Sci & Educ (2014) 23:2069–2093
DOI 10.1007/s11191-014-9704-z

"The Scientific Method" as Myth and Ideal

Brian A. Woodcock

Published online: 27 May 2014

Abstract "The Scientific Method" as it has been portrayed in popular and introductory contexts has been declared a myth. The variation that one finds in introductory presentations of "The Scientific Method" is explained by the fact that there is no canonical account among historians and philosophers of science. What, in particular, is wrong with "The Scientific Method"? This essay provides a fairly comprehensive survey of shortcomings of "The Scientific Method". Included are corrections to several misconceptions that often accompany such presentations. Rather than treating "The Scientific Method" as a useful approximation or an ideal, the myth should be discarded. Lessons can be learned for introductory pedagogical contexts from considering the shortcomings of the myth.

1 Introduction

In his book *Scientific Literacy and the Myth of the Scientific Method* (1992), Henry H. Bauer claimed that the popular conception of how science works is misleading and deserves to be called a myth. Although there has been some improvement, the myth persists in popular and introductory accounts of science. After considering a number of misunderstandings about science and science literacy, Bauer's subsequent conclusion was that it should be considered an unattainable ideal. This raises the questions: What exactly is wrong with "The Scientific Method"? Should it be considered an ideal or an approximation, or should it be discarded? And, if the latter, what should replace it?

It was not new to conclude that the very notion of "The Scientific Method" is problematic. The former president of Harvard University, James Bryant Conant, writing as both a scientist and a historian of science had declared in 1951, "I believe almost all modern historians of the natural sciences would agree…There is *no* such thing as *the* scientific method. If there were, surely an examination of the history of physics, chemistry, and biology would reveal it…" (Conant 1951, p. 45). Similar assessments were reached by the

B. A. Woodcock (✉)
Department of Philosophy and Religious Studies, University of Wisconsin-Eau Claire, Eau Claire, WI 54702-4004, USA
e-mail: woodcoba@uwec.edu

philosophers of science Thomas Kuhn (1996 [1962]) and Paul Feyerabend (1975). Peter Caws, writing the entry on 'scientific method' for *The Encyclopedia of Philosophy*, summarized the logic of the situation:

> The term "scientific method," if applied to scientific investigation in general or to something allegedly embodied in the practice of every branch of science, can only refer to the lowest common denominator of a range of methods devised to cope with problems as diverse as classifying stars and curing diseases. If such a lowest common denominator exists—that is, if some recognizable characteristics are shared by the extremes of the continuum of methods plausibly called "scientific"—it can amount to little more than fidelity to empirical evidence and simplicity of logical formulation, fidelity to the evidence taking precedence in cases of conflict. However, these two overriding requirements for scientific activity do not constitute a specification of steps to be taken by scientists… (Caws 1967, p. 339)

In the widest sense, 'The Scientific Method' could refer to all of the activities that scientists do as scientists. But this yields too much. Scientists apply for grants, supervise employees, prepare budgets, and much more. In popular and introductory accounts, it is supposed that there is some core sequence of activities (the lowest common denominator that Caws wrote of) that is shared by paradigm examples of scientific practice. Under this conception, this sequence of activities is both descriptive and normative. The sequence purportedly describes what good scientists do. And, it is thought, it sets a standard for anything that would qualify as science.

Historically, there have been several attempts by philosophers and scientists at unearthing and communicating the *real* scientific method. These attempts share the belief that there is a single method yielding scientific knowledge and that this method can be discovered and articulated. (As we will see, whether there really is such a method and, if so, how to characterize it is still a discussion among contemporary philosophers of science.) Academic attempts at articulating scientific methodology are not the targets of this essay, however. The targets of this essay are popular and introductory descriptions of scientific methodology that utilize the label 'The Scientific Method'—what someone might described as the *so-called* scientific method. Scare quotes are used in much of this essay around the words 'The Scientific Method' to indicate that the concern is with current popular and introductory conceptions about scientific method that are described under this label.

Two comments about the use of the word 'myth' are in order. First, although in scholarly contexts calling some descriptive account a myth need not imply a judgment—positive or negative—about the accuracy of the account (e.g., consider academic discussions of religious beliefs as myths), in popular usage it often does carry the connotation of a negative assessment of accuracy. Clearly, because of his treatment of misunderstandings about science, Bauer's use of the term carries this connotation. Second, a descriptive account—but not the object the account attempts to describe—is the sort of thing that can properly be said to be a myth. For example, a particular account of George Washington's life may be said to be a myth; but George Washington's life is not a myth since it is not an account. Similarly, it is sensible to claim '"The Scientific Method" is a myth' or 'The so-called scientific method is a myth' but not 'The scientific method is a myth' if the phrase 'the scientific method' is used to attempt to refer to a supposed real process for acquiring scientific knowledge. People who make the latter claim probably intend their assertion to be interpreted in the former ways. (This is how one might make sense of Bauer's undifferentiated use of the phrase in his book.)

Are common presentations of "The Scientific Method" accurate portrayals of scientific activity and reasoning? Appealing to authority, this question can be answered quickly by noting that the experts—historians and philosophers of science (including philosophically-minded scientists)—do not think so. Although this answers the question, it does not provide insight into the specific shortcomings of such presentations. In particular, it does not address the issue of whether such presentations, though inaccurate, still deserve to be taught. One might imagine the following as a response to the claims of Bauer, Conant, Kuhn, Feyerabend, and Caws: "Of course, popular and introductory treatments offer an oversimplified view of scientific activity. But what they offer is a good approximation for pedagogical purposes." Bauer himself, after concluding that the common conception is a myth, advocated treating it as an unattainable ideal rather than discarding it.

What, in particular, is wrong with "The Scientific Method"? The aim of the following essay is to provide a fairly comprehensive survey of shortcomings of "The Scientific Method" as it is depicted in many popular and introductory contexts. Included are corrections to several misconceptions about scientific methodology that often accompany such presentations. The essay ends with a consideration of conclusions to be drawn for introductory pedagogical contexts. The following essay does not attempt to be a discussion of all misconceptions about science, however. There are plenty of misconceptions about science and scientists; not all of them are directly associated with "The Scientific Method". It is hoped that the following meditation might be of use to scientists, science educators, interested intellectuals, and students of science. It is presumed that historians and philosophers of science will be familiar with many of the ideas; it is hoped that they will find the following meditation a clear and faithful expression of what they are inclined to think on this topic.

2 Lack of Consensus

2.1 No Historical or Contemporary Consensus

The use of the definite article 'The' in calling some account "The Scientific Method" implies that there is one and only one—a single, unique method. This is problematic since there have been several views on scientific methodology historically. The two most influential on statements of "The Scientific Method" are inductivism and hypothetico-deductivism. The focus in both of these views is on the logical structure behind scientific reasoning.

Deductive reasoning is defined today as reasoning in which the reasons, or premises, necessitate the conclusion. In other words, it is impossible for the premises to be true and the conclusion false. For example, if the statements 'All dogs are mammals' and 'Fido is a dog' are both true, then the statement 'Fido is a mammal' must also be true. The first two statements cannot be true and the third one false. On the other hand, the term *inductive reasoning* is often used in a broad sense today to signify any form of non-deductive, or probabilistic, reasoning. More narrow uses of the term focus on inferences that extrapolate similarities from observational data to the unobserved. Because what we have not observed could be different from our data, it is logically possible in such reasoning for the premises to be true and yet the conclusion false. Historically, the form of inductive reasoning of greatest interest to philosophers thinking about scientific methodology is known as *inductive generalization*. From a number of observed instances—'Swan 1 is white', 'Swan 2 is white',…, 'Swan n is white'—a generalization is made—'All swans are white'. The generalization extends to cases beyond the observational data. In this particular example, the generalization turns out to be false; there are black swans. Unlike deductive reasoning, such reasoning is risky because it is possible to infer a false conclusion from the data.

The ancient philosopher, Aristotle, proposed a view of scientific method that combined both induction and deduction. On his view, scientific principles are the result of inductive generalizations made on the basis of observations. Scientific explanation consists in the deduction of particular events from very general scientific principles (Losee 1993).

In the seventeenth century, Francis Bacon sought to make corrections to the way Aristotle's view was being applied, especially as it concerned the inductive portions of Aristotle's methodology. Bacon is often credited, perhaps not accurately, with the view called *inductivism*. Likewise, Isaac Newton's methodological statements suggest that he took an inductivist view of scientific methodology.[1] Classical inductivism considered scientific knowledge to consist in scientific laws and principles built by inductive reasoning on the foundation of a wide variety of objective observations. In order to be objective, observations were to be free from bias and presuppositions.

Although it is sometimes claimed or suggested that the successes of science since the scientific revolution are due to the discovery of "The Scientific Method", a careful look at figures in this period reveals that they held a variety of opinions on scientific methodology. Rene Descartes, for example, wrote extensively on method at roughly the same time period as Bacon, but his views are an expression of his rationalist philosophy. He trusted the light of reason over the senses. Descartes's methodological recommendations appear more like steps for solving mathematics problems than for doing empirical science. In the period of the scientific revolution, there was no statement of scientific method that all significant contributors held.

