# Stanford Encyclopedia of Philosophy Scientific Method

#### First published Fri Nov 13, 2015

Science is an enormously successful human enterprise. The study of scientific method is the attempt to discern the activities by which that success is achieved. Among the activities often identified as characteristic of science are systematic observation and experimentation, inductive and deductive reasoning, and the formation and testing of hypotheses and theories. How these are carried out in detail can vary greatly, but characteristics like these have been looked to as a way of demarcating scientific activity from non-science, where only enterprises which employ some canonical form of scientific method or methods should be considered science (see also the entry on science and pseudo-science). On the other hand, more recent debate has questioned whether there is anything like a fixed toolkit of methods which is common across science and only science.

Scientific method should be distinguished from the aims and products of science, such as knowledge, predictions, or control. Methods are the means by which those goals are achieved. Scientific method should also be distinguished from meta-methodology, which includes the values and justifications behind a particular characterization of scientific method (i.e., a methodology) — values such as objectivity, reproducibility, simplicity, or past successes. Methodological rules are proposed to govern method and it is a meta-methodological question whether methods obeying those rules satisfy given values. Finally, method is distinct, to some degree, from the detailed and contextual practices through which methods are implemented. The latter might range over: specific laboratory techniques; mathematical formalisms or other specialized languages used in descriptions and reasoning; technological or other material means; ways of communicating and sharing results, whether with other scientists or with the public at large; or the conventions, habits, enforced customs, and institutional controls over how and what science is carried out.

While it is important to recognize these distinctions, their boundaries are fuzzy. Hence, accounts of method cannot be entirely divorced from their methodological and meta-methodological motivations or justifications, Moreover, each aspect plays a crucial role in identifying methods. Disputes about method have therefore played out at the detail, rule, and meta-rule levels. Changes in beliefs about the certainty or fallibility of scientific knowledge, for instance (which is a meta-methodological consideration of what we can hope for methods to deliver), have meant different emphases on deductive and inductive reasoning, or on the relative importance attached to reasoning over observation (i.e., differences over particular methods.) Beliefs about the role of science in society will affect the place one gives to values in scientific method.

The issue which has shaped debates over scientific method the most in the last half century is the question of how pluralist do we need to be about method? Unificationists continue to hold out for one method essential to science; nihilism is a form of radical pluralism, which considers the effectiveness of any methodological prescription to be so context sensitive as to render it not explanatory on its own. Some middle degree of pluralism regarding the methods embodied in scientific practice seems appropriate. But the details of scientific practice vary with time and place, from institution to institution, across scientists and their subjects of investigation. How significant are the variations for understanding science and its success? How much can method be abstracted from practice? This entry describes some of the attempts to characterize scientific method or methods, as well as arguments for a more context-sensitive approach to methods embedded in actual scientific practices.

- <u>1. Overview and organizing themes</u>
- 2. Historical Review: Aristotle to Mill
- 3. Logic of method and critical responses
  - 3.1 Logical constructionism and Operationalism
  - 3.2. H-D as a logic of confirmation
  - 3.3. Popper and falsificationism
  - 3.4 Meta-methodology and the end of method
- 4. Statistical methods for hypothesis testing
- 5. Method in Practice
  - 5.1 Creative and exploratory practices
  - 5.2 Computer methods and the 'third way' of doing science
- <u>6. Discourse on scientific method</u>
  - 6.1 "The scientific method" in science education and as seen by scientists
  - <u>6.2 Privileged methods and 'gold standards'</u>
  - <u>6.3 Scientific method in the court room</u>
  - 6.4 Deviating practices
- 7. Conclusion
- Bibliography
- <u>Academic Tools</u>
- Other Internet Resources
- <u>Related Entries</u>

# 1. Overview and organizing themes

This entry could have been given the title Scientific Methods and gone on to fill volumes, or it could have been extremely short, consisting of a brief summary rejection of the idea that there is any such thing as a unique Scientific Method at all. Both unhappy prospects are due to the fact that scientific activity varies so much across disciplines, times, places, and scientists that any account which manages to unify it all will either consist of overwhelming descriptive detail, or trivial generalizations.

The choice of scope for the present entry is more optimistic, taking a cue from the recent movement in philosophy of science toward a greater attention to practice: to what scientists actually do. This "turn to practice" can be seen as the latest form of studies of methods in science, insofar as it represents an attempt at understanding scientific activity, but through accounts that are neither meant to be universal and unified, nor singular and narrowly descriptive. To some extent, different scientists at different times and places can be said to be using the same method even though, in practice, the details are different.

Whether the context in which methods are carried out will be at all relevant, or to what extent it will be so, will depend largely on what one takes the aims of science to be and what one's own aims are. For most of the history of scientific methodology the assumption has been that the most important output of science is knowledge and so the aim of methodology should be to discover those methods by which scientific knowledge is generated.

Science was seen to embody the most successful form of reasoning (but which form?) to the most certain knowledge claims (but how certain?) on the basis of systematically collected evidence (but what counts as evidence and, in particular, should the evidence of the senses or rather of rational insight take precedence?) <u>Section 2</u> surveys some of the history, pointing to two major themes. One theme is seeking the right balance between observation and reasoning (and the attendant forms of reasoning which employ them); the other is how certain scientific knowledge is or can be.

Section 3 turns to 20<sup>th</sup> century debates on scientific method. In the second half of the 20<sup>th</sup> century the epistemic privilege of science faced several challenges and many philosophers of science abandoned the reconstruction of the logic of scientific method. Views changed significantly regarding which functions of science ought to be captured and why. For some, the success of science was better identified with social or cultural features. Historical and sociological turns in the philosophy of science were made, with a demand that greater attention be paid to the non-epistemic aspects of science, such as sociological, institutional, material, and political factors. Even outside of those movements there was an increased specialization in the philosophy of science. Sections 3 and 4 surveys the main positions on scientific method in 20<sup>th</sup> century philosophy of science, focusing on where they differ in their preference for confirmation or falsification or for waiving the idea of a special scientific method altogether.

In recent decades, attention has primarily been paid to scientific activities traditionally falling under the rubric of method, such as experimental design and general laboratory practice, the use of statistics, the construction and use of models and diagrams, interdisciplinary collaboration, and science communication. Sections 4–6 attempt to construct a map of the current domains of the study of methods in science.

As these sections illustrate, the question of method is still central to the discourse about science. Scientific method remains a topic for education, for science policy, and among scientists. It arises in the public domain where the demarcation of science is at issue. Some philosophers have recently returned, therefore, to the question of what it is that makes science a unique cultural product. This entry will close with some of these recent attempts at discerning and encapsulating the activities by which scientific knowledge is achieved.

# 2. Historical Review: Aristotle to Mill

Attempting a history of scientific method compounds the vast scope of the topic. This section briefly surveys the background to modern methodological debates. What can be called the classical view goes back to antiquity, and represents a point of departure for later divergences.<sup>[1]</sup>

We begin with a point made by Laudan (1968) in his historical survey of scientific method:

Perhaps the most serious inhibition to the emergence of the history of theories of scientific method as a respectable area of study has been the tendency to conflate it with the general history of epistemology, thereby assuming that the narrative categories and classificatory pigeon-holes applied to the latter are also basic to the former. (1968: 5)

To see knowledge about the natural world as falling under knowledge more generally is an understandable conflation. Histories of theories of method would naturally employ the same narrative categories and classificatory pigeon holes. An important theme of the history of epistemology, for example, is the unification of knowledge, a theme reflected in the question of the unification of method in science. Those who have identified differences in kinds of knowledge have often likewise identified different methods for achieving that kind of knowledge (see the entry on the <u>unity of science</u>).

Related to the diversities of what is known, and how, are differences over what can be known. Plato (429–347 B.C.E) distinguished the realms of things into the visible and the intelligible. Only the latter, the Forms, could be objects of knowledge. The intelligible truths could be known with the certainty of geometry and deductive reasoning. What could be observed of the material world, however, was by definition imperfect and deceptive, not ideal. The Platonic way of knowledge therefore emphasized reasoning as a method, downplaying the importance of observation. Aristotle (384–322 B.C.E) disagreed, locating the Forms in the natural world as the fundamental principles to be discovered through the inquiry into nature.

Aristotle is recognized as giving the earliest systematic treatise on the nature of scientific inquiry in the western tradition, one which embraced observation and reasoning about the natural world. In the *Prior* and *Posterior Analytics*, Aristotle reflects first on the aims and then the methods of inquiry into nature. A number of features can be found which are still considered by most to be essential to science. For Aristotle, empiricism, careful observation (but passive observation, not controlled experiment), is the starting point, though the aim is not merely recording of facts. Science (*epistêmê*), for Aristotle, is a body of properly arranged knowledge or learning—the empirical facts, but also their ordering and display are of crucial importance. The aims of discovery, ordering, and display of facts partly determine the methods required of successful scientific inquiry. Also determinant is the nature of the knowledge being sought, and the explanatory causes proper to that kind of knowledge (see the discussion of the four causes in the entry on <u>Aristotle on causality</u>).

