Confirmation versus Falsificationism

Ray Scott Percival
Philosopher, Author and Founder of the Karl Popper Web, U.K.

Confirmation and falsification are different strategies for testing theories and characterizing the outcomes of those tests. Roughly speaking, confirmation is the act of using evidence or reason to verify or certify that a statement is true, definite, or approximately true, whereas falsification is the act of classifying a statement as false in the light of observation reports.

In clinical practice, the issue of confirmation versus falsification was highlighted by Karl Popper (1902–94), in his critique of Freud's and Adler's depth psychologies. Wishing to demarcate science from nonscientific pursuits, Popper contrasted Einstein's falsifiable theory with the "confirmable" but unfalsifiable theories of Freud and Adler. On the one hand, almost any behavior could be interpreted as "confirming" Freud's or Adler's theory, because the behavioral evidence is itself interpreted in the light of the theories. This makes confirmations worthless. Moreover, Freud failed to stipulate conditions under which the theory would be falsified. For example, what observation would show that the id does not exist? In contrast, Einstein boldly stipulated decisive experimental observations that, if satisfied, would refute his theory. For example, if the path of light from a distant star had not been observed (by Eddington in 1919) to bend by a certain degree when passing close to the sun, Einstein would have abandoned his theory.

Popper's proposed method is falsificationism. Instead of proposing hypotheses and then seeing if they can be confirmed by evidence, Popper said that science ought to involve making conjectures that can potentially be refuted. Stimulated by a problem, the scientist advances unfounded theories, which are then subjected to relentless attempts to falsify them by observation; confirmation is seen as infeasible and pointless, irrelevant. This barrage of criticism then leads to further, deeper problems, which in turn stimulate further theories, potentially nearer to the truth. Because conjectures are freely made and retained until falsified, no justifications or inductive confirmation rules are needed. This is not a description of science, but a proposal about what scientists ought to do to advance science.

The problem of confirmation versus falsification is an intense focus of debate within a wider context of fundamentally competing epistemologies: justificationism and critical rationalism. Justificationism is the idea that one should accept all and only those positions that one can justify by logic or experience. The heyday of justificationism was the Vienna Circle of the 1920s and 1930s, when its members, such as Rudolph Carnap, Otto Neurath, Moritz Schlick, and Friedrich Waismann, held that there are only two types of knowledge: analytic knowledge, which is justified by formal proof, and scientific knowledge, which is justified by empirical verification. Only those statements that are in principle justifiable by one or other of these methods are, they maintained, amenable to rational discussion. Most contemporary philosophers, though spurning the Vienna Circle's views on verification, are still justificationists as they hold on to the traditional conception of knowledge as justified true belief. However, under the impact of skepticism, many have jettisoned the pursuit of truth as unattainable, settling for simply justified positions.

Popper's answer to justificationism is critical rationalism. Critical rationalism shuns justification as both impossible and unnecessary, substituting the pursuit of truth as the fallible but sometimes attainable aim. Critical

The Encyclopedia of Clinical Psychology, First Edition. Edited by Robin L. Cautin and Scott O. Lilienfeld.

rationalism regards knowledge, at least scientific knowledge, as unjustified, largely untrue, unbelief. Untrue because merely close to the truth; unbelief because it is largely embodied not in belief but in books and computers.

The scientific method of falsificationism is the father of critical rationalism. It is the generalization to all knowledge, whether empirical or not, of falsificationism, the empirical scientific method. Unlike the Vienna Circle, for example, critical rationalism regards all positions as amenable to rational discussion. Critical rationalism flows from a fallibilist but optimistic attitude. Each of us is infinitely ignorant, we are always prone to error, and we only know little bits of knowledge. I may be wrong and you may be right, and through the cooperative competition of debate, we may get nearer to the truth. We may attain the truth, but not "know" it in the traditional sense, just as one may reach the summit of a mountain (the goal of one's journey) but not know it because one is in fog. For the critical rationalist, getting to the summit is still worth the effort, even with uncertainty.

The issue of confirmation in science goes back to Aristotle, one of the earliest justificationists. Aristotle wished to provide a method for establishing knowledge of nature, demonstrable knowledge, *episteme*, as distinguished from mere opinion, *doxa.* In this Aristotle broke from a long line of philosophers (the pre-Socratics and even Socrates himself) who thought that knowledge was something only the gods could possess.