In the nineteenth century, inductivism was opposed by a view that today is called hypothetico-deductivism. The idea is that scientists come up with hypotheses using whatever creative means they can. What matters is that these hypotheses have testable consequences that could be used as a basis to judge the hypothesis. Hypothetical reasoning had been recognized as a form of reasoning for a long time, but so long as science was viewed as delivering certain knowledge, it did not seem to be a method appropriate to science. It was well known that from different hypotheses, the same observational consequences could be deduced. So, the occurrence of those consequences could not be decisive evidence for one of the hypotheses. In the nineteenth century, people began to recognize that the knowledge the natural sciences deliver to us is not absolutely certain. That is, they began to embrace a *fallibilist* view of scientific knowledge. Hypothetical reasoning was no longer suspect simply because it failed to yield absolute certainty.

In the twentieth century, the philosopher of science Karl Popper advocated the hypothetico-deductive approach to understanding scientific reasoning but because of his distrust of non-deductive reasoning, he thought that the approach could only be used to falsify theories, not to confirm them. In other words, if a hypothesis yields a successful prediction, this does nothing in Popper's view to confirm (or support the truth of) the hypothesis. A successful prediction just shows that the hypothesis survived another test, avoiding falsification. Popper's view is called *falsificationism*. About the same time, other philosophers of science developed theories of scientific confirmation: Carl Hempel

[1] Newton thought of his own empirical propositions as being the result of the gathering of phenomena by induction. There are well-known criticisms of this method as an accurate portrayal of Newton's own activity, however. For example, Newton's First Law of Motion—the law of inertia—can hardly be a mere general summary of observations since no one has ever observed bodies that are not acted upon by any forces and yet the First Law tells us what would happen to material bodies in such situations (Whewell 1847, p. 49; Ladyman 2002, p. 55).

proposed a positive instance account and Rudolf Carnap worked on a probabilistic theory of inductive logic. By contrast, in the 1970s, the philosopher of science Paul Feyerabend advocated methodological anarchism, the rejection of the view that there are universal rules that describe and regulate good scientific activity (Feyerabend 1975).

In contemporary discussions of scientific methodology, perhaps the leading theory of scientific reasoning is Bayesian confirmation theory (Howson and Urbach 2005), but it is not without critics (e.g., Clark Glymour, Deborah Mayo, and John Norton). Inference to the best explanation (Lipton 2004) also receives a lot of attention but it too has critics (e.g., Bas van Fraassen). Currently, one does not find a consensus among philosophers of science about how to best understand what it is that scientists do, either in forming theories or in evaluating them.

2.2 Variations in Popular and Introductory Accounts

When one looks at popular and introductory accounts of "The Scientific Method", one finds some similarities but one also finds a surprising degree of variation. This is not what one would expect if there were some well-recognized, canonical account that they all were trying to reflect. They do not all agree on the number of steps (see Table 1).[2] One popular statement of "The Scientific Method" has as many as eleven steps (Edmund 2011).[3] They do not all agree on the starting point: Some begin with recognizing a problem (Marshak 2004); some begin with making observations (Bailey 2013; Edmund 2011; Pilar 1979). They do not all agree on the end point: Some end with taking action or publishing (Edmund 2011); others end before reaching this point. Although most accounts today mention forming a hypothesis and testing it—reflecting the influence of hypothetico-deductivism—older accounts can be found that make no mention of hypotheses and tests. Rather, they reflect inductivism (Pilar 1979). Since most contemporary presentations of "The Scientific Method" mention forming hypotheses and testing them, one might consider those two elements the core of "The Scientific Method" today.[4]

With any method, it is appropriate to ask what the method is intended to accomplish. Philosophers of science have traditionally distinguished how scientific ideas are formed (or generated) from how they are justified (or confirmed). That is, they have typically distinguished the *context of discovery* from the *context of justification* (Losee 1993, pp. 121–126; Ladyman 2002, pp. 74–76). Many presentations of "The Scientific Method"

[2] The outside choices in the table were selected to provide extremes. The inner ones are more representative of what one finds usually. As an assignment in a course, I have students find examples of popular and introductory descriptions of "The Scientific Method". After a few iterations of the course, I've collected a healthy stack of examples. A more systematic examination would be an interesting project and better support my claim about variation. (Blachowicz (2009) does something of this sort, looking at 70 textbooks.) For the moment, my support can rest with a challenge: If you doubt it, take an hour just to examine what you find being said about "The Scientific Method" online.

[3] A very enthusiastic presentation of "The Scientific Method" is to be found at www.scientificmethod.com. The website is authored by Norman W. Edmund, the founder of Edmund Scientific—a company well known as a supplier of optical and other scientific gadgetry for educational and amateur uses. The website claims to offer "today's most up-to-date, complete, clear, concise, and reliable information about the scientific method and scientific method activities that has ever been offered." It presents "The Scientific Method" in eleven steps or stages along with three supporting ingredients.

[4] Recognizing that "different sources describe the steps of the scientific method in different ways", William Harris (2012) assures his readers that "Fundamentally, however, they incorporate the same concepts and principles". The view that presentations of "The Scientific Method" say the same things essentially is rather typical. But it is accurate only if one neglects a number of details.

Table 1 Four presentations of steps of "The Scientific Method"

Chemistry—The universal science Pilar (1979)	Essentials of geology Marshak (2004)	About.com Bailey (2013)	Scientificmethod.com Edmund (2011)
The accumulation of observations and their classification into various categories	Recognizing the problem	Observation	Curious observation
The summarization or generalization of the observations in the form of succinct statements called laws.	Collecting data	Question	Is there a problem?
An attempt to find some overall pattern or design which, if assumed to exist, accounts for the laws in a deductive manner	Proposing hypotheses	Hypothesis	Goals and planning
	Testing hypotheses	Experiment	Search, explore, and gather the evidence
		Results	Generate creative and logical alternative solutions
		Conclusion	Evaluate the evidence
			Make the educated guess (hypothesis)
			Challenge the hypothesis
			Reach a conclusion
			Suspend judgment
			Take action

today appear to be a blending of elements from both inductivism and hypothetico-deductivism (for example, Marshak 2004; Bailey 2013). Emphasizing observational data, they begin with something akin to an inductivist statement about how to form a hypothesis. They then proceed to describe in hypothetico-deductive terms how to evaluate and ultimately justify the hypothesis with some sort of empirical test. That is, they present an inductivist method of discovery fused to a hypothetico-deductivist method of justification.[5] Although this is one way to maintain that there is a single method, philosophers of science would be inclined to distinguish the two modes of reasoning—inductive and hypothetico-deductive—as well as the two aims—discovery and justification.

When one looks, then, what one finds is (a) there is significant variation in contemporary popular and introductory presentations of "The Scientific Method". Compare this situation with textbook accounts of classical Newtonian mechanics. Uniformly, these presentations follow Newton's own at some level. Every presentation includes the same number of laws of motion—three laws (not four)—and includes them in the same order. The variation in presentations of "The Scientific Method" seems to be explained partly by the fact that (b) there is no canonical statement of scientific methodology historically and

[5] In their general order, these blended statements of method are not unlike John Stuart Mill's description of the steps—(1) direct induction, (2) ratiocination, and (3) verification—of what he called "the Deductive Method" (Mill 1881), albeit less subtle. This does raise the issue of how to individuate methods. When a method requires induction for the development of a hypothesis while hypothetico-deductive reasoning is required for its evaluation, is this one method involving two kinds of reasoning or two methods utilized in a sequence?

(c) there is no consensus account of scientific methodology among contemporary philosophers of science.

In summary, the use of the definite article 'The' in the label "The Scientific Method" implies that there is a single method that scientists use. But, introductory presentations do not agree in detail on what that one method is and neither have historians and philosophers of science. In what follows (Sect. 3), we will first consider some general misconceptions—namely, that the method is mechanical and individualistic—that were the target of criticism by Bauer in his book. Since it is very common for statements of "The Scientific Method" today to begin with an inductivist-inspired procedure for the discovery (or invention) of a hypothesis followed by a hypothetico-deductivist description of how to evaluate such hypotheses, we will turn next (Sect. 4) to consider problems with the inductivist method of discovery. Subsequently (Sect. 5), we will explore various problems that arise for presentations that take a hypothetico-deductive account of hypothesis evaluation as the core of "The Scientific Method". The latter leads to considering and correcting some common misconceptions about hypotheses and reasoning (Sect. 6). In Sect. 7, Bauer's claim that "The Scientific Method" should be considered an ideal is analyzed and evaluated. Finally, the essay ends with some suggestions for introductory pedagogical contexts.

3 General Misconceptions

In his book, *Scientific Literacy and the Myth of the Scientific Method*, Henry H. Bauer undertook to correct a number of misunderstandings about science and science literacy. Significantly, Bauer made the claim that "the scientific method is a myth" (p. 39).[6] One reason that Bauer had for denying that there is a single method for doing science stemmed from his belief that "there is not any single thing that one can usefully and globally call science; rather, there are many different sorts of science" (p. 28).

