

Falsifiability

Much discussion of the empirical status of social science in the past has revolved around Karl Popper's formulation of the doctrine of falsifiability. This idea has had a particularly noticeable influence on discussions of methodology in the social sciences. Popper's requirement is that all scientific hypotheses must in principle be falsifiable: that is, it must be possible to specify in advance a set of empirical circumstances which would demonstrate the falsity of the hypothesis. Popper writes, "A theory which is not refutable by any conceivable event is non-scientific. Irrefutability is not a virtue of a theory (as people often think) but a vice" (Popper, 1965, p. 36). This criterion is often used to fault social science research on the ground that social scientists are often prepared to adjust their hypotheses in such a way as to render them compatible with unexpected empirical results (anomalies). Popper develops these criticisms in some detail in his discussion of Marx's social theory. Popper writes,

The Marxist theory of history . . . ultimately adopted this soothsaying practice. In some of its earlier formulations . . . their predictions were testable, and in fact falsified. Yet instead of accepting the refutations the followers of Marx re-interpreted both the theory and the evidence in order to make them agree. In this way they rescued the theory from refutation; but they did so at the price of adopting a device which made it irrefutable. They thus gave a "conventionalist twist" to the theory; and by this strategem they destroyed its much advertised claim to scientific status. (Popper, 1965, p. 32)

Popper's charge of unfalsifiability finds its strongest ground in Marx's willingness to modify his hypotheses in order to save them from direct empirical refutation--for example, in his use of the idea countervailing tendencies to account for the mixed record of the rate of profit over time. Marx's economics predicts that there will be a falling rate of profit within capitalist economies over time. Marx notes, however, that this is often not the case, and attempts to account for this failure by referring to countervailing tendencies: causal factors that work to prop up the tendency for the rate of profit to fall over time. (A countervailing tendency is a previously unknown factor that is hypothesized in order to account for discrepancies between theoretical expectations and observed facts.) Popper believes that appeal to such tendencies is itself a conventionalist twist that deprives the theory of empirical content. But is it scientifically irrational to appeal to countervailing tendencies?

Popper's attack on countervailing tendencies derives from the fact that it is always possible to save a theory from false consequences by referring to factors not accounted for in the theory. But he believes that such appeals reduce or eliminate the empirical content of the theory. If the scientist is prepared to make some such appeal in every anomalous case, does he or she not relinquish claim to having provided an empirically significant hypothesis? Is it not reasonable in such a case to conclude that the theory is unfalsifiable by stipulation, and therefore devoid of empirical content? This conclusion would be justified only if it were impossible to impose limits on the appeal to such tendencies--only, that is, if it were impossible to show how to distinguish between ad hoc and progressive modifications of the theory. In order to evaluate

these charges, however, it will be necessary to consider the issues of anomaly and theory change in greater detail.

Popper's falsifiability thesis arises in response to the general problem of anomaly in science. Anomalies--facts or discoveries that appear inconsistent with accepted theory--are found everywhere in the history of science, since scientific inquiry is inherently fallible. If a theory implies some sentence *S* and *S* is false, it follows that the theory must be false as well. In such a case the scientist is faced with a range of choices. He or she can reject the theory as a whole; reject some portion of the theory in order to avoid the conclusion *S*; modify the theory to avoid the conclusion *S*; or introduce some additional assumption to show how the theory is consistent with "not *S*." A strict falsificationist would presumably require that we disavow the theory, but this response is both insensitive to actual scientific practice and implausible as a principle of methodology.

When faced with anomaly, the scientist must choose whether to abandon the theory altogether or modify it to make it consistent with the contrary observations. If the theory has a wide range of supporting evidence (aside from the contrary experience), there is a powerful incentive in favor of salvaging the theory, that is, of supplementing it with some further principle restricting the application of its laws, or modifying the laws themselves, to reconcile theory with experience. Ideally the scientist ought to proceed by attempting to locate the source of error in the original theory. Theory modification in the face of contrary evidence should result in a more realistic description of the world, either through the correction of false theoretical principles or through the description of further factors at work that were hitherto unrecognized.