The method was induction, by which Aristotle meant a way of getting to an outlook from which one could see or "intuit" the essence of the thing in question. This essence would then be embodied in a definition that could be used as a premise in arguments purportedly demonstrating a conclusion about some matter. Aristotle was driven to this route because he was aware that his syllogisms, although valid, did not establish or prove the truth of their conclusion, but only asserted that if the premises were true then the conclusion would also be true. For example, all ravens are black; Edgar is a raven; therefore, Edgar is black. However, if the truth of the premises is not demonstrated, then we require another argument to demonstrate those, and this argument in turn would require its premises to be established, and so on ad infinitum. Aristotle's essential definitions were supposed to provide proven starting points that would avoid such a vicious infinite regress. In this case, if one takes being black as part of the essential definition of a raven, then one may stop at that premises and avoid the infinite regress. Aristotle's view held sway for 2000 years.

Francis Bacon rejected Aristotle's approach to proof. In contrast, In his *Novum Organum* Bacon (1620) argued that one must start, not from intuition, but from observation to discover the laws of hidden forms and natures of matter that explain phenomena. Bacon also called his method "induction." On the standard account, Bacon said that provided we are careful to divest the mind of "idols" or prejudices, a proposed law of nature could be confirmed by the collection of many and varied observations. For example, observing many black ravens and only black ravens leads to the inference that all ravens are black. Every new black raven is then said to confirm the general statement. One could then infer from this observational starting point laws of higher and higher generality and depth, ultimately revealing the hidden nature of things.

Bacon argued that confirming instances and falsifying instances were both feasible and important, but emphasized falsifying instances to safeguard against premature generalizations. Regularities per se may be misleading. We may put this clearer than Bacon did. For example, it is possible that every time you observe a clock, the pendulum is on the right side. The inference that the pendulum is always on the right side would be incorrect. The example may be extended infinitely with the same result. That is, even if you start with an infinite number of correct statements of the pendulum being on the right side, it still does not license the inference that the pendulum is always on the right. This point of logic is independent of whether the pendulum is being observed or not. The regularities of nature may also be misleading, but not in so obvious a way. One can also see in this example why confirming instances may be less valuable than one might at first grant. There is some dispute (see Urbach, 1987) whether it was Bacon who promulgated the induction by numerous, varied observations, but it is nevertheless an important early rudimentary development of the theory of confirmation that deserves comment.

David Hume (1711–76) (1976/1739, book 1, part 3, section 6) raised the problem of induction. He raised a fundamental problem for the inductive confirmation of universal theories of science and the prediction of new phenomena from earlier phenomena. Hume does not keep these two aspects sufficiently apart, but they both make it difficult to uphold induction.

Hume starts with a puzzle. There is only one source of new knowledge: experience. However, our claims to knowledge seem to go far beyond what we can infer from experience. Hume does not use the word "induction," but discusses what has become known as the principle of induction: "that instances of which we have had no experience must resemble those of which we have had experience, and that the course of nature continues always uniformly the same." Hume argued that this goes far beyond our experience. For example, all ravens of which we have had experience have been black. We therefore expect the next one also to be black. However, Hume insists, this is deductively invalid.

Hume pointed out that laws of nature are universal. They speak about the whole of space and time. Thus the statement "All ravens are black" covers everything in the universe that is an raven, all past, present and future ravens, and all non-ravens too if it is understood as the universal conditional: for anything y, if y is a raven, then y is black. A deductively valid argument is one in which if the premises were true then the conclusion must be true. For example, all owls are nocturnal; this bird in my garden is an owl; therefore, this bird in my garden is nocturnal. If we accept the truth of the premises, we are committed, on pain of contradiction, to the truth of the conclusion. However, in the case of an induction, this is not correct, since no matter how many nocturnal owls one has observed, the very next one could be mostly active during the day and sleep at night.