To be clear, Bauer used the phrase 'the scientific method' ambiguously. At times he used the phrase very narrowly to refer to the popular conception of a formula or set of rules that scientists follow (p. 82). At other times, his discussion and criticisms of "The Scientific Method" seemed to comprehend something much broader—the whole range of "popular ideas about how science works" (p. 23). The broader target includes not only the conception of (a) a formula or set of rules for doing science but also popular conceptions about (b) the consequences for science of the use of the formula, and (c) general character traits of scientists themselves. Because he did not distinguish these, the general impression in the book is that a criticism of the popular conception of any of these items contributes support for the conclusion that "the scientific method is a myth". It is at least a logical possibility, however, that the popular formulaic conception of the scientific method presents an accurate description of scientific practice even though there are several popular misconceptions about its consequences for science and about scientists themselves.

When summing up his position, the reasons Bauer explicitly referred to in support of his conclusion that "the scientific method is a myth" are (1) the so-called method does not explain the success of science and (2) scientists in practice do not follow the method (pp. 39, 147). Notice that (1) concerns a conception about the consequences of following the method. Only (2) deals with the accuracy of the so-called method as a description of what

[6] Bauer did not capitalize the words 'the scientific method' nor did he put them in scare quotes, as is the convention in this essay. However, it is clear that his intended target in the book is the so-called scientific method; he used the latter phrase for the title of Chap. 2.

scientists do. Since he never targets a specific presentation of steps for criticism, his claim that scientists do not actually follow the method can seem puzzling. Did he mean to suggest that scientists do not frame hypotheses? That they do not carry out careful tests? Bauer never presented reasons to think that there is anything wrong with these as descriptions of scientific activity. What he did criticize is the idea that (A) scientists always rank empirical data more highly than theory. What scientists do not follow, according to Bauer, is the conception that scientists are "supposed to put evidence first and theorizing second" (p. 23). To correct this, Bauer pointed to cases of scientists who held to a theory in spite of data that seemed to tell against it and were vindicated in the end for doing so. [For more that bears on (A), see Sects. 5.2–5.4 below.] The other misconceptions that Bauer addressed concerning the narrow formulaic understanding of the so-called scientific method are more general—that (B) scientific activity is impersonal and mechanical and that (C) scientific activity can be successfully performed by solitary individuals.

3.1 Formulaic and Mechanical

Pedagogical presentations of "The Scientific Method" often make it seem like a formula or a recipe for doing science, thereby encouraging the impression of science as an impersonal mechanical process. This is presumably motivated by the desire to present science as objective. (It may also have pedagogical roots in the need for a procedure to guide students through a science project.) Rather than being impersonal and mechanical, however, Bauer emphasized the degree to which scientific activity requires subjective judgments—for example, about what hypotheses to test, about what research to do, about when support for one's views is strong enough to warrant publication, and about which details to publish and which to withhold (Bauer 1992, pp. 81–82).[7] In fact, Bauer argued plausibly that science functions better because of individual differences. Some scientists are more open-minded and some are more skeptical. The former, Bauer labeled "discoverers"—those readily pursuing and even accepting bolds ideas on slight evidence—whereas the latter he labeled "kibitzers"—those requiring a lot of evidence to shift their commitment from the conventional wisdom (Bauer 1992, pp. 54–55). Science as a whole benefits from the inclusion of both types of individuals. In addition to these points, as we will see, the creation of scientific ideas is not something that can be reduced to a formula; and, likewise, because of the actual complexities involved in evaluating a hypothesis, scientific reasoning resists reduction to a simple formula as well.

3.2 An Individualistic Stereotype

Introductory accounts of "The Scientific Method"—especially those aimed at preparing young students to create a science fair project—often portray it as something like a scientist's job description. "Anyone who is inquisitive can be a scientist. All you have to do is answer questions by following a simple, logical and straightforward prescription that's called the scientific method" (Cowens 2006, p. 42). This fits with the stereotype of the lone scientist working in a lab first formulating hypotheses and then testing them. But science, as it is practiced today, is often a social enterprise. Many scientists work on their research project as part of a team. Large science experiments involving particle accelerators, for

[7] The subjectivity described here is not opposed to rationality. The judgments described are not ones that are characteristically whimsical or a matter of mere taste; they are subjective only in the sense that they depend upon individual expertise.

example, can involve hundreds of scientists each devoting their career to some particular aspect of the experiment. The degree to which the lone scientist stereotype is accurate varies a great deal across the sciences. The scientist who formulates a hypothesis may not be the same one who devises a test of the hypothesis; and, the one who devises the test may not be the one to perform it.[8] Bauer pointed out that, because scientists become invested in their ideas, the best tests of a hypothesis are not usually produced by its originators. (In his book, Bauer considered the case of cold fusion an illustration of this pp. 52–54.)

To correct this individualistic misconception about science, Bauer latched onto two analogies—the jigsaw puzzle and filter analogies—in order to illustrate the importance of social dynamics for the scientific enterprise. They are instructive and memorable. The first analogy is from Michael Polanyi. Scientific activity, it is claimed, is like a group of people putting together a jigsaw puzzle in sight of one another (Bauer 1992, pp. 42–44). Because everyone can see the attempted contributions by others, individual initiatives can be checked, extended, and corrected. Bauer saw in such cooperative action the explanation for the origins of modern science. Modern science had its origin in the seventeenth century he thought, not because of the discovery of some method that had not been employed before, but because it is then that "the puzzlers began to organize themselves, to specialize, to communicate rapidly with one another, and to act as critics for one another, to work on the puzzle together with increasing effectiveness" (p. 52).[9]

The other analogy that Bauer considered to be particularly instructive for understanding science is the knowledge filter (pp. 44–48). Scientific inquiry on some topic begins with all sorts of speculative ideas subject to all sorts of biases and preconceptions. Over time, these ideas are filtered by the scientific community. Peer review is the biggest factor here. Peer review decides which proposals get funded and which papers get published. Further filtering occurs as only certain ideas from the published literature make their way into the secondary literature of review articles and graduate-level textbooks. Finally, those ideas that make their way into undergraduate textbooks are usually accepted by the entire scientific community. The filter analogy helps to explain the different degrees of reliability one finds among ideas in the scientific community. Although scientific ideas at the frontier are often quite unreliable, those which survive the gauntlet of communal scrutiny to become textbook science are usually very reliable.

In summary, scientific activity is not reducible to the aggregation of single individuals each following a mechanical formula. Subjective judgments and personal differences matter. Moreover, the focus on the individual ignores the important role of social dynamics for challenging, refining, and filtering scientific ideas.

4 Problems with the Inductivist Method of Discovery

As noted previously, it is common for statements of "The Scientific Method" to begin with an inductivist-inspired statement of hypothesis formation. What is (potentially) misleading about such presentations?

[8] This is nicely recognized in (Purves et al. 1992, p. 7).

[9] Understanding the importance of social interaction for science also provided the basis for Bauer's way of distinguishing pseudoscience from science. An inquiry is pseudoscientific not because of a failure to follow the scientific method but when it is pursued in isolation from "the competent, relevant scientific community" (Bauer 1992, p. 60).

Inductivism in its classical expression united both a method of discovery with a method of justification. On the inductivist procedure, both activities occur simultaneously. The classical inductivist methodology begins with making a wide variety of objective observations. Whatever patterns are found in the observations are then to be inductively generalized into a principle or law. Since this describes how to come up with scientific principles, it is a method of discovery. And, since the resulting principles are purportedly justified, it is a method of justification. The latter assumes that inductive generalization is itself a rational form of inference—something that the philosopher David Hume famously denied.

4.1 Other Sources of Inquiry Besides Observation

The first step in many presentations of "The Scientific Method" involves observation, either as a direct precursor to recognizing a problem or to forming a hypothesis. In others, the recognition of a problem or question is the forerunner to observation. (See the examples in Table 1.)[10] As philosophers of science have pointed out, observation and data gathering in science is seldom undertaken without some purpose (or presuppositions) guiding what to observe (Ladyman 2002, p. 57), suggesting some reason to prefer the second formulation. On the other hand, discovery by accident seems to fit the first formulation—for example, consider Roentgen's discovery of X-rays as a result of making observations while investigating cathode rays (Kuhn 1996, p. 57). Perhaps the question—'Which comes first, observation or the recognition of a problem?'—is not a question with a single, straightforward answer.

Regardless, in both formulations, observation is portrayed as a necessary step on the road to hypothesis formation, suggesting some influence from the classical inductivist view of science in which observation was viewed as the starting point and foundation for building knowledge. But, scientists learn about research problems in their discipline in a variety of ways besides making observations themselves: from textbooks, from journal articles, from conference lectures, from discussions with other researchers, from analyzing data obtained by others, etc. Similarly, scientists derive inspiration in formulating a hypothesis from a variety of sources. To be fair, perhaps the intent of these presentations of "The Scientific Method" is not to describe an algorithm for forming a hypothesis. (As a step in a prescription for producing a science project, simply telling someone to make observations and then formulate a hypothesis provides little insight into how to do the latter.) Perhaps the aim is simply to describe how scientists stoke their creativity. As the Nobel Prize winning biologist and advocate of hypothetico-deductivism, Sir Peter Medawar, pointed out:

> [Creativity] cannot be learned perhaps, but it can certainly be encouraged and abetted. We can put ourselves in the way of having ideas, by reading and discussion and by acquiring the habit of reflection, guided by the familiar principle that we are not likely to find answers to questions not yet formulated in the mind. (Medawar 1969, p. 57)

[10] In his well known and insightful piece about methods of belief formation and the method of science, the nineteenth century philosopher Charles Sanders Peirce observed that inquiry begins with doubt: "The irritation of doubt causes a struggle to attain a state of belief. I shall term this struggle Inquiry" (Peirce 1877). Question-first accounts seem a better fit with Peirce's view, although in Peirce's view merely proposing a question is not sufficient to motivate inquiry. "There must be a real and living doubt."