In addition to careful observation, then, scientific method requires a logic as a system of reasoning for properly arranging, but also inferring beyond, what is known by observation. Methods of reasoning may include induction, prediction, or analogy, among others. Aristotle's system (along with his catalogue of fallacious reasoning) was collected under the title the *Organon*. This title would be echoed in later works on scientific reasoning, such as *Novum Organon* by Francis Bacon, and *Novum Organon Restorum* by William Whewell (see below). In the *Organon* reasoning is divided primarily into two forms, a rough division which persists into modern times. The division, known most commonly today as deductive versus inductive method, appears in other eras and methodologies as analysis/synthesis, non-ampliative/ ampliative, or even confirmation/verification. The basic idea is that there are two "directions" to proceed in our methods of inquiry: one away from what is observed, to the more fundamental, general, and encompassing principles; the other, leads from the fundamental and general to other possible specific instantiations of those principles.

The basic aim and method of inquiry identified here can be seen as a theme running throughout the next two millennia of reflection on the correct way to seek after knowledge: carefully observe nature and then seek rules or principles which explain or predict its operation. The Aristotelian corpus provided the framework for a commentary tradition on scientific method independent of the science itself (its physics and cosmos.) During the medieval period, figures such as Albertus Magnus (1206–1280), Thomas Aquinas (1225–1274), Robert Grosseteste (1175–1253), Roger Bacon (1214/1220–1292), William of Ockham (1287–1347), Andreas Vesalius (1514–1546), Giacomo Zabarella (1533–1589) all worked to clarify the kind of knowledge which could be obtained by observation and induction, the source of justification of induction, and the best rules for its application.<sup>[2]</sup> Many of their contributions we now think of as essential to science (see also Laudan 1968). As Aristotle and Plato had employed a framework of reasoning either "to the forms" or "away from the forms", medieval thinkers employed directions away from the phenomena or back to the phenomena. In analysis, a phenomena was examined to discover its basic explanatory principles; in synthesis, explanations of a phenomena were constructed from first principles.

During the Scientific Revolution these various strands of argument, experiment, and reason were forged into a dominant epistemic authority. The 16<sup>th</sup>–18<sup>th</sup> centuries were a period of not only dramatic advance in knowledge about the operation of the natural world—advances in mechanical, medical, biological, political, economic explanations—but also of self-awareness of the revolutionary changes taking place, and intense reflection on the source and legitimation of the method by which the advances were made. The struggle to establish the new authority included methodological moves. The Book of Nature, according to the metaphor of Galileo Galilei (1564–1642) or Francis Bacon (1561–1626), was written in the language of mathematics, of geometry and number. This motivated an emphasis on mathematical description and mechanical explanation as important aspects of scientific method. Through figures such as Henry More and Ralph Cudworth, a neo-Platonic emphasis on the importance of metaphysical reflection on nature behind appearances, particularly regarding the spiritual as a complement to the purely mechanical, remained an important methodological thread of the Scientific Revolution (see the entries on Cambridge platonists; Boyle; Henry More; Galileo).

In *Novum Organum* (1620), Bacon was critical of the Aristotelian method for proceeding too quickly and leaping from particulars to universals, largely as dictated by the syllogistic form of reasoning which regularly mixed those two types of propositions. Bacon aimed at the invention of new arts, of principles, of designations and directions for works. His method would be grounded in methodical collection of data and observations, coupled with correction of our senses (and particularly, strictures for the avoidance of the Idols, as he called them, kinds of systematic errors to which naïve observers are prone.) The community of scientists could then climb, by a careful, gradual and unbroken ascent, to reliable general claims.

Bacon's method has been criticized as impractical and too inflexible for any living, practicing scientist. Whewell would later criticize Bacon in his *System of Logic* for paying too little attention to the practices of scientists. It is hard to find convincing examples of Bacon's method being put in to practice in the history of science, but there are a few who have been held up as real examples of 16<sup>th</sup> century scientific, inductive method, even if not in the rigid Baconian mold: figures such as Robert Boyle (1627–1691) and William Harvey (1578–1657) (see the entry on Bacon).

It is to Isaac Newton (1642–1727), however, that historians of science and methodologists have paid the greatest attention, by far. Given the enormous success of his *Principia Mathematica* and *Opticks*, this is understandable. The study of Newton's method has had two main thrusts: the implicit method of the experiments and reasoning presented in the Opticks, and the explicit methodological rules given as the Rules for Philosophising (the Regulae) in Book III of the *Principia*.<sup>[3]</sup> Newton's law of gravitation, the linchpin of his new cosmology, broke with explanatory conventions of natural philosophy, first for apparently proposing action at a distance, but more generally for not providing "true", physical causes. The argument for his System of the World (*Principia*, Book III) was based on phenomena, not reasoned first principles. This was viewed (mainly on the continent) as insufficient for proper natural philosophy. The Regulae counter this objection, re-defining the aims of natural philosophy by re-defining the method natural philosophers should follow.

- Rule I: No more causes of natural things should be admitted than are both true and sufficient to explain their phenomena.
- Rule II: Therefore, the causes assigned to natural effects of the same kind must be, so far as possible, the same.
- Rule III: Those qualities of bodies that cannot be intended and remitted and that belong to all bodies on which experiments can be made should be taken as qualities of all bodies universally.
- Rule IV: In experimental philosophy, propositions gathered from phenomena by induction should be considered either exactly or very nearly true notwithstanding any contrary hypotheses, until yet other phenomena make such propositions either more exact or liable to exceptions.

To his list of methodological prescriptions should be added Newton's famous phrase "*hypotheses non fingo*" (commonly translated as "I frame no hypotheses".) The scientist was not to invent systems but infer explanations from observations, as Bacon had advocated. This would come to be known as inductivism. In the century after Newton, significant clarifications of the Newtonian method were made. Colin Maclaurin (1698–1746), for instance, reconstructed the essential structure of the method as having complementary analysis and synthesis phases, one proceeding away from the phenomena in generalization, the other from the general propositions to derive explanations of new phenomena. Denis Diderot (1713–1784) and editors of the *Encyclopédie* did much to consolidate and popularize Newtonianism, as did Francesco Algarotti (1721–1764). The emphasis was often the same, as much on the character of the scientist as on their process, a character which is still commonly assumed. The scientist is humble in the face of nature, not beholden to dogma, obeys only his eyes, and follows the truth wherever it leads. It was certainly Voltaire (1694–1778) and du Chatelet (1706–1749) who were most influential in propagating the latter vision of the scientist and their craft, with Newton as hero. Scientific method became a revolutionary force of the Enlightenment. (See also the entries on <u>Newton, Leibniz, Descartes, Boyle, Hume, enlightenment</u>, as well as Shank 2008 for a historical overview.)

Not all 18<sup>th</sup> century reflections on scientific method were so celebratory. Famous also are George Berkeley's (1685–1753) attack on the mathematics of the new science, as well as the over-emphasis of Newtonians on observation; and David Hume's (1711–1776) undermining of the warrant offered for scientific claims by inductive justification (see the entries on: <u>George Berkeley; David Hume; Hume's Newtonianism and Anti-Newtonianism</u>). Hume's problem of induction motivated Immanuel Kant (1724–1804) to seek new foundations for empirical method, though as an epistemic reconstruction, not as any set of practical guidelines for scientists. Both Hume and Kant influenced the methodological reflections of the next century, such as the debate between Mill and Whewell over the certainty of inductive inferences in science.

The debate between John Stuart Mill (1806–1873) and William Whewell (1794–1866) has become the canonical methodological debate of the 19<sup>th</sup> century. Although often characterized as a debate between inductivism and hypothetico-deductivism, the role of the two methods on each side is actually more complex. On the hypothetico-deductive account, scientists work to come up with hypotheses from which true observational consequences can be deduced—hence, hypothetico-deductive. Because Whewell emphasizes both hypotheses and deduction in his account of method, he can be seen as a convenient foil to the inductivism of Mill. However, equally if not more important to Whewell's portrayal of scientific method is what he calls the "fundamental antithesis". Knowledge is a product of the objective (what we see in the world around us) and subjective (the contributions of our mind to how we perceive and understand what we experience, which he called the Fundamental Ideas). Both elements are essential according to Whewell, and he was therefore critical of Kant for too much focus on the subjective. An idea can be fundamental even if it is necessary for knowledge only within a given scientific discipline (e.g., chemical affinity for chemistry). (This distinguishes fundamental ideas from the forms and categories of intuition of Kant. See <u>Whewell</u> entry.)

Clarifying fundamental ideas is therefore an essential part of scientific method and scientific progress. Whewell called this process "Discoverer's Induction". It was induction, following Bacon or Newton, but Whewell sought to revive Bacon's account by emphasising the role of ideas in the clear and careful formulation of inductive hypotheses. Whewell's induction is not merely the collecting of objective facts. The subjective plays a role through what Whewell calls the Colligation of Facts, a creative act of the scientist, the invention of a theory. A theory is then confirmed by testing, where more facts are brought under the theory, called the Consilience of Inductions. Whewell felt that this was the method by which the true laws of nature could be discovered: clarification of fundamental concepts, clever invention of explanations, and careful testing. Mill, in his critique of Whewell, and others who have cast Whewell as a fore-runner of the hypothetico-deductivist view, seem to have under-estimated the importance of this discovery phase in Whewell's understanding of method (Snyder 1997a,b, 1999). Down-playing the discovery phase would come to characterize methodology of the early 20<sup>th</sup> century (see section 3).