Hume's devastating attack on induction did not stop there. A champion of induction might retreat to the claim that at least an inductive argument makes the conclusion probably true, and that the more positive observations we collect, the more probable our result. However, Hume pointed out that this would only delay the impact of his point, since how are we supposed to confirm this rule of induction? If we say that it is confirmed by its having worked in the past, then we cannot do this without falling into circular reasoning. To rely on induction to justify induction is to argue in a vicious circle. Further, simply postulating a general principle of the uniformity of nature to somehow endorse particular inductions is to abandon the empirical point of view that it is experience alone that ought to be the judge of our theories. However, it is unclear how such a vague postulate can add anything specific to our choice between particular causal hypotheses. Some thinkers have suggested that there is such a principle, but it remains an unanswered question how we are to falsify (or confirm) these principles themselves. For if we do not have any such potential falsifications (or confirmations) how are we to choose between alternative candidates for the job?

Hume's argument is reinforced by the observation that there have been many regularities we have observed without exception for hundreds or even thousands of years that have been rudely interrupted by an eventual exception. Consider the following examples. Newton's theory held its position with many confirmations for about 250 years. However, Einstein's new theory refuted it. For thousands of years it was thought that all life needs light to survive. However, in the last couple of decades we have discovered bacteria—so-called extremophiles—living far below the earth's surface, which live on purely chemical energy.

Bertrand Russell (1985/1918, p. 101) reinforced Hume's argument by pointing out that induction cannot even work in a finite universe. Because even if one had, by hypothesis, observed every relevant particular case in a universe, it is not only an extra step to say that there are no more cases, it is to assert a universal statement, again going beyond all possible finite collections of facts.

Contemporary conceptions of laws of nature make Hume's argument even stronger. Hume conceived of laws of nature as simply unbroken series of contiguous conjunctions of phenomena. For example, consider, under suitable conditions, the scratching of a dry match is always conjoined with the succeeding event of flames. However, recent philosophical conceptions of laws sees them as implying subjunctive conditionals, expressed in the form: if p were to happen, then q would happen. In this case, one such subjunctive conditional would be: "if any dry match were scratched, then there would be a flame." It is clear that this goes beyond all possible observation, because it covers all the actual and possible matches that could have been scratched in the past, but were not, and all possible and actual matches that could, but will not, be scratched now or in the future.

The attraction of Bacon's induction was that it held out the possibility that provided one had divested the mind of what Bacon called "idols" (prejudices, preconceptions, hypotheses) one could see the manifest regularities of natural phenomena. However, the nineteenth-century polymath William Whewell (1794–1866) pointed out that nature does not come ready labeled. We have to classify in advance the types of events or aspects that are relevant to our investigation. Without any such prior conceptual division of the world, we are without a properly defined aim. It would be like telling someone to observe without also telling what to observe and with what aim. For example, if we are trying to understand how planetary objects behave under the influence of gravity, do we observe their shape, mass, density, bulk, or weight? It is clear that even these conceptual classifications are highly theoretical.

In contrast to justificationism, falsificationism escapes Hume's problem of induction. Because conjectures are freely made and then retained until falsified, no induction-like rules are needed. In addition, falsificationism escapes Whewell's strictures because it fully accepts that we start with theories, not with data.

The Paradoxes of Confirmation and Falsification

In comparing confirmation and falsificationism it is helpful to see how each strategy copes with some paradoxes.

The Raven Paradox

Carl Hempel first reminded us of Nicod's seemingly harmless principle of confirmation that the universal $(x)(Ax \rightarrow Cx)$ is confirmed by objects that are A and C. "$(x)(Ax \rightarrow Cx)$" means: for any object x, if it is A, then it is C. Second, Hempel pointed out that the universal theory "All ravens are black" is logically equivalent to "All non-black things are non-ravens." When one is true, the other is true; when one is false the other is false. Hempel then invoked what to the inductivist is the plausible principle that whatever confirms a statement also confirms any logically equivalent statement. However, that would mean that observing a white shoe (a non-black non-raven) would confirm the theory that all ravens are black. This upsets the confirmationist's intuition.

Hempel argued that we rule out such paradoxes by invoking our background knowledge, which has already ruled out shoes as a relevant thing to observe. However, this maneuver only shifts the problem of inductive confirmation back to these background conjectures because it leaves unanswered the question how we are supposed to confirm that background knowledge without falling into a vicious regress? In contrast, as Popper (1977/1934) pointed out, falsification does not have a similar vicious regress. Even a falsifying statement may be tested in its turn. But whereas falsificationism requires that all scientific statements have to be testable, that does not require that they are all tested. One can and has to stop somewhere.