Notice that Medawar mentions reading, discussion, and reflection—but not observation—in his examples of sources of intellectual creativity. The picture of the scientist who begins research on a topic by making a bunch of observations is too narrow.

4.2 No Guarantee of Objectivity

Historically, it was important for classical inductivists that science begins with observations since these were to be the objective foundation for knowledge. In order for observations to be objective, it was thought that they should be made free from bias, previous commitments, or presuppositions. In the twentieth century, this view of scientific observation came to be seen as incredibly naive. Hence, classical inductivism is sometimes referred to as *naive inductivism*. As noted previously, scientists rarely, if ever, collect data without some theoretical purpose in mind. Decisions on what to record and what not to record as well as how to classify and organize the data one records all involve prior (theoretical) presuppositions about what factors are relevant (Ladyman 2002, pp. 56–57). Observation is complex—as opposed to simple. And it is active—as opposed to being the mere passive receiving of information. Even the language scientists use to record their observations can reflect theoretical commitments. If observations are to be recorded in some language, then this already presupposes some system of categories determined by the language used and involving theoretical commitments (Feyerabend 1975, chs. 6 and 17; Harré 1985, pp. 18–34). To twentieth-century philosophers of science the idea of "pure observations" seemed elusive. They concluded that all observation—at least of the sort involved in science—is *theory-laden*.[11]

4.3 A Limited Method of Discovery

Classical inductivism has been subjected to criticism in another way as well. Inductive generalizations are limited to the terminology and ideas of the observational data collected. The inductive generalization one obtains from observing that "Swan 1 is white", "Swan 2 is white", etc., is "All swans are white". One does not obtain a hypothesis about the DNA of swans by this means. Likewise, from data about the positions of the planets at various times one does not get a hypothesis, by inductive generalization, about the forces acting on the planets. Newton took part of the empirical basis for his Three Laws of Motion and his Law of Universal Gravitation to reside in Kepler's Laws regarding the orbits of the planets. But Kepler's Laws—unlike Newton's—do not mention "forces" or "quantity of matter" (mass). Therefore, Newton's Laws cannot be induced from Kepler's because Newton's Laws contain new concepts and new terminology (Ladyman 2002, p. 55–56).

It is now typical for scientific theories to appeal to unobservable objects—atoms, DNA, electric fields, spacetime curvature, quantum states, etc.—as a means to explain and unify observational data and principles. In doing so, these theories use new concepts and new terminology. If such theories were the product of inductive generalization from observations, they would be framed strictly using the language of our observations, but they are not. The formation of these theories is not explained by the inductivist method of discovery (Blachowicz 2009, pp. 318, 322). Their formation seems to be a matter of intuition and insight rather than the application of the inductivist algorithm (Medawar 1969). Recognizing this flaw in the inductivist story, proponents of hypothetico-deductivism have been

[11] A very influential expression of the theory-ladeness of observations can be found in section X, "Revolutions as Changes of World View", in Thomas Kuhn's *The Structure of Scientific Revolutions*.

suspicious of the idea of a method of discovery. In contrast to classical inductivism, the focus of their philosophy of science has been on the logic and justification—not the invention—of scientific concepts. The latter was considered a matter of psychology.

In conclusion, taking observation as the starting point for scientific inquiry is (a) misleading as a description of the actual sources of hypothesis formation and (b) prescriptively without motivation once it is realized that observation is theory-laden. Furthermore, (c) the inductive algorithm can at best account for a limited range of hypotheses.

5 Problems with Presentations of Hypothesis Testing

5.1 Alternative Temporal Sequences

The core element in the vast majority of accounts of "The Scientific Method" is hypothesis testing. These accounts, under the influence of hypothetico-deductivism, provide a schematic description (or flowchart) of the *hypothesis testing procedure*.

1. Formulate a hypothesis H.
2. Deduce observational consequences O from H.
3. Perform a test (by observation or experimentation) to determine whether O obtains under the conditions required to deduce O.
4. Evaluate the results of the test. If O occurred, then H is said to be "confirmed". If O did not occur, then H is said to be "disconfirmed".

Although a careful description would include each of the four steps, it is not uncommon for step 2 to be overlooked and step 4 to be implicitly merged with step 3 so that the procedure is summarized as "formulate a hypothesis and test it".

When philosophers of science consider hypothetico-deductivism, their focus is primarily on the structure of the reasoning involved as opposed to an operational procedure understood either descriptively or prescriptively. That is, they are concerned with the logic of *hypothetico-deductive (HD) reasoning*.[12] Depending on the result of the test of the hypothesis, HD reasoning splits into two different forms of inference. If the observational consequence(s) O of the hypothesis H were observed to occur, then the form of inference (oversimplified as we will see shortly) is:

H deductively entails O.
O occurred.
Therefore, H is confirmed.[13]

If the observational consequence(s) O of the hypothesis H were observed NOT to occur, then the form of inference (again oversimplified) is:

H deductively entails O.

[12] For a critical, technical analysis of the hypothetico-deductive model of inductive support see Lipton 2004, pp. 15–16 and chs. 5 and 6. Lipton accuses the hypothetico-deductive model of being both too permissive and too restrictive. On one hand, according to him, some deductive consequences of a hypothesis are not evidentially relevant; and, on the other, sometimes relevant evidence is not related to a hypothesis by deductive reasoning. Lipton contrasts the hypothetico-deductive model with his theory of inference to the best explanation, arguing for the superiority of the latter with respect to the former.

[13] The terms 'confirmation' and 'disconfirmation' are preferable to 'verification' and 'falsification', although sometimes one might encounter the latter pair. The latter terms suggest an absolute verdict on the truth or falsity of the hypothesis H whereas the former terms are more susceptible to interpretations of degree.

O did not occur.
Therefore, H is disconfirmed.

The first pattern of inference may be called *HD confirmation* and the second *HD disconfirmation*.

A logical inference pattern is not the same as a temporal procedure. The order of the premises in a pattern of inference is logically irrelevant. Changing the ordering of the steps in a procedure, however, yields a different procedure. The order of operations in the statement of the hypothesis testing procedure may seem like a "logical" ordering, but reflection shows that it is not necessary. Logic does not require testing to occur only after the formation of the hypothesis in time. In other words, there is no reason why step 3 cannot occur before step 1. Philosophers of science have recognized this possibility for some time; it is presupposed in a debate over the value of "old evidence"—evidence already obtained before the hypothesis was formed. Sometimes, the debate is framed in terms of the value of prediction as opposed to explanation (or accommodation).[14] Clearly, old evidence sometimes counts as a "test" of a theory. For example, the theory of general relativity benefitted from its ability to account for the previously discovered "anomalous" precession of Mercury's perihelion (Kuhn 1996, pp. 81, 155).

5.2 Overlooked Complexities of Evaluation

There are several problems with HD reasoning as outlined above as a complete model of scientific reasoning. First, positive instances of a hypothesis may not always confirm a hypothesis. Consider the instructor grading a 100-question multiple-choice exam for a class of thirty students where achieving a score of 60 or greater is required to pass the exam. If the first five "randomly selected" exams all yield scores in the 60s, this would seem to disconfirm the hypothesis that all of the students passed the exam even though the first five scores are all positive instances of the hypothesis.[15]

Second, where statistical theories and claims are involved, special techniques are required for testing them. The claim that a particular edible substance is carcinogenic does not imply that everyone who ingests it will get cancer. So, the claim is not falsified by producing someone who routinely ate the substance but did not get cancer, nor is it confirmed simply by producing someone who did. One does not deduce (in the strict sense) a determinate observational consequence from a statistical theory. However, one can deduce probabilities; and, by comparing those with population studies, assess the theory—although this involves inductive besides deductive reasoning (Earman and Salmon 1992, p. 49). In a course on statistical research methods, this is known as *statistical hypothesis testing*.

[14] There is concern about old evidence especially when the hypothesis is tailored for the very purpose of explaining the data. As compared to the prediction of the results of a test heretofore unperformed, that should count less it would seem, if at all, toward confirming the hypothesis. However, the logic above makes no such distinction (Lipton 2004, ch. 10). When the hypothesis is formulated without regard to, or in ignorance of the data, this concern would not be apt. Such would be the case when the hypothesis is formed without knowledge of the data obtained by someone else. The latter possibility undercuts a presupposition often present in introductory accounts of hypothesis testing—namely, that the same agent is involved in every step of the procedure. Where it is natural to assume that a single agent is involved, placing the testing step after hypothesis formation may represent a way of encoding the belief that a hypothesis ought not be evaluated by data already acquired.

[15] Of course, background beliefs play a role in this example. Thanks go to Casey Swank for bringing to my attention a similar example to this one. The origin of the discussion about cases of a hypothesis that do not support the hypothesis is I. J. Good (1967).

Third, once empirical data is obtained, the evaluation of the impact of that data on a given hypothesis is not as straightforward a matter as many presentations make it seem. The process of acquiring data—either from observations or experiments—must be evaluated. One must determine what degree of confidence should be placed in the data. Was the equipment functioning properly? Were proper techniques used to acquire the data? What degree of error is acceptable given the set-up?