It is possible, however, to modify a theory in ways that do not reflect any additional insight into the real nature of the phenomena in question, but are rather merely mechanical modifications of the theory made to bring it into line with the contrary evidence. Such modifications are common in the history of science; Carl Hempel cites the example of phlogiston theorists under attack by Lavoisier. After Lavoisier's discovery that metals weighed more after combustion than prior (thereby apparently falsifying phlogiston theory), some proponents of that hypothesis modified their concept of phlogiston by assuming that it possessed negative weight. This alteration reconciled the phlogiston theory with Lavoisier's evidence; nevertheless it seems on fairly intuitive grounds to be an illegitimate modification. It is "introduced ad hoc--i.e., for the sole purpose of saving a hypothesis seriously threatened by adverse evidence; it would not be called for by other findings, and, roughly speaking, it leads to no additional test implications" (Hempel, 1966, pp. 28-30).

The problem of avoiding adhocness by devising a set of methodological standards suitable for governing the modification of hypotheses in light of contrary evidence is a substantial one. As Hempel observes, the clearest judgments of adhocness are made with the benefit of hindsight; what may have been a rational modification given current beliefs is with the benefit of later knowledge a transparent case of ad hoc modification. However, we may advance a rough set of guidelines for the introduction of modifications: "Is the hypothesis proposed just for the purpose of saving some current conception against adverse evidence, or does it also account for other phenomena, does it yield further significant test implications" (Hempel, 1966, p. 30)? Does it contribute to a theory that affords simple explanations of a wide range of phenomena? Does it appear to represent an increased knowledge of the real mechanisms that underlie

observable phenomena? Does it merely repeat the evidence already available, or is it amenable to independent tests (Popper, 1965, p. 288)? These considerations fall far short of a definition of adhocness, and recent work in the philosophy of science has substantially extended these ideas by introducing the notion of a research program.¹

Postpositivist philosophy of science has directed much of its efforts to formulating more adequate standards for modifying theory in the light of anomaly. Its chief insights have resulted from a shift of attention from the level of finished theories to the level of the research program, that is, from the formal laws and principles of a theory to the more encompassing set of presuppositions, methodological commitments, and research interests that guide scientists in the conduct of research and theory formation. The central focus of neopositivist theory of science was the scientific theory, conceived ideally as a formal system of axioms and deductive consequences. Neopositivists distinguished between the context of discovery and that of justification, and they argued that only the latter fell within the scope of rational control. This meant that only finished theories could be rationally evaluated, whereas the conduct of research was conceived of as an exercise of pure, unregulated imagination (Popper, 1968, p. 31). From this judgment followed falsificationism, verificationism, and various forms of confirmation theory.

The "new" philosophy of science focuses on the "context of discovery"--the assumptions and research goals that guide scientists in their research. Philosophers of science in this area reject the idea that the conduct of research is an unstructured, nonrational process, and they have tried to formulate a theory of the rules that distinguish good scientific research from bad. From this starting point, the "research program" becomes the central interest.²

What is a research program? It is the framework of assumptions, experimental procedures, explanatory paradigms, and theoretical principles that guide the conduct of research. Lakatos has provided an especially clear formulation of the concept. First, he argues that any research program possesses a "hard core" of theoretical principles that constitute its central insight into its subject matter. This core is taken as fixed; the "negative heuristic" of the program forbids the scientist from interpreting anomalous results as falsifying this core. Instead, the scientist is directed to construct a "protective belt" of auxiliary assumptions intended to secure the correctness of the theoretical principles at the core, and scientific research becomes an effort to modify or replace the assumptions included in the belt so as to make the core consistent with experimental results. Finally, the research program includes what Lakatos calls a "positive heuristic": a set of principles and assumptions that provide guidance in extending and developing the belt. This conception of scientific inquiry could be summarized in the form of a slogan: Defend and extend! Built into the view is a rejection of falsificationism, for, far from seeking to refute the central theoretical principles, the scientist is directed to defend and extend them as forcefully as possible.