There is no initial paradox afflicting falsificationism because there is nothing odd about saying that when one refutes the theory "All ravens are black" one is also refuting the statement "All non-black things are non-ravens." More importantly, there is no falsificationist counterpart to Nicod's principle. When one refutes a theory, there is no requirement that one also refute every single one of its implications.

The Grue Paradox

Goodman (1955/1979) raised an interesting paradox concerning confirmation with the concept of grue. Grue means "is first examined before 2020 and is green or is examined after 2020 and is blue." Given that definition, it becomes a puzzle why we are prepared to countenance the inference from the fact that all observed emeralds have been green to "All emeralds are green" but not to "All emeralds are grue," since every observed green emerald up to 2020 confirms both universal gem theories.

What vexes inductive philosophers is that inferences about grue seem just as valid as inferences about green by a supposed inductive logic. Goodman argued that the reason we choose green over grue is that the former concept is more projectible, meaning that it is more entrenched in our linguistic habits.

One falsificationist answer to this conundrum is to say that scientific method requires the comparison of different hypotheses that we can subject to a decisive experiment. Grue-like hypotheses do not have this characteristic and so should not be admitted to the body of scientific theories under test. Some inductivists (e.g. Worrall, 1989) say that this approach makes the choice of green over grue "arbitrary," even an "historical accident." However, as Quine (1977) explains, the evolution of our conjectures about gems, minerals, and chemistry over thousands of years is far from arbitrary. Another falsificationist response is to say that the grue hypothesis does not solve any problem that our familiar green emerald theory cannot solve just as well (see Bartley, 1968).

The Asymmetry of Falsification and Confirmation

Popper (1977/1934, section 6, 1983, section 22) argued that there is a fundamental logical asymmetry between falsification and verification which follows from the logical form of scientific theories, which are universal statements. The statement "The raven in my room now is white" is fully decidable in principle because we can either verify it or falsify it. This is typical of our test statements. But our theories are only partially decidable, because even though we can falsify them, we cannot verify them.

Given a universal statement and initial conditions, predicted specific observations can be validly inferred. If these observations fail to occur, it follows that the theory is false. This valid inference is called *modus tollens.* An example of the valid *modus tollens* is: If someone has a cold, then they'll sneeze; John is not sneezing; therefore, John does not have a cold. However, if the predicted observations do occur, inferring that the theory must therefore be true would involve the fallacy of *affirming the consequent*, but that is what confirmation seems to require. An example of the invalid *affirming the consequent* is: If someone has a cold, then they'll sneeze; John is sneezing; therefore, John has a cold. (People sneeze for other reasons too.) The asymmetry is that falsifying observations allow us to draw valid conclusions about the theory under test, but the confirming observations do not.

One objection to this holds that the asymmetry is an illusion, because whenever we refute a universal statement we thereby verify its negation. A universal statement "All x are y" is equivalent to "There is no non-y x." Therefore, when we refute "All apples are green" we automatically verify "There is a non-green apple." We can, therefore, equally speak about verification instead of falsification and so the two are perfectly symmetrical. Popper (1977/1934) had already covered the logical equivalence of nonexistence statements and universal laws, and replied that this misses the point, rather as if someone had said because positive and negative numbers have many symmetrical properties, therefore they are perfectly symmetrical.

Duhem/Quine Problem

Duhem (1991/1914) and Quine (1961/1951) each raised a problem with how one assigns the fault of a falsification. They argued that in a scientific test of a theory, one must first work out a testable implication of this theory, an observable event that ought to present itself if the theory is true. One uses a valid argument to tease out the expected observation statement from the theory. However, to work this implication out one typically needs a host of other auxiliary assumptions regarding equipment, initial conditions, and background theories. So the question is, when the expected event fails to appear, which of the premises is at fault—the main theory by itself, and/or one (or more) of the other assumptions used in the deduction? The fact that our argument is valid only tells us that at least one of the premises is at fault. For example, if an electric door opener is installed, when you walk up to the door it will open. If the door does not open, that may mean that a door opener is not installed. But we are assuming that the power is on, that the door opener is hooked up properly, that the door is not locked, that we are not dreaming, etc. Duhem and Quine said that, without appeal to other principles such as simplicity, this is undecidable.