Finally, deriving empirical consequences from a theory often requires the assistance of other theories along with other pieces of information (for example, initial or boundary conditions). These are called *auxiliary assumptions*. Accordingly, the first premise in the two patterns of HD reasoning above should really be '(H & A) deductively entails O' where A stands for the appropriate auxiliary assumptions; likewise, in the conclusions, H should be replaced with the conjunction (H & A). When calculations were made on the basis of Newtonian mechanics about the orbit of Uranus, it was assumed that Uranus was the last of the planets in the solar system. But, in the nineteenth century, the calculated predictions about the orbit of Uranus were found not to fit the observational data. One could have taken this incident as an occasion to revise or reject Newtonian mechanics. But since Newtonian mechanics had been so successful in explaining a variety of phenomena, astronomers began looking for another planet and found Neptune. The assumption that Uranus was the last of the planets—an auxiliary assumption—turned out to be mistaken (Earman and Salmon 1992, p. 47).

The early twentieth century physicist Pierre Duhem noted that physical theories are not tested in isolation: "[T]he physicist can never subject an isolated hypothesis to experimental test, but only a whole group of hypotheses" (Duhem 1914, p. 187). The philosopher W.V. Quine generalized Duhem's insight (Quine 1951), conceiving of our whole conceptual scheme as a web of interconnected beliefs which includes some that are sharply empirical and some that are highly theoretical like our beliefs about logic and the basic constituents of reality (ontology). Quine pointed out that in fitting our conceptual scheme with the data of experience (or experiment), the more theoretical parts are connected with the more empirical parts. In accommodating some bit of recalcitrant experience (e.g., a negative test result), one could make adjustments at various places in the web of belief in order to do so. This insight into the interconnectedness of our theories and beliefs is called the *Duhem–Quine thesis*. Both Duhem and Quine recognized that when faced with a negative test result, one faces a choice about what to revise or give up—the hypothesis or one of the auxiliary assumptions used to derive consequences from the hypothesis. Quine even considered the possibility that it might be rational to revise our system of logic in order to accommodate experimental results (for example, to simplify quantum mechanics in its description of quantum phenomena). When scientists choose to revise or reject a hypothesis, it is because they express their relative confidence in the rest of the auxiliary assumptions relevant at the time as compared with the hypothesis. Logic alone does not dictate the choice.[16]

[16] The philosopher of science, Peter Kosso, claims that the main problem with the standard textbook model of scientific method is that it treats scientific theories and hypotheses in a piecemeal fashion, ignoring the large-scale structure of scientific method: "Scientific ideas and practices are essentially interconnected and networked. To understand science, one must consider not just the links between a theory and observations, but between a theory and other theories, and the influence of theories on observations" (Kosso 2009, p. 36).

5.3 Narrow Context of Evaluation

A successful prediction is accorded little weight when it is predicted by an alternative competing hypothesis. This shows that not all empirical predictions are significant tests of a hypothesis. Accounts of "The Scientific Method" often treat scientific inquiry as an activity involving the comparison of a single theory with the real world. The situation often faced by scientists, however, is one in which alternative theories are being considered. Discussing broad conceptual schemes he called *paradigms*, the philosopher of science, Thomas Kuhn, remarked "[T]he testing situation never consists…simply in the comparison of a single paradigm with nature. Instead, testing occurs as part of the competition between two rival paradigms for the allegiance of the scientific community" (Kuhn 1966, p. 145). Following Kuhn, the philosopher of science Imre Lakatos described these situations in which rival theories and experiment are placed in juxtaposition with one another as "three-cornered fights" (Lakatos 1970, p. 115). This recognition encourages viewing theory acceptance as a Darwinian affair—a survival among the fittest between alternative theories (Kuhn 1966, pp. 172–173).

What is missing in the standard portrayal of "The Scientific Method" is the big picture (Kosso 2009). The focus is routinely on a single hypothesis and its relationship with data from observation and experiment. But there are other relationships of significance for a full understanding—in particular, the interaction of the hypothesis with other accepted theories and assumptions as well as the competition between the hypothesis and alternative hypotheses.[17]

5.4 Naive Empiricism

By only mentioning observation and experiment in the evaluation of a hypothesis, hypothesis testing accounts of "The Scientific Method" treat the acceptance of scientific theories as a purely empirical matter. This reflects a *naive empiricist* view of theory choice. Other considerations besides empirical adequacy enter into the evaluation of theories; these include such factors as simplicity, internal logical consistency, external consistency (i.e., consistency with accepted theories in other domains), explanatory scope, predictive power, and fruitfulness for future research.[18] These features are called *theoretical virtues*.

What the history of science reveals is that scientists often appeal to theoretical virtues when first deciding whether a theory is pursuit-worthy and then when comparing alternative theories with one another. The logical process involved in such comparisons has been described as *inference to the best explanation*, where the best explanation is the hypothesis that is best on the basis of a total comparison (Lipton 2004). Certainly, empirical adequacy is important for convincing scientists to accept a theory, but it is not the sole criterion involved in theory selection. At times, the other virtues may be cause for scientists to continue to pursue a theory in spite of empirical difficulties it faces. Moreover, when alternative theories are compatible with the same available data (given the allowed margins of error), these other considerations become paramount. Although situations

[17] Blachowicz (2009) noted that a few of out of the 70 texts he surveyed recognized that "confirmation cannot be conclusive if alternative hypotheses are available" (p. 325). Windschitl et al. (2008) propose a view of investigative science they call "model-based inquiry" that recognizes the importance of entertaining competing hypotheses.

[18] Blachowicz (2009, p. 337) found that, in the 70 science texts he investigated, there was little awareness that non-empirical (theoretical) criteria could play a role in hypothesis evaluation.

involving such underdetermination by the evidence are sometimes resolved empirically through a so-called crucial experiment, these other dimensions of comparison play a role nonetheless.

5.5 A Restrictive Stereotype

Philosophical views of scientific methodology have historically focused upon scientific reasoning, seeking to find some single pattern of reasoning that captures how scientific knowledge is acquired. It can be dangerous, however, to promote a single form of reasoning as the essence of science. If HD reasoning is presented as the core of “The Scientific Method”, then this can suggest that inductive generalization is not an important form of scientific inference. An article in *Wired Magazine* with the title, “The End of Theory: The Data Deluge Makes the Scientific Method Obsolete”, illustrates this well. The article assumes that “The Scientific Method” requires the formation of a hypothesis before consulting data. It also assumes that the hypothesis must provide a causal-mechanical explanation of the data. As the author—the editor in chief of Wired—noted:

> The scientific method is built around testable hypotheses…But faced with massive data, this approach to science – hypothesize, model, test – is becoming obsolete… Petabytes allow us to say: “Correlation is enough.” We can stop looking for models. We can analyze the data without hypotheses about what it might show. We can throw the numbers into the biggest computing clusters the world has ever seen and let statistical algorithms find patterns where science cannot… Correlation supersedes causation, and science can advance even without coherent models, unified theories, or really any mechanistic explanation at all. (Anderson 2008)

Data-mining (using computers to extract patterns from large data sets) is portrayed as a problem for scientific methodology because there is no specific hypothesis in view beforehand. The activity itself is intended to generate a hypothesis. This process seems better described as an inductive process of discovery than an instance of hypothesis testing. Rather than being a problem for scientific methodology, what the author has really pointed out is the inadequacy of overly-restrictive portrayals of “The Scientific Method” which equate it solely with the hypothesis testing procedure inspired by HD reasoning.[19]

Scientists use various patterns of reasoning—e.g., inductive generalization, HD reasoning, and inference to the best explanation. This suggests that the correct view to take is pluralism about methods of reasoning. As the Nobel Prize winning physicist, Percy Bridgman, insightfully pointed out, scientists are opportunists:

> [T]he working scientist…feels complete freedom to utilize any method or device whatever which in the particular situation before him seems likely to yield the correct answer. In his attack on his specific problem he suffers no inhibitions of precedent or authority, but is completely free to adopt any course that his ingenuity is capable of suggesting to him. (Bridgman 1955, pp. 81–82)

5.6 An Incomplete Picture of Scientific Activity

As noted, the inspiration for the hypothesis testing procedure at the core of most statements of “The Scientific Method” is the hypothetico-deductive view of scientific reasoning. The philosophical aim had been to find some form of reasoning that is universally employed in the sciences and distinctive of how science is done. We’ve seen some reasons to conclude

[19] Campbell et al. (2006) recognize that science uses more than one approach to inquiry and reasoning. Inquiry based on inductive reasoning they describe as “Discovery Science” and inquiry based on deductive reasoning they describe as “Hypothesis-Based Science” (p. 9).

that scientific reasoning is not captured by a single simple pattern.[20] (We might also note that all the forms of reasoning considered so far are ones that occur in everyday reasoning as well and, hence, cannot be what is distinctive about science.) But a global method of reasoning is not what most scientists are concerned with when they consider the topic of methodology (a topic often discussed as a section in a research report). There are numerous local methods of great importance for particular disciplines. There are mathematical methods: certain branches of mathematics are of particular importance for one discipline as opposed to another. There are methods of measurement, e.g., random sampling. And, there are experimental methods, e.g., the use of double blind experiments. A more balanced representation of scientific activity would seem to recognize the great variety of methods that are local to various disciplines and sub-disciplines within the sciences. Again, this suggests methodological pluralism.