With this fundamentally different starting point, the new philosophers of science have posed a different question for themselves. Rather than the positivists' question--What is the

¹ Larry Laudan's *Progress and Its Problems* (1977) provides a good critical summary of the views of philosophers of science who give pride of place to the idea of a research program or research tradition.

² See particularly Imre Lakatos, "Falsification and the Methodology of Scientific Research Programmes" (1970a).

criterion of an empirically adequate theory?--they have asked, What are the features that distinguish a rational and progressive program of research from its contrary? The problem of theory adequacy does not disappear, but it becomes a subordinate concern. This broader approach to empirical rationality lays the emphasis on the degree to which the commitments of the research program successfully direct research productively and suggest empirically adequate theories--rather than on the narrower question of the criterion of empirical adequacy of theories. On this view, empirical rationality is a feature of the program of research rather than the finished theory; theories are tools for understanding empirical phenomena created by the scientist within the context of a framework of methodological and substantive assumptions.

From this research-oriented point of view, falsificationism is an unsound principle of theory choice, since it is an extreme principle that requires the rejection of any theory with false consequences. A more conservative strategy is required, one that allows the scientist to preserve the old theory at minimum cost. On this view, it is generally a reasonable methodological principle to try to formulate a hypothesis that would account for the truth of a theory and the falsehood of one of its consequences S --either by supposing S is really true (i.e., experimental error) or by modifying the theory or by positing some unobserved factor that, together with the theory, predicts "not S ". (Consider, for example, the anomalies arising in the Newtonian description of the planetary orbits that led to the subsequent positing of Uranus.) It is reasonable, that is, to take as a research strategy the maxim of least harm: to try to produce a reconciliation of theory and observation that requires the least change in the theory. And the general success of scientific theory formation guided by this maxim vindicates the strategy.

The problem with the principle of least harm is that it allows us to stave off rejection of the theory indefinitely; it potentially makes the theory irrefutable. Once we widen our vision from theories to research programs, however, we find that the key problem is not how to keep a theory falsifiable but rather how to impose a set of rational constraints on the principle of least harm: how to avoid ad hoc modifications of theory that fail to advance the theory's empirical power and explanatory adequacy.

Lakatos has discussed this question in detail. His account is not altogether adequate, but it gives an indication of the sort of criterion of adequacy that seems to have some promise of success. On his view, the problem is how to define the notion of a "progressive problem shift," that is, a modification of theory in light of conflicting experience that improves the empirical adequacy of the theory. Lakatos gives a twofold criterion of progressiveness. A modification of theory is theoretically progressive if the modification has some excess empirical content over its predecessor, and it is empirically progressive if some of this content is corroborated. If the change is not progressive in these senses, the research tradition is in a state of degeneration and ought to be replaced (Lakatos, 1970a, pp. 116-122).

This conception of a progressive research tradition may be amplified into a more specific criterion of rational adherence to an empirical theory in the face of anomaly. First, the theory in question must have achieved some empirical success. That is, it must produce empirically adequate explanations of phenomena in areas other than those affected by anomaly; otherwise it would be irrational to remain committed to the theory. And second, the modifications of the theory must themselves be, at least potentially, empirically significant. (1) They must give rise to other consequences besides the range of phenomena they were

introduced to explain, and (2) they must be amenable to further investigation. If these conditions obtain, and if independent justification is produced for the new factors, both they and the earlier theory are vindicated.

These considerations show, then, that the core value of empirical standards does not entail commitment to the doctrine of falsifiability in its narrow form: it is not scientifically irrational to modify hypotheses so as to render them compatible with empirical observations. The crucial question is not falsifiability, but rather whether the research program continues to broaden its empirical scope. It is reasonable for scientists to modify their hypotheses and theories in light of anomalous findings; the crucial empirical constraint is that the modified theory ought to have additional empirical scope. The standard of strict falsifiability, then, is not a reasonable constraint on hypothesis formation in scientific research.