Mill, in his *System of Logic*, puts forward instead a narrower view of induction as the essence of scientific method. For Mill, induction is the search first for regularities among events. Among those regularities, some will continue to hold for further observations, eventually gaining the status of laws. One can also look for

regularities among the laws discovered in one domain, i.e., for a law of laws. Which "law law" will hold is time and discipline dependent and should be held open to revision. One example is the Law of Universal Causation, and Mill put forward specific methods for identifying causes—now commonly known as Mill's methods. These five methods look for circumstances which are common among the phenomena of interest, those which are absent when the phenomena are, or those for which both vary together. Mill's methods are still seen as capturing basic intuitions about experimental methods for finding the relevant explanatory factors (*System of Logic* (1843), see <u>Mill</u> entry). The methods advocated by Whewell and Mill, in the end, look similar. Both involve induction and generalization to covering laws. They differ dramatically, however, with respect to the necessity of the knowledge arrived at; that is, at the meta-methodological level (see the entries on <u>Whewell</u> and <u>Mill</u> entries).

# 3. Logic of method and critical responses

The quantum and relativistic revolutions in physics in the early 20<sup>th</sup> century had a profound effect on methodology. The conceptual foundations of both of these physical theories were taken to show the defeasibility of even the most seemingly secure commonsense intuitions about space, time and physical bodies. Certainty of knowledge about the natural world was therefore recognized as unattainable, and instead a renewed empiricism was sought, which rendered science fallible but at the same time rationally justified.

In support of this, analysis of the reasoning of scientists emerged according to which the aspects of scientific method which were of primary importance were the means of testing and confirming of theories. A distinction in methodology was made between the contexts of discovery and of justification. The distinction could be used as a wedge between, on the one hand the particularities of where and how theories or hypotheses are arrived at and, on the other, the underlying reasoning scientists use (whether or not they are aware of it) when assessing theories and judging their adequacy on the basis of the available evidence. By and large, for most of the 20<sup>th</sup> century, philosophy of science focused on the second context, although philosophers differed on whether to focus on confirmation or refutation as well as on the many details of how confirmation or refutation could or could not be brought about. By the mid-20<sup>th</sup> century these attempts at defining the method of justification and the context distinction itself came under pressure. During the same period, philosophy of science developed rapidly, and from section 4 this entry will therefore shift from a primarily historical treatment of the scientific method towards a primarily thematic one.

#### 3.1 Logical constructionism and Operationalism

Advances in logic and probability held out promise of the possibility of elaborate reconstructions of scientific theories and empirical methods. The best example of this is Rudolf Carnap's *The Logical Structure of the World* (1928) Here, Carnap attempted to show that a scientific theory could be understood as a formal axiomatic system—that is, a logic. Insofar as that system referred to the world, it did so because some of its basic sentences could be understood in terms of observations or operations which one could perform to test them. The rest of the theoretical system, including sentences using theoretical or unobservable terms (like electron or force) would then either be meaningful because they could be reduced to observations, or they had purely logical meanings (called analytic, like mathematical identities). This has been referred to as the verifiability criterion of meaning. According to the criterion, any statement not either analytic or verifiable was strictly meaningless. Although the view was endorsed by Carnap in 1928, he would later come to see it as too restrictive (Carnap 1956). Another familiar version of this idea is operationalism of Percy William Bridgman. In *The Logic of Modern Physics* (1927) Bridgman asserted that every physical concept could be defined in terms of the operations one would perform to verify the application of that concept. Making good on the operationalisation of a concept even as simple as length, however, can easily become enormously complex (for measuring very small lengths, for instance) or impractical (measuring large distances like light years.)

Carl Hempel's (1950, 1951) criticisms of the verifiability criterion of meaning had enormous influence. He pointed out that universal generalizations, such as most scientific laws, were not strictly meaningful on the criterion. Verifiability and operationalism both seemed too restrictive to capture standard scientific aims and practice. And the tenuous connection between these reconstructions and actual scientific practice was criticized in another way. In both approaches, what are scientific methods are instead recast in methodological roles. Measurements, for example, were looked to as ways of giving meanings to terms. The aim of the philosopher of science was not to understand the methods *per se*, but to use them to reconstruct theories, their meanings, and their relation to the world. When scientists perform these operations, however, they will not report that they are doing them to give meaning to terms in a formal axiomatic system. This disconnect between methodology and the details of actual scientific practice would seem to violate the empiricism the Logical Positivists, or Bridgman, were committed to. The view that methodology should correspond to practice (to some extent) has been called historicism, or intuitionism. We turn to these criticisms and responses in section 3.4.<sup>[4]</sup>

Positivism also had to contend with the recognition that a purely inductivist approach, along the lines of Bacon-Newton-Mill, was untenable. There was no pure observation, for starters. All observation was theory laden. Theory is required to make any observation, therefore not all theory can be derived from observation alone. (See also the entry on theory and observation in science). Even granting an observational basis, Hume had already pointed out that one could not argue for inductive conclusions without begging the question by presuming the success of the inductive method. Likewise, positivist attempts at analyzing how a generalization can be confirmed by observations of its instances were subject to a number of criticisms. In his riddle of induction, Goodman (1965) pointed out that for a set of observations, there will be multiple hypotheses that are equally supported. For example, the observation that all emeralds examined before today were green would support equally the two generalization 'all emeralds are green' and 'all emeralds are grue' where 'x is grue' iff either x has been examined before today and is green or x has not been examined before today and is blue. Goodman suggested that one could distinguish between generalizations that were supported by their instances and those that were not by comparing the entrenchment of their predicates-that is, the degree to which they have formed part of generalizations that have successfully been projected to account for new instances. In this way 'all emeralds are green' could be distinguished as more entrenched than 'all emeralds are grue'. In the 'Raven Paradox', Hempel (1965) pointed out that if an observation confirms a given hypothesis, it also confirms all other hypotheses that are logically equivalent to it. For example, the generalization 'all ravens are black' is logically equivalent to the generalization 'all non-black objects are non-ravens', and the observation of a black raven, a red herring and a white shoe would therefore all confirm the hypothesis that ravens are black. Many find this paradoxical, but Hempel maintained that our intuition is based on a tacit appeal to background knowledge on the prevalence of ravens and non-ravens that prompt us to give more weight to evidence of ravens being black than to evidence of non-black items being non-ravens. (for more on these points of criticism as well as how they have been met, see the entries on confirmation and the problem of induction). We shall return to more recent attempts at explaining how observations can serve to confirm a scientific theory in section 4 below.

#### 3.2. H-D as a logic of confirmation

The standard starting point for a non-inductive analysis of the logic of confirmation is known as the Hypothetico-Deductive (H-D) method. In its simplest form, the idea is that a theory, or more specifically a sentence of that theory which expresses some hypothesis, is confirmed by its true consequences. As noted in <u>section 2</u>, this method had been advanced by Whewell in the 19<sup>th</sup> century, as well as Nicod (1924) and others in the 20<sup>th</sup> century. Often, Hempel's (1966) description of the H-D method illustrated by the case of Semmelweiss' inferential procedures in establishing the cause of childbed fever has been presented as a key account of H-D as well as a foil for criticism of the H-D account of confirmation (see, for example, Lipton's (2004) discussion of inference to the best explanation; also the entry on <u>confirmation</u>). Hempel described Semmelsweiss' procedure as examining various hypotheses that would answer the question about the cause of childbed fever. Some hypotheses conflicted with observable facts and could be rejected as false immediately. Others needed to be tested experimentally by deducing which observable events should follow if the hypothesis were true (what Hempel called the test implications of the hypothesis, then conducting an experiment and observing whether or not the test implications occurred. If the experiment showed the test implication to be false, the hypothesis could be rejected. On the other hand, if the experiment showed the test implication to a test implication does not verify a hypothesis, Hempel did

allow that "it provides at least some support, some corroboration or confirmation for it" (Hempel 1966: 8). The degree of this support then depends on the quantity, variety and precision of the supporting evidence.

#### 3.3. Popper and falsificationism

Another approach that took off from the difficulties with inductive inference was <u>Karl Popper's</u> critical rationalism or falsificationism (Popper 1959, 1963). Falsification is deductive and similar to H-D in that it involves scientists deducing observational consequences from the hypothesis under test. For Popper, however, the important point was not whatever confirmation successful prediction offered to the hypotheses but rather the logical asymmetry between such confirmations, which require an inductive inference, versus falsification, which can be based on a deductive inference. This simple opposition was later questioned, by Lakatos, among others. (See the entry on <u>historicist theories of scientific rationality</u>.)

Popper stressed that, regardless of the amount of confirming evidence, we can never be certain that a hypothesis is true without committing the fallacy of affirming the consequent. Instead, Popper introduced the notion of corroboration as a measure for how well a theory or hypothesis has survived previous testing—but without implying that this is also a measure for the probability that it is true.

Popper was also motivated by his doubts about the scientific status of theories like the Marxist theory of history or psycho-analysis, and so wanted to draw a line of demarcation between science and pseudo-science. Popper saw this as an importantly different distinction than demarcating science from metaphysics. The latter demarcation was the primary concern of many logical empiricists. Popper used the idea of falsification to draw a line instead between pseudo and proper science. Science was science because it subjected its theories to rigorous tests which offered a high probability of failing and thus refuting the theory. The aim was not, in this way, to verify a theory. This could be done all too easily, even in cases where observations were at first inconsistent with the deduced consequences of the theory, for example by introducing auxiliary hypotheses designed explicitly to save the theory, so-called *ad hoc* modifications. This was what he saw done in pseudo-science where the theories appeared to be able to explain anything that happened within the field to which they applied. In contrast, science is risky; if observations showed the predictions from a theory to be absent, the theory would be refuted. Hence, scientific hypotheses must be falsifiable. Not only must there exist some possible observation statement which could falsify the hypothesis or theory, were it observed, (Popper called these the hypothesis' potential falsifiers) it is crucial to the Popperian scientific method that such falsifications be sincerely attempted on a regular basis.