Thomas Kuhn is famous for having made a distinction between *normal science* and *revolutionary science* (1996). Normal science consists of research performed under the guidance of an accepted *paradigm* where a paradigm (in one sense) is a constellation of group commitments (including symbolic generalizations, model problem-solutions, metaphysical and heuristic commitments, and scientific values). A period of revolutionary science ensues when the current paradigm is increasingly questioned because of accumulating anomalies that resist resolution and bring about a sense of crisis in the respective scientific community. Whether or not one thinks this distinction has the importance that Kuhn placed upon it, if Kuhn is correct that there are periods of normal science—i.e., periods of scientific activity during which the current paradigm is accepted by the community and research is focused, not upon testing the current paradigm, but, upon applying and extending it—then the focus of "The Scientific Method" on hypothesis testing seems to miss out on normal science altogether. Kuhn describes three classes of problems (with both theoretical and experimental versions) undertaken during periods of normal science (1996, ch. 3). (1) Determination of significant fact: This involves increasing the accuracy and scope of facts that the paradigm reveals to be particularly significant, e.g., in astronomy, stellar positions and magnitudes along with the periods of eclipsing binaries and of planets. (2) Matching of facts with theory: This involves developing instruments or theoretical techniques that extend the degree to which the paradigm theory can be directly compared with the facts, e.g., the development of Atwood's machine to give the first unequivocal demonstration of Newton's second law. (3) Articulation of the paradigm theory: For example, the determination of physical constants or the reformulation of the theory in a more logically and aesthetically satisfying form. So, if Kuhn is correct about normal science, a lot of genuine scientific activity is not aimed at new discoveries nor is it aimed at testing the current paradigm. The emphasis of "The Scientific Method" upon hypothesis testing fails to capture a lot of scientific activity.

6 Specific Misconceptions about Hypotheses and Reasoning

The word 'hypothesis' has become standard in most presentations of "The Scientific Method". A hypothesis is often described as an *educated* guess aimed at being a possible answer to a question or solution to a problem. What presumably makes it an educated guess is that it is based on prior data or research. In genuine scientific contexts, one first

[20] This is not to deny that some overarching framework, like the Bayesian framework, might be able to subsume all these forms into one coherent theory of scientific reasoning.

spends a great deal of time learning about the discipline—its theories and concepts as well as its outstanding problems—in order to be in a position to formulate interesting hypotheses to address the problems in the discipline. This is why example hypotheses for a young audience [e.g., "Snails prefer a smooth path rather than a rough one" (Hicks et al. 1993, p. 5)] can seem so far removed from the sorts of hypotheses practicing scientists are concerned about. They must be generated without much prior education.

Introductory presentations of "The Scientific Method" sometimes make claims which assume that all scientific hypotheses are the same in some respect. But scientific problems and questions are of different sorts and, thus, the answers scientists produce are unlikely to all be the same in some non-trivial respect. For example, sometimes it is claimed that well-tested hypotheses become scientific laws (Hewitt 2002, p. 9). But this is problematic; it presupposes that all hypotheses take the form of a generalization—i.e., the logical form of a lawlike statement. Yet, not all scientific questions are answered by appeal to a general principle, such as a scientific law. Some scientific questions seek to know the particular causes of historic events (for example, "What brought about the extinction of the dinosaurs?"). Because scientific questions can be quite varied, answers to them may take the form of descriptions, correlations, causal statements, laws, general theoretical frameworks, initial conditions, boundary conditions, models, and so on.

6.1 "Hypotheses are possible explanations"

Sometimes hypotheses are defined as tentative explanations—"a guess about how or why something happens" (Cowens 2006) or "a possible explanation, involving only naturally occurring processes, that can explain a set of observations" (Marshak 2004).[21] But not all questions scientists ask seek an explanation. Some are simply questions about the facts or questions about relationships between variables. If that is the case, then either not all proposed answers to scientific questions (i.e., "educated guesses") are hypotheses or not all hypotheses are explanatory.

Although many introductory presentations of "The Scientific Method" claim that hypotheses are possible explanations, it is not difficult to find presentations that provide examples of hypotheses which, upon reflection, are not explanatory.

> "All objects fall to the ground when released" (Karplus 1969, p. 4).
> "Moths react positively to light" (Storer 1972, p. 6).
> "Brightly colored caterpillars are unpalatable [to birds]" (Purves et al. 1992, p. 8).
> "Elementary schoolchildren who watch more violent TV shows tend to behave more aggressively toward their peers" (Rathus 1999, p. 37).

6.2 "Hypotheses must be 'if…then…' statements"

Sometimes it is claimed that hypotheses must be worded as "if…then…" statements (Science Buddies 2013). For example, "If I add salt to fresh water, then the water will take longer to freeze" (Wysession 2007, p. 164). Presumably, the motivation for this claim about the logical structure of hypotheses is to enforce the idea that hypotheses are testable. Notice, first, that such a hypothesis does not qualify as a tentative explanation. A

[21] It is not extremely common, but one occasionally finds the bold claim made that "Science can deal only with things that can be observed" (Kramer 1987, p. 18; see also Keeton 1969, p. 1). Since contemporary science abounds in appeals to unobservables—like subatomic particles, magnetic fields, and spacetime curvature—to explain observable processes, it should be obvious that a claim like this does not sit well with the view that hypotheses are explanatory.

conditional statement like this—“if I do A, then I will observe outcome O”—connects some action with an observable consequence. As such, it is a very localized claim. But, scientific ideas are created with different degrees of generality. General conceptual schemes—for example, that air is like a fluid or that material bodies are composed of particles—are only remotely related to observational consequences.[22]

James Bryant Conant (1951) located the difference between science and common sense, insofar as method is concerned, in the difference between broad versus limited working hypotheses. According to Conant, both everyday and scientific experimentation are similar in their employment of limited working hypotheses. For example, when faced with a locked door and a bunch of keys on the floor, the everyday use of the method of trial and error experimentation can be described by a limited working hypothesis of the following sort for each key: “This key will open the door” or “If I place this key in this lock, then the door will open”. By contrast, the general conceptual scheme, or broad hypothesis, that we live in a sea of air that exerts pressure is only remotely conceptually related to the behavior of a particular piece of laboratory apparatus with a stopcock. The limited working hypothesis—“if I turn the stopcock, then such and such will happen”—is connected with the broad hypothesis only by “a highly complex process of thought and action which brings in many other concepts and conceptual schemes”; there is a “complicated chain of reasoning connecting the consequence deduced from a broad hypothesis and the actual experimental manipulation” (Conant 1951, p. 52).[23]

According to Conant, what distinguishes science from everyday experimentation—and experimentation in the practical arts—are these general conceptual schemes. In everyday experimentation, the aim is to solve a practical problem in close proximity to the limited working hypothesis. In scientific experimentation, he claimed, the aim is to test the general conceptual scheme. If Conant’s analysis of what distinguishes the scientific enterprise from everyday reasoning is correct, then the requirement that hypotheses must take the form of a fairly specific “if…then…” statement has the effect of highlighting the similarities between science and everyday problem-solving to the neglect of that which is distinctive of science: its broad conceptual schemes.

This purported requirement for hypotheses to be “if…then…” statements is related to two kinds of absences one often finds in presentations of “The Scientific Method”. First, many presentations of “The Scientific Method” omit the deduction of observational consequences as a step altogether. They proceed from the formulation of a hypothesis directly to testing it. Of course, if the hypothesis is framed in very local terms—“if I do A, then I will observe O”—there is no work to be done in deducing observational consequences; so, there is no need for a deductive step. Secondly, although presentations of

[22] Sometimes it is claimed that deductive reasoning is “if, then” reasoning. (See Blachowicz 2009, p. 317, for discussion on this.) Examples in Sect. 6.4 show that deductive reasoning does not require the use of “if…then…” statements. What such authors seem to be trying to convey is the idea that from a hypothesis H, observable consequences O should be derived. The result of this process can be summarized using a conditional statement—“if H, then O”—for example, “if organisms are composed of cells, then microscopic examination of any part of an organism should reveal cells” (Mader 2010, p. 11). In a single statement, the conditional relates H to the observational consequence O inferred from it (Saladin 2010, p. 8).

[23] Once a chain of reasoning has connected a hypothesis with an observation via some experimental manipulation, a more complex conditional can be formed that relates the hypothesis to the experimental manipulation and the observation in a single statement. For example, ‘if we live in a sea of air that exerts pressure, then if I turn the stopcock, then such and such will happen’. This has the form, ‘if H is true, then if I do A, then I will observe outcome O’. Windschitl et al. (2008) take this to be a preferable form of conditional because it includes the context of the hypothesis (or model). They, unfortunately, continue the confusing practice of calling the whole conditional a “hypothesis”.

"The Scientific Method" often include extra elaboration on observational and experimental techniques, not much is said about the difficult conceptual work involved in some scientific projects. In more mathematical sciences, theoreticians may expend much effort in deducing consequences from a general hypothesis or in using it to model some phenomenon. That is, a lot of effort can go into producing such conditionals.

