that Newton's cautionary note, which was in keeping with an experimental approach espoused by Francis Bacon, is inapplicable to current experimental method since, in accord with the philosopher Karl Popper, modern-day hypotheses are framed to serve as instruments of falsification, as opposed to verification. But Popper's "critical rationalist" framework too is problematic. It has been accused of being: inconsistent on philosophical grounds; unworkable for modern "large science," such as systems biology; inconsistent with the actual goals of experimental science, which is verification and not falsification; and harmful to the process of discovery as a practical matter. A criticism of the hypothesis as a framework for experimentation is offered. Presented is an alternative framework—the query/model approach—which many scientists may discover is the framework they are actually using, despite being required to give lip service to the hypothesis.

© 2010 American Association for Clinical Chemistry

In the early 1600s, Francis Bacon, the Lord Chancellor to King James I of England, was arrested for accepting bribes, briefly imprisoned, and forbidden thenceforth to hold office. He thus found himself with ample time to write his promised treatise on scientific method. Bacon's *Novum Organum*, or *New Method*, was so titled to draw a distinction between his work and Aristotle's *Organon*, which he criticized for its nonempirical approach to scientific exploration (1).

Bacon's greatest contributions to scientific method can be boiled down to 2 main facets: (a) a call for an inductive, rather than a deductive, approach to

science and (b) an advocacy for a reliance on experiments rather than dogma for induction.

"Induction" refers to 2 distinct types of reasoning, both of which are germane to the scientist. First, induction means that what happened can be said to be predictive of what will happen, so that one can induce from the experience that a released apple falls toward the ground that the next time one releases the apple it will fall again. Second, induction refers to the ability to generalize a result, which is based on a principle uncovered by experiments on a specific case, so the fact that an apple fell might allow the scientist to induce that an orange would fall toward the earth as well, even though one did the original experiments on an apple, not an orange (2).

As for then-current practice, Bacon further pointed out the problem with starting with an unproven premise and deducing rules from that premise. If the premise is unfounded, then the resulting deductions from that premise will be equally unfounded. Although he did not explicitly use the term "hypothesis," it was the thing he was criticizing, hypothesis being defined as "whatever is not derived from phenomena, an unproven premise, advanced without evidence, as a tentative explanation" (3). The hypothesis as a framework was in turn explicitly rejected by Isaac Newton, a Baconian, despite the fact that the first edition of Newton's *Principia* had been organized around hypotheses (4). It has been argued that Newton's evolution from an alchemist to a scientist forced him to move away from the hypothesis construct in later editions of his great work and then in his *Opticks* to eschew the hypothesis in favor of rules he could prove, rules derived inductively from experiments. Newton wrote in the *Principia* (4) (translated from Latin):

I have not as yet been able to discover the reason for these properties of gravity from phenomena, and I do not frame hypotheses. For whatever is not deduced from the phenomena must be called a hypothesis; and hypotheses, whether metaphysical or physical, or based on occult qualities, or mechanical, have no place in experimental philosophy. In this philosophy particular propositions are inferred from the phenomena, and afterwards rendered general by induction.

Furthermore, when Newton was writing the *Opticks*, he started out Part I as follows: "My design in this book is not to explain the properties of light by

Novartis Institutes for Biomedical Research, Cambridge, MA.

* Address correspondence to the author at: Novartis Institutes for Biomedical Research, 100 Technology Square, Cambridge, MA 02139. E-mail david.glass@novartis.com.

Received April 5, 2010; accepted April 7, 2010.

Previously published online at DOI: 10.1373/clinchem.2010.144477

hypotheses, but to propose and prove them by reason and experiments” (5).

Bacon explained the problem with starting with an unproven explanation before one did actual experiments (6) (translated from Latin):

Once a man’s understanding has settled on something, . . . it draws everything else also to support and agree with it. And if it encounters a large number of more powerful countervailing examples, it either fails to notice them, or disregards them, or makes fine distinctions to dismiss and reject them, and all this with much dangerous prejudice, to preserve the authority of its first conceptions. . . . And even apart from the pleasure and vanity we mentioned, it is an innate and constant mistake in the human understanding to be much more moved and excited by affirmatives than by negatives, when rightly and properly it should make itself equally open to both. . . .