Duhem and Quine's positions are different in some respects. Whereas Duhem confines the group of problematic auxiliary assumptions to physics, Quine extends the group to the whole of human knowledge and also sees it as applying to both verification and falsification. For Quine, in the face of recalcitrant observations, one may even consider abandoning logical or mathematical assumptions such as the law of the excluded middle (that a statement is either true or false, with no third option). Some interpreters of quantum theory suggest that this law does not apply at the quantum level.

The falsificationist's answer is that fallibility prevails even in the interpretation of test results. We may be wrong in assigning the fault of a refutation. However, this does not mean that we cannot successfully classify statements as true, as opposed to certifying them. One ought to choose auxiliary assumptions that are unproblematic, and it is only a rampant undiscriminating skepticism that could rank everything as problematic. Falsificationists assume that scientists—though fallible—can at least sometimes classify trivial observation reports as true, for example, correctly reading an on/off light. If someone conjectures that some subset of the auxiliary assumptions is problematic, then they are free to set up an experiment to test that set using what they regard as unproblematic assumptions.

Probability and Inductive Confirmation

Having surrendered to Hume's attack, most justificationists have sought solace in the hope that induction might at least provide some theories that are more probable and reliable than others. The most popular contenders for some time for the title of probabilistic induction are the Bayesian interpretations of probability. To evaluate the probability of a hypothesis, the Bayesian specifies some prior probability for a given hypothesis, which is then updated by new information. These are subjective degrees of probability ranging from 0 (no belief) to 1 (certainty). The Bayesian interpretation provides a standard set of procedures and formulas to perform this calculation. One of its strong points is that it conforms to the dominant axiom system for probability, which was developed by Kolmogorov.

The formula used to calculate the probability of a hypothesis on the evidence is read: the probability of the hypothesis H given the evidence E equals the probability of the evidence given the hypothesis multiplied by the probability of the hypothesis taken alone divided by the probability of the evidence taken alone.

$$P(\mathrm{H}/\mathrm{E}) = \frac{P(\mathrm{E}/\mathrm{H})P(\mathrm{H})}{P(\mathrm{E})}$$

The first point to note here is that prior probabilities have to be entered into the formula before any evidence is available. However, it is generally accepted that there is no non-arbitrary way to choose these prior probabilities. More fundamental is one of Popper's criticisms. If we search for the most probable theories, then we shall abandon the most informative theories of science because there is an inverse relationship between probability and information content. Glymour (1980) pointed out that it is unclear how to connect Bayesianism with the history of science. Neither Copernicus, Newton, nor Kepler—to name but a few of the greats—gave probabilistic arguments for their theories.

Another criticism, based on an evolutionary theory of belief, is that at any given moment, we cannot decide what we believe. We can research an issue more or less thoroughly, but we cannot voluntarily adopt, reject, or change the degree to which we adhere to our beliefs (Percival, 2012). Evidence and argument have their impact on belief, and the soundness of an argument enhances its persuasiveness. However, when a telling argument has its effect, it is as if the mind had been infected by it, rather like a catchy tune. Belief is not something that we do. We cannot therefore obey the prescriptions that issue from the Bayesian evaluations of data. In contrast, falsificationism instructs one to perform the fallible but feasible task of classifying basic observation statements as true or false. This is something we can do. Imagine putting true and false statements into separate boxes. Belief is also a capricious individual state of mind, spontaneously varying from moment to moment and with changing risks and outcomes. On the other hand, since statements, as opposed to beliefs, are permanent records independent of our psychology, falsificationism is not troubled by the vagaries of a scientist's momentary state of mind. Therefore falsificationism is a far more stable and publicly testable method than assigning degrees of probability to subjective belief. It also harmonizes better with Popper's point that the overwhelming amount of scientific knowledge is and must be embodied in libraries and computers, not in scientists' minds.