The more potential falsifiers of a hypothesis, the more falsifiable it would be, and the more the hypothesis claimed. Conversely, hypotheses without falsifiers claimed very little or nothing at all. Originally, Popper thought that this meant the introduction of *ad hoc* hypotheses only to save a theory should not be countenanced as good scientific method. These would undermine the falsifiabilility of a theory. However, Popper later came to recognize that the introduction of modifications (immunizations, he called them) was often an important part of scientific development. Responding to surprising or apparently falsifying observations often generated important new scientific insights. Popper's own example was the observed motion of Uranus which originally did not agree with Newtonian predictions, but the *ad hoc* hypothesis of an outer planet explained the disagreement and led to further falsifiable predictions. Popper sought to reconcile the view by blurring the distinction between falsifiable and not falsifiable, and speaking instead of degrees of testability (Popper 1985: 41f.).

#### 3.4 Meta-methodology and the end of method

From the 1960s on, sustained meta-methodological criticism emerged that drove the philosophical focus away from scientific method. Something brief about those criticisms must be said here, but recommendations for further reading can be found at the end of the entry.

Thomas Kuhn's The Structure of Scientific Revolutions (1962) begins with a well-known shot across the bow for philosophers of science:

History, if viewed as a repository for more than anecdote or chronology, could produce a decisive transformation in the image of science by which we are now possessed. (1962: 1)

The kind of image Kuhn wanted to transform was the a-historical, rational reconstruction sought by many of the Logical Positivists, though Carnap and other positivists were actually quite sympathetic to Kuhn's views. (See the entry on the <u>Vienna Circle</u>). Kuhn shares with other of his contemporaries, such as Feyerabend and Lakatos, a commitment to a more empirical approach to philosophy of science. Namely, the history of science provides important data, and necessary checks, for philosophy of science, including any theory of scientific method.

An examination of the history of science reveals, according to Kuhn, that scientific development occurs in alternating phases. During normal science, the members of the scientific community adhere to the paradigm in place. Their commitment to the paradigm means a commitment to the puzzles to be solved and the acceptable ways of solving them. Confidence in the paradigm remains so long as steady progress is made in solving the shared puzzles. Method in this normal phase operates within a disciplinary matrix (Kuhn's later concept of a paradigm) which includes standards for problem solving, as well as defines the range of problems the method should be applied to. An important part of a disciplinary matrix is the set of values which provide the norms and aims for scientific method. The main values that Kuhn identifies are prediction, problem solving, simplicity, consistency, and plausibility.

An important by-product of normal science, however, is the accumulation of puzzles which cannot be solved utilizing the resources of the current paradigm. Once the accumulation of these anomalies has reached some critical mass, it can trigger a communal shift to a new paradigm and a new phase of normal science. Importantly, the values that provide the norms and aims for scientific method may have transformed in the meantime. Method may therefore be relative to discipline, time or place

Feyerabend also identified the aims of science as progress, but argued that any methodological prescription would only stifle that progress (Feyerabend 1988). His arguments are grounded in re-examining accepted "myths" about the history of science. Heroes of science, like Galileo, are shown to be just as reliant on rhetoric and persuasion as they are on reason and demonstration. Others, like Aristotle, are shown to be far more reasonable and far-reaching in their outlooks then they are given credit for. As a consequence, the only rule that could provide what he took to be sufficient freedom was the vacuous "anything goes". More generally, even the methodological restriction that science is the best way to pursue knowledge, and to increase knowledge, is too restrictive. Feyerabend suggested instead that science might, in fact, be a threat to a free society, because it and its myth had become so dominant (Feyerabend 1978).

An even more fundamental kind of criticism was offered by several sociologists of science from the 1970s onwards who dismissed what they saw as a false distinction between philosophical accounts of the rational development of science and sociological accounts of the irrational mistakes. Instead, they adhered to a symmetry thesis on which any causal explanation of how scientific knowledge is established needs to be symmetrical in explaining truth and falsity, rationality and irrationality, success and mistakes by the same causal factors (see, e.g., Barnes and Bloor 1982, Bloor 1991). Movements in the Sociology of Science, like the Strong Programme, or in the social dimensions and causes of knowledge more generally led to extended and close examination of detailed case studies in contemporary science and its history. (See the entries on the social dimensions of scientific knowledge and social epistemology.) Well-known examinations by Latour and Woolgar (1979/1986), Knorr-Cetina (1981), Pickering (1984), Shapin and Schaffer (1985) seemed to bear out that it was social ideologies (on a macro-scale) or individual interactions and circumstances (on a micro-scale) which were the primary causal factors in determining which beliefs gained the status of scientific knowledge. As they saw it, in other words, explanatory appeals to scientific method were not empirically well grounded.

By the close of the 20<sup>th</sup> century the search by philosophers for the scientific method was flagging. Nola and Sankey (2000b) could introduce their volume on method by remarking that "For some, the whole idea of a theory of scientific method is yester-year's debate ...".

# 4. Statistical methods for hypothesis testing

Despite the many difficulties that philosophers encountered in trying to providing a clear methodology of conformation (or refutation), still important progress has been made on understanding how observation can provide evidence for a given theory. Work in statistics has been crucial for understanding how theories can be tested empirically, and in recent decades a huge literature has developed that attempts to recast confirmation in Bayesian terms. Here these developments can be covered only briefly, and we refer to the entry on <u>confirmation</u> for further details and references.

Statistics has come to play an increasingly important role in the methodology of the experimental sciences from the 19<sup>th</sup> century onwards. At that time, statistics and probability theory took on a methodological role as an analysis of inductive inference, and attempts to ground the rationality of induction in the axioms of probability theory have continued throughout the 20<sup>th</sup> century and in to the present. Developments in the theory of statistics itself, meanwhile, have had a direct and immense influence on the experimental method, including methods for measuring the uncertainty of observations such as the Method of Least Squares developed by Legendre and Gauss in the early 19<sup>th</sup> century, criteria for the rejection of outliers proposed by Peirce by the mid-19<sup>th</sup> century, and the significance tests developed by Gosset (a.k.a. "Student"), Fisher, Neyman & Pearson and others in the 1920s and 1930s (see, e.g., Swijtink 1987 for a brief historical overview; and also the entry on <u>C.S.</u> Peirce).

These developments within statistics then in turn led to a reflective discussion among both statisticians and philosophers of science on how to perceive the process of hypothesis testing: whether it was a rigorous statistical inference that could provide a numerical expression of the degree of confidence in the tested hypothesis, or if it should be seen as a decision between different courses of actions that also involved a value component. This led to a major controversy among Fisher on the one side and Neyman and Pearson on the other (see especially Fisher 1955, Neyman 1956 and Pearson 1955, and for analyses of the controversy, e.g., Howie 2002, Marks 2000, Lenhard 2006). On Fisher's view, hypothesis testing was a methodology for when to accept or reject a statistical hypothesis, namely that a hypothesis should be rejected by evidence if this evidence would be unlikely relative to other possible outcomes, given the hypothesis were true. In contrast, on Neyman and Pearson's view, the consequence of error also had to play a role when deciding between hypotheses. Introducing the distinction between the error of rejecting a true hypothesis (type II error), they argued that it depends on the consequences of the error to decide whether it is more important to avoid rejecting a true hypothesis or accepting a false one. Hence, Fisher aimed for a theory of inductive inference that enabled a numerical expression of confidence in a hypothesis. To him, the important point was the search for truth, not utility. In contrast, the Neyman-Pearson approach provide a strategy of inductive behaviour for deciding between different courses of action. Here, the important point was not whether a hypothesis was true, but whether one should act as if it was.

Similar discussions are found in the philosophical literature. On the one side, Churchman (1948) and Rudner (1953) argued that because scientific hypotheses can never be completely verified, a complete analysis of the methods of scientific inference includes ethical judgments in which the scientists must decide whether the evidence is sufficiently strong or that the probability is sufficiently high to warrant the acceptance of the hypothesis, which again will depend on the importance of making a mistake in accepting or rejecting the hypothesis. Others, such as Jeffrey (1956) and Levi (1960) disagreed and instead defended a value-neutral view of science on which scientists should bracket their attitudes, preferences, temperament, and values when assessing the correctness of their inferences. For more details on this value-free ideal in the philosophy of science and its historical development, see Douglas (2009) and Howard (2003).