It is worth taking some time to unpack Bacon’s aphorism into a series of arguments and apply them to present-day science.

1. It Is a Mistake to Frame an Experimental Project with a Hypothesis Because the Experimentalist Will Filter Data through the Lens of That Hypothesis, Rejecting Contradicting Evidence in Favor of Validating Evidence.

Such an action may not happen because of malfeasance, but rather because the existence of a hypothesis drives a particular type of experimental methodology, including a filter for data interpretation. It creates the expectation of a particular result and thus implicitly rejects other possibilities before the actual experiment is performed.

For an illustration of the issue, consider the hypothesis that caffeine increases blood pressure (7). It should be clear that this hypothesis creates the expectation of a particular result, an increase in blood pressure, as opposed to a decrease or no change. Therefore, it should not be controversial to point out, first, that a scientist starting out with such a framework may be relatively more likely to search for a setting in which the premise is found than if the framework was simply the question, “What is the effect of caffeine on blood pressure?” Second, this hypothesis actually inculcates a bias in favor of discovering an increase in blood pressure upon treatment with caffeine by forcing the scientist to look for an increase, as opposed to a decrease. In other words, this hypothesis establishes a data filter that asks the scientist to determine whether an increase in particular happened and therefore forces methodology to determine an increase in particular.

What might a scientist do to “confirm” the hypothesis that caffeine increases blood pressure? One might keep increasing the caffeine dosage until a “pos-

itive” result is seen, rejecting the finding that lower doses of caffeine do not significantly increase blood pressure as “negative” and therefore not germane. Next, a scientist with a requirement to validate this particular hypothesis might accept even very small increases in blood pressure as positive evidence, even if such increases are not physiologically relevant. Third, any discounting evidence, such as the lack of a dose response, might be ignored in light of single “positive” points. All of these actions seem more likely if an investigator has put in place a preexisting requirement to arrive at a particular result, which a hypothesis as a framework seems to establish more than does a question.

2. The Hypothesis Establishes a Dysfunctional Positive/Negative Binary, in Which the Scientist Is Often Forced to State That the Experiment’s Goal Is Negation, When the True Goal Is Affirmation. Furthermore, the Hypothesis As Currently Used Rejects As Untenable Inductive Reasoning, a Rejection That Is Itself Untenable.

The first claim, that the true goal of the scientist is affirmation, is straightforward to demonstrate. Consider the entire set of clinical trials. Can anyone doubt that hypothesis verification and not falsification is the goal of these trials? Why would anyone spend hundreds of millions of dollars on such an experiment if the goal were to disprove the framing hypothesis? In addition, what would be the point of doing a clinical trial if it were unacceptable to induce from the results, if it were unacceptable to claim that the trial predicts future outcomes? This issue does not apply just to medical trials. Consider any scientist publishing in one of the top journals. What is the percentage of reports that claim premise falsification, as opposed to a new finding that demonstrates a rule or principle thought to stably explain how things work?

Given this simple reality, that scientists do experiments usually because they want to derive an experimental model that has inductive, or predictive, power—that scientists use their data inductively—why are we often made to frame our experiments within a philosophical paradigm that explicitly rejects inductive reasoning?

How did this situation develop? This question is complicated and probably requires a book rather than an essay to answer in detail. But here is the bare-bones tale. Shortly after Newton died, David Hume, in *A Treatise of Human Nature*, rejected inductive reasoning by declaring that the fact that something behaved in a certain way in the past is no guarantee that it will do so in the future, rejecting even probability as a rationale for engaging in inductive reasoning (8, 9). For the

reader to appreciate how extreme Hume was in this claim, it is worth quoting several lines in full.