6.3 "Hypotheses must be testable"

It is very common to be told that hypotheses must be testable or falsifiable (Bailey 2013; Hewitt 2002).[24] And, undoubtedly, testing is important to science. However, testability is not as simple a matter as introductory presentations suggest. Hypotheses, especially of the kind Conant identified as distinctive of the scientific enterprise, do not wear their testable credentials on their sleeve. As Duhem and Quine had noted, explanatory hypotheses—especially about unobservable entities—do not yield testable consequences apart from a context of auxiliary assumptions and other theories. So, as a point of mere logic, testability is not a property of a scientific hypothesis taken in isolation. In addition, testability as a prerequisite for serious scientific engagement places the cart before the horse. General conceptual schemes (like general relativity, quantum theory, and now string theory) when first proposed are often vague or poorly-understood and need time to be refined and explored before predictions (beyond what they were designed to explain) can be made. Hypotheses as proposed solutions to a problem need to display some virtues (like internal consistency, aesthetic appeal, and explanatory power, for example) in order to make them seem attractive enough to engage seriously, but testability need not be one of them. It can take much effort to devise suitable tests of a hypothesis. Testing is important to science, but testability as a prerequisite for initial engagement—a prerequisite for the status of "scientific hypothesis" in the first place—would be a premature requirement.[25]

6.4 "Deductive reasoning is reasoning from the general to the particular"

Sometimes introductory accounts of "The Scientific Method" expand on the logic involved. It is not uncommon to find the statement that "deductive reasoning is reasoning from the general to the particular". And, so as to give the appearance of symmetry, 'inductive reasoning' is defined as "reasoning from the particular to the general".[26] These definitions, however, represent an outdated view going back to Aristotle. Unfortunately, one still encounters them. As pointed out in Sect. 2, a *deductively valid* inference is one in which the premises logically necessitate the conclusion—that is, it is impossible for the premises to be true and the conclusion false. Note that the definition does not require the premises to be true or for someone to know them. Logic is merely concerned with the connection between the premises and the conclusion. In addition, the definition poses no restriction on the particularity or generality of the premises and the conclusion. One can reason deductively from general premises to a general conclusion:

[24] Blachowicz (2009) investigated 70 fairly recent high school and college science textbooks and found that the "great majority…stress that genuinely scientific hypotheses must be falsifiable" (p. 313).

[25] For schoolchildren producing a science fair project, however, it may be wise to require them to propose a hypothesis whose proximity to observable consequences is fairly close in order to ensure manageability.

[26] For example, one finds this in more recent treatments—Campbell et al. (2006), Cline (2013), Harris (2012)—as well as in some older texts—Keeton (1969, p. 1), Pilar (1979, pp. 4–5).

All humans are mammals.
All mammals are animals.
Therefore, all humans are animals.

One can also reason deductively from particular premises to a general conclusion:

The first of the three birds we observed today was a robin.
The second of the three birds we observed today was a robin.
The third of the three birds we observed today was a robin.
Therefore, all three birds we observed today were robins.

All that matters in deductive reasoning on the modern understanding is whether it is logically possible for the premises to be true and the conclusion false, in which case the premises do not logically necessitate the conclusion. The view that deduction proceeds from the general to the particular comes from Aristotle's view of scientific explanation. To explain a particular fact is to "deduce" it from a general scientific principle (Losee 1993, ch. 1).

Unlike the deductive case, there is some basis today for describing induction as reasoning from the particular to the general. This description could be understood as a reference to the specific form of inductive reasoning (in the modern sense) which is often called *inductive generalization*. From observing some, but not all, members of a class of objects and that they each share some property P, one infers that all members of the class have P. Since the inference extrapolates from observed cases to a conclusion that includes unobserved ones, the conclusion of the inference could be mistaken. So, the inference is deductively invalid. The premises of such an inference supply evidence for the conclusion and, at best, make the conclusion probable. In the modern understanding, inductive reasoning in the broad sense is just probabilistic reasoning, where the inferences of interest are ones that under the scrutiny of deductive reasoning would be declared invalid. Inductive logic is concerned with techniques for evaluating the inductive strength of such inferences. Nothing stands in the way then of reasoning inductively from a general premise to a particular conclusion:

All birds observed today have been robins.
Therefore, the next bird we observe today will be a robin.

Logic as it is understood today, whether inductive or deductive, is concerned with the kind of evidential relationship that exists between the premises and the conclusion. Whether they are particular or general is incidental (Hurley 1991, pp. 34–35).

7 "The Scientific Method" as Myth and Ideal

Some myths, once recognized as such, deserve to be discarded. But this was not Bauer's conclusion in his book *Scientific Literacy and the Myth of the Scientific Method*. Bauer thought that the myth promotes certain ideals which are important for science to function well—for example, objectivity and systematic rationality. Instead of discarding the myth, Bauer advocated understanding and treating the so-called method as an unattainable ideal. The myth should be qualified, according to Bauer, so that people understand that it is an unattainable ideal and not a description of reality. Why? "Those who hold ideals, no matter that they are unattainable, are likely to behave more in accord with them than will people who do not hold those ideals…Myths, after all, even if not literally true, are stories that embody moral truths" (p. 39).

Bauer's conclusion that we should qualify the myth raises a problem. Can someone qualify the myth in the way that Bauer suggested and still hold it? The qualification that Bauer advocated would require a person to no longer believe that the scientific method is an accurate description of what scientists do and how science works. But that belief is the essence of the myth itself. In claiming that the scientific method is an unattainable ideal, was Bauer claiming that scientists do not formulate hypotheses and test them? No. And, surely, the misconceptions about the scientific method that he corrected—that data always trumps theory, that it is mechanical, and that it is individualistic—are not aspects of the myth that should be considered as ideals. What Bauer took to be an unattainable ideal is not the method in the myth but the objectivity and systematic rationality associated with the method. What Bauer *should have concluded* is that we discard the myth of "The Scientific Method" while continuing to promote the ideals that are associated with it. The way to promote these ideals is not to root them in a false story about "The Scientific Method" but—as Bauer in fact did in his book—to make clear the importance of objectivity and rationality in science as well as the factors that contribute to the realization of these ideals as science is actually performed.

For example, the popular conception of science locates the objectivity of science either in the individuals involved or in following an individualistic, mechanical method. Bauer, however, located the source of science's objectivity in the social dynamics of the scientific community, illustrated by the puzzle and filter analogies (see Sect. 3). "Objectivity comes into science because ideas and results are exposed to the criticism of people with disparate and conflicting and competing intellectual approaches and beliefs, personal biases, social goals, hidden agendas, and the rest" (p. 102). This is objectivity obtained through the process of achieving intersubjective consensus. It is an ideal because social structure can be more or less conducive to this goal.

In another thread in his book, Bauer recognized a further ideal that the myth communicates about science besides the ideals of objectivity and rationality. This occurs in his discussion of what he called "reality therapy" (a term he attributed to Richard Burian). He used the term to describe the way in which scientific consensus is sensitive to the test of nature. Although the myth would have it that empirical evidence decisively determines theory,

> [N]ature does still constrain observation and experiment and thereby also interpretation (or theory, or scientific belief). It does so less directly, less precisely, less automatically, and less quickly than is envisaged in the classical formulations of the scientific method; nevertheless, nature cannot but remain the ultimate and entirely firm arbiter. (p. 89)

The term 'therapy' is meant to capture the idea that "reality does not determine belief all at once" and that it is often a "lengthy, slow, nonautomatic process" (pp. 89–90). The role that reality (via empirical evidence) plays in bringing about consensus adds a further dimension of objectivity to science.

Although misleading in many ways, the myth of "The Scientific Method" communicates a very important ideal of science. It communicates an epistemic stance, the importance of which Bauer recognized with his discussion of reality therapy and that one might call, using other terms, *the empirical stance*. As opposed to other epistemic stances which would attempt to influence belief by appeal to authority, pure reason, or revelation, the empirical stance focuses on the importance of empirical evidence—even if it is not the whole story—for belief formation and consensus building in the scientific community. The myth functions as a way of communicating that stance, but it does so by communicating much more besides. The ideals of objectivity and rationality along with the empirical stance can, and should, be communicated without all the baggage the myth brings with it. This is what Bauer should have concluded.

8 Conclusion

We have seen various respects in which "The Scientific Method" in its popular and introductory presentations is inadequate and misleading as a description of scientific activity. These reasons not only justify calling it a "myth", with the common negative evaluation associated with that term, but they also justify the conclusion that it ought to be discarded. Sometimes pictures that are strictly incorrect or incomplete can be useful as first approximations on the road to understanding. This is not the case with "The Scientific Method"; it misleads. We do better by avoiding the myth altogether. Likewise, it should not be considered an unattainable ideal.

The lack of expert consensus on an ultimate overarching perspective on scientific methodology (Sect. 2) is a problem for pedagogy at the introductory level. One cannot simply look to the experts for a single unifying theme. This does not imply that nothing can be responsibly conveyed about scientific methodology, but it does suggest avoiding the impression that any one idea is the core to understanding all scientific activity. What lessons can be drawn for introductory pedagogical contexts from the preceding meditations?