In recent decades, philosophical discussions of the evaluation of probabilistic hypotheses by statistical inference have largely focused on Bayesianism that understands probability as a measure of a person's degree of belief in an event, given the available information, and frequentism that instead understands probability as a long-run frequency of a repeatable event. Hence, for Bayesians probabilities refer to a state of knowledge, whereas for frequentists probabilities refer to frequencies of events (see, e.g., Sober 2008, chapter 1 for a detailed introduction to Bayesianism and frequentism as well as to likelihoodism). Bayesianism aims at providing a quantifiable, algorithmic representation of belief revision, where belief revision is a function of prior beliefs (i.e., background knowledge) and incoming evidence. Bayesianism employs a rule based on Bayes' theorem, a theorem of the probability calculus which relates conditional probabilities. The probability that a particular hypothesis is true is interpreted as a degree of belief, or credence, of the scientist. There will also be a probability and a degree of belief that a hypothesis to that conditional on a piece of evidence (an observation, say) being true. Bayesianism proscribes that it is rational for the scientist to update their belief in the hypothesis to that conditional probability should it turn out that the evidence is, in fact, observed. Originating in the work of Neyman and Person, frequentism aims at providing the tools for reducing long-run error rates, such as the error-statistical approach developed by Mayo (1996) that focuses on how experimenters can avoid both type I and type II errors by building up a repertoire of procedures that detect errors if and only if they are present. Both Bayesianism and frequentism have developed over time, they are interpreted in different ways by its various proponents, and their relations to previous criticism to attempts at defining scientific method are seen differently by proponents and critics. The liter

# 5. Method in Practice

Attention to scientific practice, as we have seen, is not itself new. However, the turn to practice in the philosophy of science of late can be seen as a correction to the pessimism with respect to method in philosophy of science in later parts of the 20<sup>th</sup> century, and as an attempted reconciliation between sociological and rationalist explanations of scientific knowledge. Much of this work sees method as detailed and context specific problem-solving procedures, and methodological analyses to be at the same time descriptive, critical and advisory (see Nickles 1987 for an exposition of this view). The following section contains a survey of some of the practice focuses. In this section we turn fully to topics rather than chronology.

#### 5.1 Creative and exploratory practices

A problem with the distinction between the contexts of discovery and justification that figured so prominently in philosophy of science in the first half of the 20<sup>th</sup> century (see section 2) is that no such distinction can be clearly seen in scientific activity (see Arabatzis 2006). Thus, in recent decades, it has been recognized that study of conceptual innovation and change should not be confined to psychology and sociology of science, but are also important aspects of scientific practice which philosophy of science should address (see also the entry on scientific discovery). Looking for the practices that drive conceptual innovation has led philosophers to examine both the reasoning practices of scientists and the wide realm of experimental practices that are not directed narrowly at testing hypotheses, that is, exploratory experimentation.

Examining the reasoning practices of historical and contemporary scientists, Nersessian (2008) has argued that new scientific concepts are constructed as solutions to specific problems by systematic reasoning, and that of analogy, visual representation and thought-experimentation are among the important reasoning practices employed. These ubiquitous forms of reasoning are reliable—but also fallible—methods of conceptual development and change. On her account, model-based reasoning consists of cycles of construction, simulation, evaluation and adaption of models that serve as interim interpretations of the target problem to be solved.

Often, this process will lead to modifications or extensions, and a new cycle of simulation and evaluation. However, Nersessian also emphasizes that

creative model-based reasoning cannot be applied as a simple recipe, is not always productive of solutions, and even its most exemplary usages can lead to incorrect solutions. (Nersessian 2008: 11)

Thus, while on the one hand she agrees with many previous philosophers that there is no logic of discovery, discoveries can derive from reasoned processes, such that a large and integral part of scientific practice is

the creation of concepts through which to comprehend, structure, and communicate about physical phenomena .... (Nersessian 1987: 11)

Similarly, work on heuristics for discovery and theory construction by scholars such as Darden (1991) and Bechtel & Richardson (1993) present science as problem solving and investigate scientific problem solving as a special case of problem-solving in general. Drawing largely on cases from the biological sciences, much of their focus has been on reasoning strategies for the generation, evaluation, and revision of mechanistic explanations of complex systems.

Addressing another aspect of the context distinction, namely the traditional view that the primary role of experiments is to test theoretical hypotheses according to the H-D model, other philosophers of science have argued for additional roles that experiments can play. The notion of exploratory experimentation was introduced to describe experiments driven by the desire to obtain empirical regularities and to develop concepts and classifications in which these regularities can be described (Steinle 1997, 2002; Burian 1997; Waters 2007)). However the difference between theory driven experimentation and exploratory experimentation should not be seen as a sharp distinction. Theory driven experiments are not always directed at testing hypothesis, but may also be directed at various kinds of fact-gathering, such as determining numerical parameters. *Vice versa*, exploratory experiments are usually informed by theory in various ways and are therefore not theory-free. Instead, in exploratory experiments phenomena are investigated without first limiting the possible outcomes of the experiment on the basis of extant theory about the phenomena.

In recent years, the development of high throughput instrumentation in molecular biology and neighbouring fields has given rise to a special type of exploratory experimentation that collects and analyses very large amounts of data, and these new 'omics' disciplines are often said to represent a break with the ideal of hypothesis-driven science (Burian 2007; Elliott 2007; Waters 2007; O'Malley 2007) and instead described as data-driven research (Leonelli 2012; Strasser 2012) or as a special kind of "convenience experimentation" in which many experiments are done simply because they are extraordinarily convenient to perform (Krohs 2012).

#### 5.2 Computer methods and the 'third way' of doing science

The field of omics just described is possible because of the ability of computers to process, in a reasonable amount of time, the huge quantities of data required. Computers allow for more elaborate experimentation (higher speed, better filtering, more variables, sophisticated coordination and control), but also, through modelling and simulations, might constitute a form of experimentation themselves. Here, too, we can pose a version of the general question of method versus practice: does the practice of using computers fundamentally change scientific method, or merely provide a more efficient means of implementing standard methods?

Because computers can be used to automate measurements, quantifications, calculations, and statistical analyses where, for practical reasons, these operations cannot be otherwise carried out, many of the steps involved in reaching a conclusion on the basis of an experiment are now made inside a "black box", without the direct involvement or awareness of a human. This has epistemological implications, regarding what we can know, and how we can know it. To have confidence in the results, computer methods are therefore subjected to tests of verification and validation.

The distinction between verification and validation is easiest to characterize in the case of computer simulations. In a typical computer simulation scenario computers are used to numerically integrate differential equations for which no analytic solution is available. The equations are part of the model the scientist uses to represent a phenomenon or system under investigation. Verifying a computer simulation means checking that the equations of the model are being correctly approximated. Validating a simulation means checking that the equations of the model are dequate for the inferences one wants to make on the basis of that model.

A number of issues related to computer simulations have been raised. The identification of validity and verification as the testing methods has been criticized. Oreskes et al. (1994) raise concerns that "validiation", because it suggests deductive inference, might lead to over-confidence in the results of simulations. The distinction itself is probably too clean, since actual practice in the testing of simulations mixes and moves back and forth between the two (Weissart 1997; Parker 2008a; Winsberg 2010). Computer simulations do seem to have a non-inductive character, given that the principles by which they operate are built in by the programmers, and any results of the simulation follow from those in-built principles in such a way that those results could, in principle, be deduced from the program code and its inputs. The status of simulations as experiments has therefore been examined (Kaufmann and Smarr 1993; Humphreys 1995; Hughes 1999; Norton and Suppe 2001). This literature considers the epistemology of these experiments: what we can learn by simulation, and also the kinds of justifications which can be given in applying that knowledge to the "real" world. (Mayo 1996; Parker 2008b). As pointed out, part of the advantage of computer simulation derives from the fact that huge numbers of calculations can be carried out without requiring direct observation by the experimenter/simulator. At the same time, many of these calculations are approximations to the calculations which would be performed first-hand in an ideal situation. Both factors introduce uncertainties into the inferences drawn from what is observed in the simulation.

For many of the reasons described above, computer simulations do not seem to belong clearly to either the experimental or theoretical domain. Rather, they seem to crucially involve aspects of both. This has led some authors, such as Fox Keller (2003: 200) to argue that we ought to consider computer simulation a "qualitatively different way of doing science". The literature in general tends to follow Kaufmann and Smarr (1993) in referring to computer simulation as a "third way" for scientific methodology (theoretical reasoning and experimental practice are the first two ways.). It should also be noted that the debates around these issues have tended to focus on the form of computer simulation typical in the physical sciences, where models are based on dynamical equations. Other forms of simulation might not have the same problems, or have problems of their own (see the entry on computer simulations in science).

## 6. Discourse on scientific method

Despite philosophical disagreements, the idea of *the* scientific method still figures prominently in contemporary discourse on many different topics, both within science and in society at large. Often, reference to scientific method is used in ways that convey either the legend of a single, universal method characteristic of all science, or grants to a particular method or set of methods privilege as a special 'gold standard', often with reference to particular philosophers to vindicate the claims. Discourse on scientific method also typically arises when there is a need to distinguish between science and other activities, or for justifying the special status conveyed to science. In these areas, the philosophical attempts at identifying a set of methods characteristic for scientific endeavors are closely related to the philosophy of science's classical problem of demarcation (see the entry on <u>science and pseudo-science</u>) and to the philosophical analysis of the social dimension of scientific knowledge and the role of science in democratic society.

### 6.1 "The scientific method" in science education and as seen by scientists

One of the settings in which the legend of a single, universal scientific method has been particularly strong is science education (see, e.g., Bauer 1992; McComas 1996; Wivagg & Allchin 2002).<sup>[5]</sup> Often, 'the scientific method' is presented in textbooks and educational web pages as a fixed four or five step procedure starting from observations and description of a phenomenon and progressing over formulation of a hypothesis which explains the phenomenon, designing and conducting experiments to test the hypothesis, analyzing the results, and ending with drawing a conclusion. Such references to a universal scientific method can be found in educational material at all levels of science education (Blachowicz 2009), and numerous studies have shown that the idea of a general and universal scientific method often form part of both students' and teachers' conception of science (see, e.g., Aikenhead 1987; Osborne et al. 2003).