Your appeal to past experience decides nothing in the present case; and at the utmost can only prove, that that very object, which produced any other, was at that very instant endowed with such a power; but can never prove, that the same power must continue in the same object or collection of sensible qualities; much less, that a like power is always conjoined with like sensible qualities. Should it be said, that we have experience, that the same power continues united with the same object, and that like objects are endowed with like powers, I would renew my question, why from this experience we form any conclusion beyond those past instances, of which we have had experience. (9)

This point was debated frequently for the next 150 years, the most frequent rejoinder being that probability was rational and allowed one to make predictive statements within appropriate constraints (10–12). Yet, to be consistent with Hume's proscription against induction, the philosopher Karl Popper in the 1930s espoused an experimental method that reinvigorated the hypothesis but advocated its use as a pure instrument of falsification, because even if one cannot say something will happen, at least one can say with certainty whether a thing did or did not happen (13). This approach placed Popper in a philosophical school known as critical rationalism.

Critical rationalism has obvious attractions, not the least of which is the pure mathematical ability to disprove something definitively, as opposed to the relative inability to absolutely prove a thing. Only a single negative example is required to disprove a hypothesis, whereas an infinite number of confirming trials still leaves open what might happen the next time the experiment is performed and thus leaves the scientist with no more than an extremely strong statistic. Therefore, although it is easy to see why the hypothesis-falsification terminology gained currency—because of its rigor and its utility at framing an issue so that it can be tested via experimentation and apparently shown at least to be false—it is this very attractiveness as a model of pure falsification that can be said to be deceptive, if that is not how the paradigm is actually being used.

What is more, Popper's approach has been shown to be unworkable, for many reasons but not the least of which is that it is not possible to avoid inductive reasoning (14, 15). In addition, falsificationalism can be subjected to the same semantic maneuvers as verification and thus does not offer the scientist any relative safety (16).

Because scientific experimentation as a rational endeavor must rely on the predictability of prior experience to proceed and because it can be demonstrated that sufficient experience is in fact reproducibly predictive of the future, within certain probabilities, and

within certain logical limits and parameters, one is left with inductive reasoning as the only alternative for science. Even if one cannot say with mathematical certainty that reality is stable, the repetition of results (such that statistics and probabilities can be determined) does in fact constitute control experiments for the time variable. The scientist demonstrates through repetition that a model is predictive by actually using it to predict an outcome and finding that the prediction is verified (again, one can continue to object that the repetition still does not predict the next iteration, which is why one must always proscribe one's claim of verification within certain probabilities). Lest you still decry this approach as irrational in absolute terms, stop and ask, when you get up in the morning, how you base the decision to steel yourself against gravity. How do you justify the expectation that a car will move when you insert a key into its ignition? How do you justify the expectation when you get on board an airplane that it will take off and land safely? All of these actions are acts of induction. Induction is the reasoning by which one derives the simple expectation that the sun will rise and then set within a set time each day. It is a particular predictive model of reality that is verified on a daily basis. Induction can thus be shown to be more than a matter of necessity or of pure convenience but actually to be workable for conclusions as to what reality is, and its stability, within accepted constraints.

3. The Insistence on a Preexisting Hypothesis That Can Be Held Up for Verification Stifles Innovation.

The US NIH requires hypotheses for most of its grant applications (17–19). Although hypotheses are constructs that need not be derived from phenomena, put forth as a tentative explanation, to justify funding for experiments that might derive the relevant phenomena, NIH review committees often require that sufficient preliminary evidence be shown to demonstrate the hypothesis is probably correct (3, 18, 19).

This situation establishes a couple of things. First, the term "hypothesis" is not being used correctly because prior experimental evidence of the premise is demanded. Second, because the hypothesis is often being used as it was in the 1500s, as a premise to frame an exercise in verification, the concept again becomes vulnerable to all of the charges originally posed against it, that it is an instrument of bias.

It might be to avoid this last problem, that the hypothesis could lead the scientist astray, that granting organizations shield themselves from the potential hazard of an ill-founded premise by insisting that the scientist know in advance that almost the entirety of what he or she proposes has already been demonstrated, that, in fact, the experiments have already been

performed. This situation is not only backwards, it is stifling, as Gina Kolata pointed out in the *New York Times* last year (19):

The institute's reviewers choose such projects because, with too little money to finance most proposals, they are timid about taking chances on ones that might not succeed. The problem, Dr. Young and others say, is that projects that could make a major difference in cancer prevention and treatment are all too often crowded out because they are too uncertain. In fact, it has become lore among cancer researchers that some game-changing discoveries involved projects deemed too unlikely to succeed and were therefore denied federal grants, forcing researchers to struggle mightily to continue.