SEE ALSO: Bayesian Analysis; Duhem–Quine Thesis; Evidence-Based Assessment; Hypothetico-Deductive Model; Laws of Nature; Natural Kinds; Popper, Karl (1902–94); Positivism and Logical Positivism

References

Bacon, F. (1620). *Novum organum.* London: Billium.

Bartley, W. W. (1968). Goodman's paradox: A simple minded solution. *Philosophical Studies,* December, 85–88.

Duhem, P. (1991/1914). *The aim and structure of physical theory.* Princeton, NJ: Princeton University Press.

Glymour, C. (1980). *Theory and evidence.* Princeton, NJ: Princeton University Press.

Goodman, N. (1955/1979). *Fact, fiction and forecast* (3rd ed.). Brighton, U.K.: Harvester Press.

Hume, D. (1976/1739). *A treatise of human nature.* Oxford, U.K.: Oxford University Press.

Percival, R. S. (2012). *Myth of the closed mind: Explaining why and how people are rational.* Chicago, IL: Open Court.

Popper, K. (1977/1934). *The logic of scientific discovery.* London: Hutchinson.

Popper, K. (1983). *Realism and the aim of science.* W. W. Bartley, III (Ed.). London: Hutchinson.

Russell, B. (1985/1918). *The philosophy of logical atomism.* David Pears (Ed.). Chicago, IL: Open Court.

Quine, W. Van O. (1961/1951). Two dogmas of empiricism. In *From a logical point of view.* New York: Harper Torchbooks.

Quine, W. Van O. (1977). Natural kinds. In *Naming, necessity and natural kinds* (chapter 6). Ithaca, NY: Cornell University Press.

Urbach, P. (1987). *Francis Bacon's philosophy of science.* Chicago, IL: Open Court.

Worrall, J. (1989). Why both Popper and Watkins fail to solve the problem of induction. In F. D'Agostino & I. C. Jarvie (Eds.), *Freedom and rationality: Essays in honor of John Watkins.* Dordrecht, The Netherlands: Kluwer.

Further Reading

Bartley, W. W. (1984). *The retreat to commitment*. Chicago, IL: Open Court.

Fetzer, J. (2002). Propensities and frequencies: Inference to the best explanation. *Syntheses, 132,* 27–61. Reprinted in Eells, E., & Fetzer, J. (Eds.). (2010). *The place of probability in science* (pp. 323–351). New York: Springer.

Harman, G. (1965). The inference to the best explanation. *The Philosophical Review, 74,* 88–95.

Howson, C., & Urbach, P. (1989). *Scientific reasoning: The Bayesian approach.* Chicago, IL: Open Court.

Karl Popper Forums and Dublin City University. *The Karl Popper web*. www.tkpw.net

Miller, D. (1994). *Critical rationalism: A restatement and defence.* Chicago, IL: Open Court.

Miller, D. (2006). *Out of error* (illustrated ed.). London: Ashgate Publishing.

Notturno, M. (2003). *On Popper* (especially chapters 1, 2, and 3). Melbourne, Australia: Thomas Wadsworth.

Popper, K., & Miller, D. (1983). A proof of the impossibility of inductive probability. *Nature, 302,* 687–688.

Please note that the abstract and keywords will not be included in the printed book, but are required for the online presentation of this book which will be published on Wiley's own online publishing platform.

If the abstract and keywords are not present below, please take this opportunity to add them now.
The abstract should be a short paragraph upto 200 words in length and keywords between 5 to 10 words.

ABSTRACT

Confirmation and falsification are different strategies for testing theories and characterizing the outcomes of those tests. Roughly speaking, confirmation is the act of using evidence or reason to verify or certify that a statement is true, definite, or approximately true, whereas falsification is the act of classifying a statement as false in the light of observation reports. This simple picture hides profound problems in the methodology of science and epistemology. In reaction to the skeptic's attack on the possibility of certain or certified truth, many writers have opted for what they regard as the next best feasible substitute: probable, reliable, or consistent theories. Within the philosophy of science there is a long-running debate about the meaning, relative importance, and feasibility of confirmation and falsification and their role in the advancement of science. Although not confined to the issue of induction in science, the debate centers on this issue.

KEYWORDS

Aristotle; critical rationalism; falsificationism; Francis Bacon; empiricism; evolution; induction; positivism