Although occasionally phrased with reference to the H-D method, important historical roots of the legend in science education of a single, universal scientific method are the American philosopher and psychologist Dewey's account of inquiry in *How We Think* (1910) and the British mathematician Karl Pearson's account of science in *Grammar of Science* (1892). On Dewey's account, inquiry is divided into the five steps of

(i) a felt difficulty, (ii) its location and definition, (iii) suggestion of a possible solution, (iv) development by reasoning of the bearing of the suggestions, (v) further observation and experiment leading to its acceptance or rejection. (Dewey 1910: 72)

Similarly, on Pearson's account, scientific investigations start with measurement of data and observation of their correction and sequence from which scientific laws can be discovered with the aid of creative imagination. These laws have to be subject to criticism, and their final acceptance will have equal validity for "all normally constituted minds". Both Dewey's and Pearson's accounts should be seen as generalized abstractions of inquiry and not restricted to the realm of science—although both Dewey and Pearson referred to their respective accounts as 'the scientific method'.

Occasionally, scientists make sweeping statements about a simple and distinct scientific method, as exemplified by Feynman's simplified version of a conjectures and refutations method presented, for example, in the last of his 1964 Cornell Messenger lectures.<sup>[6]</sup> However, just as often scientists have come to the same conclusion as recent philosophy of science that there is not any unique, easily described scientific method. For example, the physicist and Nobel Laureate Weinberg described in the paper "The Methods of Science ... And Those By Which We Live" (1995) how

The fact that the standards of scientific success shift with time does not only make the philosophy of science difficult; it also raises problems for the public understanding of science. We do not have a fixed scientific method to rally around and defend. (1995: 8)

#### 6.2 Privileged methods and 'gold standards'

Reference to the scientific method has also often been used to argue for the scientific nature or special status of a particular activity. Philosophical positions that argue for a simple and unique scientific method as a criterion of demarcation, such as Popperian falsification, have often attracted practitioners who felt that they had a need to defend their domain of practice. For example, references to conjectures and refutation as the scientific method are abundant in much of the literature on complementary and alternative medicine (CAM)—alongside the competing position that CAM, as an alternative to conventional biomedicine, needs to develop its own methodology different from that of science.

Also within mainstream science, reference to the scientific method is used in arguments regarding the internal hierarchy of disciplines and domains. A frequently seen argument is that research based on the H-D method is superior to research based on induction from observations because in deductive inferences the conclusion follows necessarily from the premises. (See, e.g., Parascandola (1998) for an analysis of how this argument has been made to downgrade epidemiology compared to the laboratory sciences.) Similarly, based on an examination of the practices of major funding institutions such as the National Institutes of Health (NIH), the National Science Foundation (NSF) and the Biomedical Sciences Research Practices (BBSRC) in the UK, O'Malley et al. (2009) have argued that funding agencies seem to have a tendency to adhere to the view that the primary activity of science is to test hypotheses, while descriptive and exploratory research is seen as merely preparatory activities that are valuable only insofar as they fuel hypothesis-driven research.

In some areas of science, scholarly publications are structured in a way that may convey the impression of a neat and linear process of inquiry from stating a question, devising the methods by which to answer it, collecting the data, to drawing a conclusion from the analysis of data. For example, the codified format of publications in most biomedical journals known as the IMRAD format (Introduction, Method, Results, Analysis, Discussion) is explicitly described by the journal editors as "not an arbitrary publication format but rather a direct reflection of the process of scientific discovery" (see the so-called "Vancouver Recommendations", ICMJE 2013: 11). However, scientific publications do not in general reflect the process by which the reported scientific results were produced. For example, under the provocative title "Is the scientific paper a fraud?", Medawar argued that scientific papers generally misrepresent how the results have been produced (Medawar 1963/1996). Similar views have been advanced by philosophers, historians and sociologists of science (Gilbert 1976; Holmes 1987; Knorr-Cetina 1981; Schickore 2008; Suppe 1998) who have argued that scientists' experimental practices are messy and often do not follow any recognizable pattern. Publications of research results, they argue, are retrospective reconstructions of these activities that often do not preserve the temporal order or the logic of these activities, but are instead often constructed in order to screen off potential criticism (see Schickore 2008 for a review of this work).

#### 6.3 Scientific method in the court room

Philosophical positions on the scientific method have also made it into the court room, especially in the US where judges have drawn on philosophy of science in deciding when to confer special status to scientific expert testimony. A key case is *Daubert vs Merrell Dow Pharmaceuticals* (92-102, 509 U.S. 579, 1993). In this case, the Supreme Court argued in its 1993 ruling that trial judges must ensure that expert testimony is reliable, and that in doing this the court must look at the expert's methodology to determine whether the proffered evidence is actually scientific knowledge. Further, referring to works of Popper and Hempel the court stated that

ordinarily, a key question to be answered in determining whether a theory or technique is scientific knowledge ... is whether it can be (and has been) tested. (Justice Blackmun, Daubert v. Merrell Dow Pharmaceuticals; see Other Internet Resources for a link to the opinion)

But as argued by Haack (2005a,b, 2010) and by Foster & Hubner (1999), by equating the question of whether a piece of testimony is reliable with the question whether it is scientific as indicated by a special methodology, the court was producing an inconsistent mixture of Popper's and Hempel's philosophies, and this has later led to considerable confusion in subsequent case rulings that drew on the Daubert case (see Haack 2010 for a detailed exposition).

#### 6.4 Deviating practices

The difficulties around identifying the methods of science are also reflected in the difficulties of identifying scientific misconduct in the form of improper application of the method or methods of science. One of the first and most influential attempts at defining misconduct in science was the US definition from 1989 that defined misconduct as

fabrication, falsification, plagiarism, or other practices that seriously deviate from those that are commonly accepted within the scientific community. (Code of Federal Regulations, part 50, subpart A., August 8, 1989, italics added)

However, the "other practices that seriously deviate" clause was heavily criticized because it could be used to suppress creative or novel science. For example, the National Academy of Science stated in their report *Responsible Science* (1992) that it

wishes to discourage the possibility that a misconduct complaint could be lodged against scientists based solely on their use of novel or unorthodox research methods. (NAS: 27)

This clause was therefore later removed from the definition. For an entry into the key philosophical literature on conduct in science, see Shamoo & Resnick (2009).

## 7. Conclusion

The question of the source of the success of science has been at the core of philosophy since the beginning of modern science. If viewed as a matter of epistemology more generally, scientific method is a part of the entire history of philosophy. Over that time, science and whatever methods its practioners may employ have changed dramatically. Today, many philosophers have taken up the banners of pluralism or of practice to focus on what are, in effect, fine-grained and contextually limited examinations of scientific method. Others hope to shift perspectives in order to provide a renewed general account of what characterizes the activity we call science.

One such perspective has been offered recently by Hoyningen-Huene (2008, 2013), who argues from the history of philosophy of science that after three lengthy phases of characterizing science by its method, we are now in a phase where the belief in the existence of a positive scientific method has eroded and what has been left to characterize science is only its fallibility. First was a phase from Plato and Aristotle up until the 17<sup>th</sup> century where the specificity of scientific knowledge was seen in its absolute certainty established by proof from evident axioms; next was a phase up to the mid-19<sup>th</sup> century in which the means to establish the certainty of scientific knowledge had been generalized to include inductive procedures as well. In the third phase, which lasted until the last decades of the 20<sup>th</sup> century, it was recognized that empirical knowledge was fallible, but it was still granted a special status due to its distinctive mode of production. But now in the fourth phase, according to Hoyningen-Huene, historical and philosophical studies have shown how "scientific methods with the characteristics as posited in the second and third phase do not exist" (2008: 168) and there is no longer any consensus among philosophers and historians of science about the nature of science. For Hoyningen-Huene, this is too negative a stance, and he therefore urges the question about the nature of science anew. His own answer to this question is that "scientific knowledge differs from other kinds of knowledge, especially everyday knowledge, primarily by being more systematic" (Hoyningen-Huene 2013: 14). Systematicity can have several different dimensions: among them are more systematic descriptions, explanations, predictions, defense of knowledge claims, epistemic connectedness, ideal of completeness, knowledge generation, representation of knowledge, etc. On this position, what characterizes science is not that the methods employed are unique to science, but that the methods are more carefully employed.

Another, similar approach has been offered by Haack (2003). She sets off, similar to Hoyningen-Huene, from a dissatisfaction with the recent clash between what she calls Old Deferentialism and New Cynicism. The Old Deferentialist position is that science progressed inductively by accumulating true theories confirmed by empirical evidence or deductively by testing conjectures against basic statements; while the New Cynics position is that science has no epistemic authority and no uniquely rational method and is merely just politics. Haack insists that contrary to the views of the New Cynics, there are objective epistemic standards, and there is something epistemologically special about science, even though the Old Deferentialists pictured this in a wrong way. Instead, she offers a new Critical Commonsensist account on which standards of good, strong, supportive evidence and well-conducted, honest, thorough and imaginative inquiry are not exclusive to the sciences, but the standards by which we judge all inquirers. In this sense, science does not differ in kind from other kinds of inquiry, but it may differ in the degree to which it requires broad and detailed background knowledge and a familiarity with a technical vocabulary that only specialists may possess.