The current granting system can be criticized as being worse than contradictory—as being actually incoherent—if on the one hand a preexisting hypothesis is mandated as an experimental framework for a grant application and then on the other hand grant applications that are less than absolutely certain to verify what is already believed to be true are rejected. What are the chances that something novel will be found and recognized, with all these layers of inculcated preexisting bias?

4. Big Science Cannot Be Framed with Hypotheses.

One development over the last dozen years has been the creation of the field of systems biology, in which scientists seek to interrogate systems more comprehensively. One example of a systems biology experiment is a set of gene expression studies in which changes to every gene in an organism can be examined in response to a series of perturbations. Another type of systems biology project could revolve around proteomic experiments, in which, for example, scientists investigate how the phosphorylation states of all proteins are altered after a stimulus. A third example is simple gene or mRNA sequencing, in which “deep sequencing” can be done to determine subtle epigenetic changes.

Such systems biology experiments cannot be usefully framed with a hypothesis aimed at falsification. The requirement to do so can be comical and is certainly unhelpful to experimental design.

The NIH deals with this “problem” by labeling such experiments as “hypothesis generating” experiments (20); however, if a large systems biology experiment is not governed by a hypothesis but is performed so that someone downstream might create hypotheses, this is proof that hypotheses are not required for big science. Furthermore, if one does not need a hypothesis to sequence a genome, why does one require a hypothesis to enter into an inquiry concerning a particular gene?

5. Hypotheses Appeal to the Ego, or to Vanity, and Are Therefore Dangerous.

One need not dwell on this point too long, but it should be mentioned because it was one of Bacon's more incisive arguments. The notion that the scientist does not actually need to do the experiment to derive the answer—that the scientist is so clever as to figure out the answer in advance so that the experiment is simply an act of confirmation—is quite seductive. After all, one might fear that if all the scientist is doing is posing a question and therefore is requiring nature to deliver the answer, then what is the scientist other than a specialized kind of accountant? Compare the role of a pure observer and recorder (and, in some special instances, one who is especially insightful in deriving a principle from recordings that can be shown to be predictive—even if this last instance is noteworthy) with that of the diviner, who can intuit the truth without the help of nature to point the way. This is the seductive appeal of the hypothesis.