8.1 Scientific Ideas and Their Formation

Scientific ideas take the form of generalizations, equations, correlations, causal statements, initial conditions, models, etc. They can concern observable as well as unobservable entities and processes. So, it is a dangerous business to place restrictions on scientific hypotheses—e.g., that they should be explanatory, falsifiable, or 'if…then…' statements, etc. (Sect. 6) Moreover, the inductivist view of the discovery (or invention) of hypotheses was too narrow as a method for the formation of the whole range of scientific ideas (Sect. 4). Some generalizations and correlations may be the result of noticing and generalizing patterns in observational data, but hypotheses about causes and unobservables are often not of that sort. Intellectual creativity is involved and finds inspiration in reading, reflection, and discussion as much as in making observations. The appeal to creativity is not a denial that there are techniques of discovery for such hypotheses—e.g., the use of analogy or corrective trial-and-error techniques (Blachowicz 2009)—but, when and how to apply these techniques may require insight and expertise.

8.2 Scientific Methods

Scientific methods can be discussed without the use of the word 'the' to denote a single method. Some have suggested that a useful pedagogical metaphor for discussing scientific methodology at the introductory level would be *the scientists' toolbox* (Wivagg and Allchin 2002). Since toolboxes often contain several tools for different purposes, this has the advantage of suggesting a plurality of methods tailored to particular aims. Besides, it is more fitting to how scientists themselves view methodology to think of methods as tools wielded by scientists in pursuit of their goals and interests rather than conceiving of a universal method that utilizes the scientist as a cog in its production process.

There are methods of various kinds. There are methods, or forms, of scientific reasoning—e.g., inductive generalization, hypothetico-deductive reasoning, and inference to the best explanation. Ultimately, as we have seen, scientific reasoning is complex and its complexity is due to the global structure of scientific ideas; scientific ideas are connected to each other to form an interlocking network or web. Besides methods of reasoning, there are also mathematical methods, computational methods, measurement methods, and experimental methods (Sect. 5).

8.3 Scientific Aims, Values, and Virtues

Scientific activity is not simply aimed at testing ideas, but also articulating and applying them (Sect. 5.6). Commitments to objectivity, rationality, and empirical evidence (the empirical stance) shape these activities (Sect. 7). The acceptance of scientific ideas is also shaped by values called "theoretical virtues". Besides empirical adequacy, these include such virtues as simplicity, elegance, predictive power, explanatory scope, coherence with other accepted theories, etc. (Sect. 5.4)

8.4 Scientific Communities

Scientists often work with others, co-authoring articles, performing experiments, discussing results, etc. They often specialize. Some disciplines divide into experimentalists and theorists. Scientists differ in their skills as well as their personalities. There is not a universal job description that outlines what each scientist does, even though each scientist contributes to the total enterprise in his or her discipline. One's contributions are subject to the review and criticism of one's peers. This has the effect of filtering out ideas and results that are not acceptable to the community as a whole—for example, ideas that do not seem to be suitably objective, rational, and well-supported by evidence (Sects. 3, 7).

Acknowledgments Special gratitude goes to the students in the various iterations of my course, Philosophy of Science, along with audience members at my talks at the 2012 and 2013 meetings of the Minnesota Philosophical Society, the 2013 meeting of the Wisconsin Philosophical Association, and the 2013 Biennial IHPST Conference. Support has been provided from the University of Wisconsin-Eau Claire Academic Affairs Professional Development Program.

References

Anderson, C. (2008). The end of theory: The data deluge makes the scientific method obsolete. *Wired Magazine*, 16(7). http://www.wired.com/science/discoveries/magazine/16-07/pb_theory.

Bailey, R. (2013). About.com: Biology: Scientific method. http://biology.about.com/od/biologysciencefair/p/sciencemethod.htm.

Bauer, H. H. (1992). *Scientific literacy and the myth of the scientific method*. Champaign, IL: University of Illinois Press.

Blachowicz, J. (2009). How science textbooks treat scientific method: A philosopher's perspective. *British Journal for the Philosophy of Science, 60*, 303–344.

Bridgman, P. W. (1955). On scientific method. In *Reflections of a Physicist* (2nd Edn.), New York: Philosophical Library.

Campbell, N. A., Reece, J. B., Taylor, M. R., & Simon, E. J. (2006). *Biology: Concepts and connections* (5th ed.). San Francisco, CA: Pearson Education, Inc.

Caws, P. (1967). Scientific method. In P. Edwards (Ed.), *The encyclopedia of philosophy* (Vol. 7, pp. 339–343). New York: Macmillan Publishing Co., Inc.

Cline, A. (2013). About.com: Agnosticism/atheism: What is the scientific method? http://atheism.about.com/od/philosophyofscience/a/ScientificMethod.htm.

Conant, J. B. (1951). *Science and common sense*. New Haven: Yale University Press.

Cowens, J. (2006). The scientific method. *Teaching Pre K-8, 37*(1), 42–46.

Duhem, P. (1954 [1914]). *The Aim and Structure of Physical Theory*. Princeton, NJ: Princeton University Press.

Earman, J., & Salmon, W. C. (1992). The confirmation of scientific hypotheses. In M. H. Salmon, et al. (Eds.), *Introduction to the philosophy of science*. Indianapolis: Hackett Publishing Company.

Edmund, N. W. (2011). The scientific method today. http://scientificmethod.com/index2.html.

Feyerabend, P. (1975). *Against method: Outline of an anarchistic theory of knowledge*. London: Verso Editions.

Good, I. J. (1967). The white shoe is a red herring. *The British Journal for the Philosophy of Science, 17*(4), 322.

Harré, R. (1985). *The philosophies of science* (2nd ed.). Oxford: Oxford University Press.

Harris, W. (2012). Howstuffworks: How the scientific method works. http://science.howstuffworks.com/innovation/scientific-experiments/scientific-method.htm.

Hewitt, P. G. (2002). *Conceptual physics* (9th ed.). San Francisco: Addison Wesley.

Hicks, L., Shimmin, D., & Rickard, G. (1993). *Science: Order and reality* (2nd ed.). Pensacola: Pensacola Christian College.

Howson, C., & Urbach, P. (2005). *Scientific reasoning: The Bayesian approach* (3rd ed.). Chicago, IL: Open Court.

Hurley, P. J. (1991). *A concise introduction to logic* (4th ed.). Belmont, CA: Wadsworth Publishing Company.

Karplus, R. (1969). *Introductory physics: A model approach*. New York: W. A. Benjamin Inc.

Keeton, W. T. (1969). *Elements of biological science*. New York: Norton.

Kosso, P. (2009). The large-scale structure of scientific method. *Science and Education, 18*, 33–42.

Kramer, S. P. (1987). *How to think like a scientist: Answering questions by the scientific method*. HarperCollins Publishers.

Kuhn, T. S. (1996 [1962]). *The structure of scientific revolutions* (3rd Edn). Chicago: University of Chicago Press.

Ladyman, J. (2002). *Understanding philosophy of science*. New York: Routledge.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In Lakatos, I., & Musgrave, A. (Eds.), *Criticism and the growth of knowledge*. Proceedings of the International Colloquium in the Philosophy of Science, London, 1965, Vol. 4 (pp. 91–195). Cambridge University Press.

Lipton, P. (2004). *Inference to the best explanation* (2nd ed.). New York: Routledge.

Losee, J. (1993). *A historical introduction to the philosophy of science* (3rd ed.). Oxford: Oxford University Press.

Mader, S. S. (2010). *Biology* (10th ed.). New York: McGraw-Hill.

Marshak, S. (2004). *Essentials of geology*. New York: W. W. Norton & Company.

Medawar, P. B. (1969). *Induction and intuition in scientific thought*. Philadelphia: American Philosophical Society.

Mill, J. S. (1881). *A system of logic* (8th Edn.). In E. Nagel (Ed.) *John Stuart mill's philosophy of scientific method*. New York: Hafner Publishing Co. (1950).

Peirce, C. S. (1877). The fixation of belief. *Popular Science Monthly, 12*, 1–15.

Pilar, F. L. (1979). *Chemistry—The universal science*. Reading, MA: Addison-Wesley Publishing.

Purves, W. K., Orians, G. H., & Heller, H. C. (1992). *Life, the science of biology* (3rd ed.). Sunderland, MA: Sinauer Associates Inc.

Quine, W. V. (1951). Two dogmas of empiricism. *The Philosophical Review, 60*(1), 20–43.

Rathus, S. A. (1999). *Psychology in the new millennium*. Orlando, FL: Harcourt Brace College Publishers.

Saladin, K. S. (2010). *Anatomy and physiology: The unity of form and function* (5th ed.). Boston, MA: McGraw-Hill.

Science Buddies. (2013). Science buddies: Steps of the scientific method. http://www.sciencebuddies.org/science-fair-projects/project_guide_index.shtml.

Storer, T. I. (1972). *General zoology* (5th ed.). New York: McGraw-Hill.

Whewell, W. (1847). *The philosophy of the inductive sciences founded upon their history* (Vol. 2). London: John W. Parker.

Windschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education, 92*(5), 941–967.

Wivagg, D., & Allchin, D. (2002). The dogma of "the" scientific method. *The American Biology Teacher, 64*(9), 645–646.

Wysession, M. (2007). *Prentice Hall Science explorer: Earth's changing surface*. Boston, MA: Pearson Education Inc.