# **Bibliography**

- Aikenhead, G.S., 1987, "High-school graduates' beliefs about science-technology-society. III. Characteristics and limitations of scientific knowledge", *Science Education*, 71(4): 459–487.
- Arabatzis, T., 2006, "On the inextricability of the context of discovery and the context of justification", in *Revisiting Discovery and Justification*, J. Schickore and F. Steinle (eds.), Dordrecht: Springer, pp. 215–230.
- Barnes, B. and D. Bloor, 1982, "Relativism, Rationalism, and the Sociology of Knowledge", in *Rationality and Relativism*, M. Hollis and S. Lukes (eds.), Cambridge: MIT Press, pp. 1–20.
- Bauer, H.H., 1992, Scientific Literacy and the Myth of the Scientific Method, Urbana: University of Illinois Press.
- Bechtel, W. and R.C. Richardson, 1993, Discovering complexity, Princeton, NJ: Princeton University Press.
- Berkeley, G., 1734, *The Analyst* in *De Motu and The Analyst: A Modern Edition with Introductions and Commentary*, D. Jesseph (trans. and ed.), Dordrecht: Kluwer Academic Publishers, 1992.
- Blachowicz, J., 2009, "How science textbooks treat scientific method: A philosopher's perspective", *The British Journal for the Philosophy of Science*, 60(2): 303–344.
- Bloor, D., 1991, Knowledge and Social Imagery, Chicago: University of Chicago Press, 2nd edition.
- Boyle, R., 1682, New experiments physico-mechanical, touching the air, Printed by Miles Flesher for Richard Davis, bookseller in Oxford.
- Bridgman, P.W., 1927, The Logic of Modern Physics, New York: Macmillan.

—, 1956, "The Methodological Character of Theoretical Concepts", in *The Foundations of Science and the Concepts of Science and Psychology*, Herbert Feigl and Michael Scriven (eds.), Minnesota: University of Minneapolis Press, pp. 38–76.

- Burian, R., 1997, "Exploratory Experimentation and the Role of Histochemical Techniques in the Work of Jean Brachet, 1938–1952", *History and Philosophy of the Life Sciences*, 19(1): 27–45.
- ----, 2007, "On microRNA and the need for exploratory experimentation in post-genomic molecular biology", *History and Philosophy of the Life Sciences*, 29(3): 285–311.
- Carnap, R., 1928, Der logische Aufbau der Welt, Berlin: Bernary, transl. by R.A. George, The Logical Structure of the World, Berkeley: University of California Press, 1967.
- ----, 1956, "The methodological character of theoretical concepts", Minnesota studies in the philosophy of science, 1: 38–76.
- Churchman, C.W., 1948, "Science, Pragmatics, Induction", Philosophy of Science, 15(3): 249-268.
- Darden, L., 1991, Theory Change in Science: Strategies from Mendelian Genetics, Oxford: Oxford University Press
- Dewey, J., 1910, How we think, New York: Dover Publications (reprinted 1997).
- Douglas, H., 2009, Science, Policy, and the Value-Free Ideal, Pittsburgh: University of Pittsburgh Press.
- Elliott, K.C., 2007, "Varieties of exploratory experimentation in nanotoxicology", History and Philosophy of the Life Sciences, 29(3): 311-334.

Falcon, Andrea, 2005, Aristotle and the science of nature: Unity without uniformity, Cambridge: Cambridge University Press.

- Feyerabend, P., 1978, Science in a Free Society, London: New Left Books
- ----, 1988, Against Method, London: Verso, 2<sup>nd</sup> edition.
- Fisher, R.A., 1955, "Statistical Methods and Scientific Induction", Journal of The Royal Statistical Society. Series B (Methodological), 17(1): 69–78.
- Foster, K. and P.W. Huber, 1999, Judging Science. Scientific Knowledge and the Federal Courts, Cambridge: MIT Press.
- Fox Keller, E., 2003, "Models, Simulation, and 'computer experiments", in *The Philosophy of Scientific Experimentation*, H. Radder (ed.), Pittsburgh: Pittsburgh University Press, 198–215.
- Gilbert, G., 1976, "The transformation of research findings into scientific knowledge", Social Studies of Science, 6: 281-306.
- Gimbel, S., 2011, Exploring the Scientific Method, Chicago: University of Chicago Press.
- Goodman, N., 1965, Fact, Fiction, and Forecast, Indianapolis: Bobbs-Merrill.
- Haack, S., 1995, "Science is neither sacred nor a confidence trick", Foundations of Science, 1(3): 323-335.
- ----, 2003, Defending science-within reason, Amherst: Prometheus.
- —, 2005a, "Disentangling Daubert: an epistemological study in theory and practice", Journal of Philosophy, Science and Law, 5, <u>Haack 2005a available online</u>. doi:10.5840/jpsl2005513
- ----, 2005b, "Trial and error: The Supreme Court's philosophy of science", American Journal of Public Health, 95: S66-S73.
- ----, 2010, "Federal Philosophy of Science: A Deconstruction-and a Reconstruction", NYUJL & Liberty, 5: 394.
- Harper, W.L., 2011, Isaac Newton's Scientific Method: Turning Data into Evidence about Gravity and Cosmology, Oxford: Oxford University Press.
- Hempel, C., 1950, "Problems and Changes in the Empiricist Criterion of Meaning", Revue Internationale de Philosophie, 41(11): 41-63.
- ----, 1951, "The Concept of Cognitive Significance: A Reconsideration", Proceedings of the American Academy of Arts and Sciences, 80(1): 61–77.
- ----, 1965, Aspects of scientific explanation and other essays in the philosophy of science, New York-London: Free Press.
- —, 1966, *Philosophy of Natural Science*, Englewood Cliffs: Prentice-Hall.
- Holmes, F.L., 1987, "Scientific writing and scientific discovery", Isis, 78(2): 220-235.
- Howard, D., 2003, "Two left turns make a right: On the curious political career of North American philosophy of science at midcentury", in *Logical Empiricism in North America*, G.L. Hardcastle & A.W. Richardson (eds), Minneapolis: University of Minnesota Press, pp. 25–93.
- Hoyningen-Huene, P., 2008, "Systematicity: The nature of science", Philosophia, 36(2): 167-180.
- ----, 2013, Systematicity. The Nature of Science, Oxford: Oxford University Press.
- Howie, D., 2002, Interpreting probability: Controversies and developments in the early twentieth century, Cambridge: Cambridge University Press.
- Hughes, R., 1999, "The Ising Model, Computer Simulation, and Universal Physics", in *Models as Mediators*, M. Morgan and M. Morrison (eds.), Cambridge: Cambridge University Press, pp. 97–145
- Hume, D., 1739, A Treatise of Human Nature, D. Fate Norton and M.J. Norton (eds.), Oxford: Oxford University Press, 2000.
- Humphreys, P., 1995, "Computational science and scientific method", Minds and Machines, 5(1): 499-512.
- ICMJE, 2013, "Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals", International Committee of Medical Journal Editors, available online, accessed August 13 2014
- Jeffrey, R.C., 1956, "Valuation and Acceptance of Scientific Hypotheses", Philosophy of Science, 23(3): 237-246.
- Kaufmann, W.J., and L.L. Smarr, 1993, Supercomputing and the Transformation of Science, New York: Scientific American Library.
- Knorr-Cetina, K., 1981, The Manufacture of Knowledge, Oxford: Pergamon Press.
- Krohs, U., 2012, "Convenience experimentation", Studies in History and Philosophy of Biological and BiomedicalSciences, 43: 52-57.
- Kuhn, T.S., 1962, The Structure of Scientific Revolutions, Chicago: University of Chicago Press
- Latour, B. and S. Woolgar, 1986, Laboratory Life: The Construction of Scientific Facts, Princeton: Princeton University Press, 2<sup>nd</sup> edition.
- Laudan, L., 1968, "Theories of scientific method from Plato to Mach", History of Science, 7(1): 1-63.
- Lenhard, J., 2006, "Models and statistical inference: The controversy between Fisher and Neyman-Pearson", *The British Journal for the Philosophy of Science*, 57(1): 69–91.
- Leonelli, S., 2012, "Making Sense of Data-Driven Research in the Biological and the Biomedical Sciences", *Studies in the History and Philosophy of the Biological and Biomedical Sciences*, 43(1): 1–3.
- Levi, I., 1960, "Must the scientist make value judgments?", Philosophy of Science, 57(11): 345-357
- Lindley, D., 1991, Theory Change in Science: Strategies from Mendelian Genetics, Oxford: Oxford University Press.
- Lipton, P., 2004, Inference to the Best Explanation, London: Routledge, 2<sup>nd</sup> edition.
- Marks, H.M., 2000, The progress of experiment: science and therapeutic reform in the United States, 1900-1990, Cambridge: Cambridge University Press.
- Mayo, D.G., 1996, Error and the Growth of Experimental Knowledge, Chicago: University of Chicago Press.
- McComas, W.F., 1996, "Ten myths of science: Reexamining what we think we know about the nature of science", School Science and Mathematics, 96(1): 10-16.
- Medawar, P.B., 1963/1996, "Is the scientific paper a fraud", in *The Strange Case of the Spotted Mouse and Other Classic Essays on Science*, Oxford: Oxford University Press, 33–39.
- Mill, J.S., 1963, Collected Works of John Stuart Mill, J. M. Robson (ed.), Toronto: University of Toronto Press
- NAS, 1992, Responsible Science: Ensuring the integrity of the research process, Washington DC: National Academy Press.
- Nersessian, N.J., 1987, "A cognitive-historical approach to meaning in scientific theories", in *The process of science*, N. Nersessian (ed.), Berlin: Springer, pp. 161–177.
- -----, 2008, Creating Scientific Concepts, Cambridge: MIT Press.
- Newton, I., 1726, *Philosophiae naturalis Principia Mathematica* (3<sup>rd</sup> edition), in *The Principia: Mathematical Principles of Natural Philosophy: A New Translation*, I.B. Cohen and A. Whitman (trans.), Berkeley: University of California Press, 1999.
- ----, 1704, Opticks or A Treatise of the Reflections, Refractions, Inflections & Colors of Light, New York: Dover Publications, 1952.
- Neyman, J., 1956, "Note on an Article by Sir Ronald Fisher", Journal of the Royal Statistical Society. Series B (Methodological), 18: 288–294.
- Nickles, T., 1987, "Methodology, heuristics, and rationality", in *Rational changes in science: Essays on Scientific Reasoning*, J.C. Pitt (ed.), Berlin: Springer, pp. 103–132.
- Nicod, J., 1924, Le problème logique de l'induction, Paris: Alcan. (Engl. transl. "The Logical Problem of Induction", in Foundations of Geometry and Induction, London: Routledge, 2000.)
- Nola, R. and H. Sankey, 2000a, "A selective survey of theories of scientific method", in Nola and Sankey 2000b: 1-65.
- ----, 2000b, After Popper, Kuhn and Feyerabend. Recent Issues in Theories of Scientific Method, London: Springer.
- -, 2007, Theories of Scientific Method, Stocksfield: Acumen.
- Norton, S., and F. Suppe, 2001, "Why atmospheric modeling is good science", in *Changing the Atmosphere: Expert Knowledge and Environmental Governance*, C. Miller and P. Edwards (eds.), Cambridge, MA: MIT Press, 88–133.
- O'Malley, M., 2007, "Exploratory experimentation and scientific practice: Metagenomics and the proteorhodopsin case", *History and Philosophy of the Life Sciences*, 29(3): 337–360.
- O'Malley, M., C. Haufe, K. Elliot, and R. Burian, 2009, "Philosophies of Funding", Cell, 138: 611-615.
- Oreskes, N., K. Shrader-Frechette, and K. Belitz, 1994, "Verification, Validation and Confirmation of Numerical Models in the Earth Sciences", Science, 263(5147):