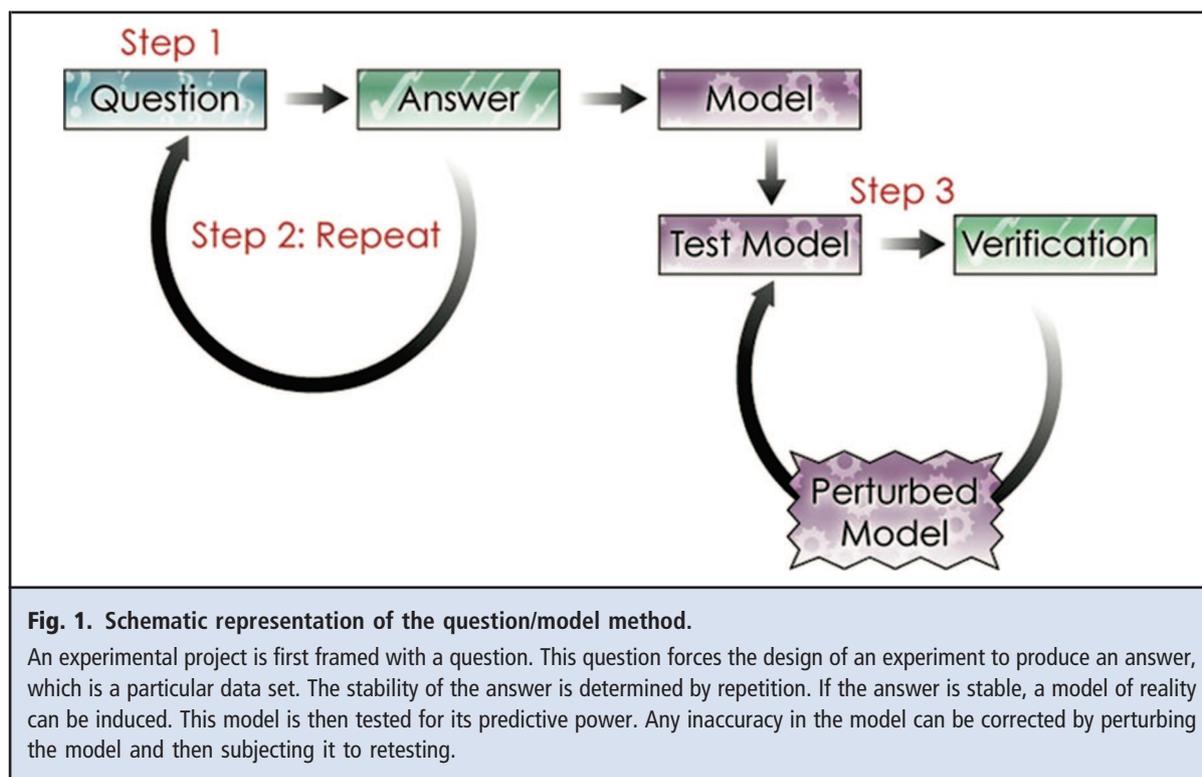
One does not need to go further than the sixteenth century literature to see the folly of such an approach. Hypotheses made without the help of nature are doomed to failure, and the desire to prove such a construct true places an individual so inclined more in sync with a preacher than a scientist.

Indeed, some might argue that one does not need to go back to the sixteenth century. The scientific literature of the twenty-first century has no shortage of hypotheses claimed to be true that turn out to be false due to the variety of problems listed, not the least of which is vanity.

If these criticisms have merit, if hypotheses are unhelpful, disadvantageous, inconsistent, or unworkable as instruments of either falsification or verification, what should the scientist use as an alternative?

The way that science seems to be actually done productively is to first conduct an experiment in response to a question, rather than a hypothesis (Fig. 1). Why else would a scientist seek to experiment and determine how nature works unless the scientist had a question about how it works? A question functions as an adequate framework for an experiment if it is posed so that it can be answered with an experiment (2).

The answer to a question might be a set of data, and from these data the scientist can build a model. A model differs from a hypothesis in several fundamental ways. First, it is data derived. Second, it can be explicitly tested for its predictive, or inductive, power. Third, it exists in a framework that accepts inductive reasoning. Fourth, it can be modified on the basis of new data and not just falsified/rejected or affirmed/accepted in a purely binary manner. Finally, a model can be said to



be correct within a probability range. It does not have to be absolutely correct, as long as the stated probability is verifiable.

A query/model approach is also appropriate for big science. For example, one can simply frame a proteomic experiment with a question, such as “How does the proteome change in response to X?” The resulting data set can then be tested for its predictive power as a model by asking whether the proteome changes the same way in subsequent experiments (Fig. 1). From that model of proteomic changes, one can determine the model’s inductive power, or generalizability, to other perturbations or settings by asking further questions, all without ever putting a hypothesis in place (Fig. 1).

The query/model approach seems to be how much of science is actually done, which begs the question, Why aren’t scientists more explicit about this?

It is time scientists embrace their role as inductive agents. An exploration of the appropriate limits and weaknesses of such a framework will afford the scientist an appropriate way both to frame experiments and to analyze data. A question as an initial framework admits the scientist into new areas where little is known, whereas the requirement for a pre-proven hypothesis explicitly closes the door to exploring the unknown.