641–646.

Osborne, J., S. Simon, and S. Collins, 2003, "Attitudes towards science: a review of the literature and its implications", *International Journal of Science Education*, 25(9): 1049–1079.

Parascandola, M., 1998, "Epidemiology-2nd-Rate Science", Public Health Reports, 113(4): 312-320.

Parker, W., 2008a, "Franklin, Holmes and the Epistemology of Computer Simulation", International Studies in the Philosophy of Science, 22(2): 165-83.

----, 2008b, "Computer Simulation through an Error-Statistical Lens", Synthese, 163(3): 371-84.

Pearson, K. 1892, The Grammar of Science, London: J.M. Dents and Sons, 1951

Pearson, E.S., 1955, "Statistical Concepts in Their Relation to Reality", Journal of the Royal Statistical Society, B, 17: 204-207.

Pickering, A., 1984, Constructing Quarks: A Sociological History of Particle Physics, Edinburgh: Edinburgh University Press.

Popper, K.R., 1959, The Logic of Scientific Discovery, London: Routledge, 2002

----, 1963, Conjectures and Refutations, London: Routledge, 2002.

----, 1985, Unended Quest: An Intellectual Autobiography, La Salle: Open Court Publishing Co..

Rudner, R., 1953, "The Scientist Qua Scientist Making Value Judgments", Philosophy of Science, 20(1): 1-6.

Rudolph, J.L., 2005, "Epistemology for the masses: The origin of 'The Scientific Method' in American Schools", *History of Education Quarterly*, 45(3): 341–376 Schickore, J., 2008, "Doing science, writing science", *Philosophy of Science*, 75: 323–343.

Shamoo, A.E. and D.B. Resnik, 2009, *Responsible Conduct of Research*, Oxford: Oxford University Press.

Shank, J.B., 2008, The Newton Wars and the Beginning of the French Enlightenment, Chicago: The University of Chicago Press.

Shapin, S. and S. Schaffer, 1985, Leviathan and the air-pump, Princeton: Princeton University Press.

Smith, G.E., 2002, "The Methodology of the Principia", in *The Cambridge Companion to Newton*, I.B. Cohen and G.E. Smith (eds.), Cambridge: Cambridge University Press, 138–173.

Snyder, L.J., 1997a, "Discoverers' Induction", Philosophy of Science, 64: 580-604.

----, 1997b, "The Mill-Whewell Debate: Much Ado About Induction", Perspectives on Science, 5: 159–198.

----, 1999, "Renovating the Novum Organum: Bacon, Whewell and Induction", Studies in History and Philosophy of Science, 30: 531–557.

Sober, E., 2008, Evidence and Evolution. The logic behind the science, Cambridge: Cambridge University Press

Steinle, F., 1997, "Entering New Fields: Exploratory Uses of Experimentation", Philosophy of Science (Proceedings), 64: S65–S74.

----, 2002, "Experiments in History and Philosophy of Science", Perspectives on Science, 10(4): 408-432.

Strasser, B.J., 2012, "Data-driven sciences: From wonder cabinets to electronic databases", Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 43(1): 85–87.

Suppe, F., 1998, "The Structure of a Scientific Paper", Philosophy of Science, 65(3): 381-405.

Swijtink, Z.G., 1987, "The objectification of observation: Measurement and statistical methods in the nineteenth century", in *The probabilistic revolution. Ideas in History, Vol. 1*, L. Kruger (ed.), Cambridge MA: MIT Press, pp. 261–285.

Waters, C.K., 2007, "The nature and context of exploratory experimentation: An introduction to three case studies of exploratory research", *History and Philosophy of the Life Sciences*, 29(3): 275–284.

Weinberg, S., 1995, "The methods of science... and those by which we live", Academic Questions, 8(2): 7-13.

Weissert, T., 1997, The Genesis of Simulation in Dynamics: Pursuing the Fermi-Pasta-Ulam Problem, New York: Springer Verlag.

William H., 1628, Exercitatio Anatomica de Motu Cordis et Sanguinis in Animalibus, in On the Motion of the Heart and Blood in Animals, R. Willis (trans.), Buffalo: Prometheus Books, 1993.

Winsberg, E., 2010, Science in the Age of Computer Simulation, Chicago: University of Chicago Press.

Wivagg, D. & D. Allchin, 2002, "The Dogma of the Scientific Method", The American Biology Teacher, 64(9): 645-646

## **Academic Tools**

How to cite this entry.

Preview the PDF version of this entry at the Friends of the SEP Society.

Look up this entry topic at the Internet Philosophy Ontology Project (InPhO).

PP Enhanced bibliography for this entry at PhilPapers, with links to its database.

# **Other Internet Resources**

- Blackmun opinion, in Daubert v. Merrell Dow Pharmaceuticals (92-102), 509 U.S. 579 (1993).
- Scientific Method at philpapers. Darrell Rowbottom (ed.).
- <u>Recent Articles | Scientific Method | The Scientist Magazine</u>

# **Related Entries**

al-Kindi | Albert the Great [= Albertus magnus] | Aquinas, Saint Thomas | Arabic and Islamic Philosophy, disciplines in: natural philosophy and natural science | Arabic and Islamic Philosophy, historical and methodological topics in: Greek sources | Arabic and Islamic Philosophy, historical and methodological topics in: influence of Arabic and Islamic Philosophy on the Latin West | Aristotle | Bacon, Francis | Bacon, Roger | Berkeley, George | biology: experiment in | Boyle, Robert | Cambridge Platonists | confirmation | Descartes, René | Enlightenment | epistemology | epistemology: Bayesian | epistemology: social | Feyerabend, Paul | Galileo Galilei | Grosseteste, Robert | Hempel, Carl | Hume, David | Hume, David: Newtonianism and Anti-Newtonianism | induction: problem of | Kant, Immanuel | Kuhn, Thomas | Leibniz, Gottfried Wilhelm | Locke, John | Mill, John Stuart | More, Henry | Neurath, Otto | Newton, Isaac | Newton, Isaac: philosophy | Ockham [Occam], William | operationalism | Peirce, Charles Sanders | Plato | Popper, Karl | rationality: historicist theories of | Reichenbach, Hans | Schlick, Moritz | science: and pseudoscience | science: theory and observation in | science: unity of | scientific discovery | scientific knowledge: social dimensions of | simulations in science | skepticism: medieval | space and time: absolute and relational theories of space and motion | Vienna Circle | Whewell, William | Zabarella, Giacomo

> <u>Copyright © 2015</u> by <u>Hanne Andersen </br/>hanne.andersen@ind.ku.dk</u>> Brian Hepburn </br/><u>brian.hepburn@wichita.edu</u>>

Open access to the SEP is made possible by a world-wide funding initiative. Please Read How You Can Help Keep the Encyclopedia Free



The Stanford Encyclopedia of Philosophy is copyright © 2020 by The Metaphysics Research Lab, Center for the Study of Language and Information (CSLI), Stanford University

Library of Congress Catalog Data: ISSN 1095-5054