Having a question as the initial framework for a project in advance of any experimentation also seems

to have a useful humbling effect on the scientist. The act of posing a question forces the scientist to admit that the answer is not yet known and that the question therefore requires an experiment. From the act of doing the experiment, one accumulates data that can be used to build a model (Fig. 1). The model seems to be the more appropriate framework for basing predictions, because it is explicitly derived from experiential data, and that experience can then be queried for its inductive power. This approach seems an accurate description of how science today is actually done and how one can use large unknowns to pose big questions, which demand novel and exciting experimentation. For example, there should be no shame in posing a question like, What is the cure to cancer? The greater shame is the requirement that we claim to know the answer in advance of engaging in such an exploration, a requirement that will limit us to what we already think we know. In contrast, the question will demand that we search for an answer we do not yet have, freeing us to see what we might discover.

Some might object that such a huge question does not point to an experimental design, but it does frame a large project and alerts the scientist to a great unknown. Subsequent to such a large framework question, the scientist can ask a more focused question, such as, What genetic markers coassociate with sensitivity and with resistance to a particular treatment? The

answers to these questions first would help in tailoring medical care and, second, would point scientists to the potential mediators of treatment resistance that still require their attention. Once these mediators are discovered, new mechanistic questions can be asked—all in a particular focused program directed toward discovery and innovation. This approach seems to be the actual process of science, a process that for good reason is called “scientific inquiry” and that at no stage requires a hypothesis.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 re-

quirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures of Potential Conflicts of Interest: No authors declared any potential conflicts of interest.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

Acknowledgments: I thank my colleagues at Novartis for their support, especially M. Fishman and B. Richardson, and the entire Muscle Group. Thanks also to S. Hall, E. Anders, D. Perkins, and A. Smart for their critical reading and suggestions. Thanks to A. Abrams for artwork.

References

1. Bacon F. *Novum organum scientiarum*. London: Adriani Wyngaerden; 1650.
2. Glass DJ, Hall N. A brief history of the hypothesis. *Cell* 2008;134:378–81.
3. Barrow J. *A new and universal dictionary of arts and sciences*. London: John Whiston; 1751.
4. Newton I. *The Principia: mathematical principles of natural philosophy*. Cohen IB, Whitman A, translators. Berkeley: University of California Press; 1999. Translation of: *Philosophiae naturalis principia mathematica*.
5. Newton I. *Opticks, or a treatise of the reflections, refractions, inflections, and colours of light*. New York: Dover; 1979.
6. Bacon F. *The Novum Organon, or a true guide to the interpretation of nature*. Cambridge (UK): Cambridge University Press; 2005.
7. Glass D. *Experimental design for biologists*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2006.
8. Hume D. *An enquiry concerning human understanding*. Beauchamp TL, ed. Oxford (UK): Oxford University Press; 1999.
9. Hume D. *A treatise of human nature*. Oxford (UK): Oxford University Press; 2000.
10. Poincaré H. *Science and hypotheses*. New York: Dover Publications; 1952.
11. Russell B. *The problems of philosophy*. New York: Cosimo Classics; 2007.
12. Whewell W. *History of the inductive sciences from the earliest to the present time*. New York: Appleton; 1866.
13. Popper K. *The logic of scientific discovery*. New York: Routledge Classics; 2002.
14. Nozick R. *Philosophical explanations*. Cambridge (MA): Belknap Press; 1981.
15. Nozick R. *Invariances: the structure of the objective world*. Cambridge (MA): Belknap Press; 2001.
16. Kuhn TS. *The essential tension; selected studies in scientific tradition and change*. Chicago: University of Chicago Press; 1977.
17. National Institute of Allergy and Infectious Diseases. Part 2. Game Plan. <http://funding.niaid.nih.gov/ncn/grants/cycle/part02.htm> (Accessed May 2010).
18. Loscalzo J. The NIH budget and the future of biomedical research. *N Engl J Med* 2006;354:1665–7.
19. Kolata G. Grant system leads cancer researchers to play it safe. *New York Times* 2009 Jun 27.
20. NIH. Systems Biology, HIV/AIDS, and Substance Abuse (R01). <http://grants.nih.gov/grants/guide/rfa-files/RFA-DA-10-014.html> (Accessed June 2010). See “Part I overview information.”