Parapsychology is defined by its practitioners as the study of extrasensory perception (ESP) and paranormal powers such as telekinesis. ESP includes such alleged psychic phenomena as telepathy, clairvoyance, and precognition. Shunned for decades by the scientific establishment, parapsychologists received official recognition in 1969 when the American Association for the Advancement of Science (the AAAS) admitted the Parapsychological Association as an affiliate member. Many scientists are unhappy with this decision, since they regard parapsychology as a pseudoscience. In 1979, the renowned physicist John A. Wheeler wrote a blistering letter to the president of the AAAS urging that the parapsychologists be expelled from the association. Wheeler wrote, "We have enough charlatanism in this country today without needing a scientific organization to prostitute itself to it. The AAAS has to make up its mind whether it is seeking popularity or whether it is strictly a scientific organization."

The debate about the nature of science—about its scope, methods, and aims—is as old as science itself. But this debate becomes especially heated when one group of practitioners accuses another group of practicing pseudoscience. In the twentieth century many individuals, groups, and theories have been accused of being pseudoscientific, including Freud and psychoanalysis, astrology, believers in the paranormal, Immanuel Velikovsky and Erich von Daniken (whose best-selling books Worlds in Collision and Chariots of the Gods excited the wrath of Carl Sagan and the scientific establishment), and, most recently, the self-styled advocates of creation-science. The proponents of astrology, the paranormal, psychoanalysis, and creation-science engage in research, write books, and publish articles, but their work is typically found in popular magazines and bookstores rather than refereed journals and science libraries. They are seldom funded by
the National Science Foundation or elected to the National Academy of Sciences. They are outside of the scientific establishment and are kept out by those who regard themselves as real scientists.

If our only concern were to label certain people "pseudoscientists," we might simply check where their work is published and how their theories have been received by the scientific community. But we are concerned with the reasons certain doctrines are considered pseudoscientific; it is those reasons that interest philosophers of science.

Some philosophers have proposed necessary conditions for genuine science. That is, they have offered characteristics that any discipline or field of study must possess in order to qualify as genuine science. These characteristics are often called demarcation criteria because they can be used to differentiate science from its counterfeit: if a discipline fails to meet one of these conditions, then it is judged to be nonscientific.

In the twentieth century, philosophers of science have often disagreed about demarcation criteria. In this chapter Karl Popper, Thomas Kuhn, Imre Lakatos, and Paul Thagard each defend a different set of necessary conditions for genuine science. Popper's view, that a scientific theory must be open to refutation by making testable predictions, has been very influential, especially among working scientists. Kuhn, Lakatos, and Thagard all reject Popper's claim that falsifiability is the hallmark of genuine science but disagree about what should replace it. All three address whether a theory or discipline's claim to scientific legitimacy depends on historical considerations, such as how theories have developed over time.

The chapter ends with an exchange of views between Michael Ruse and Larry Laudan about the credentials of creation-science. Ruse, a prominent philosopher of biology, served as an expert witness in a trial concerning the constitutionality of an Arkansas law requiring public school biology teachers to present creationism as a viable scientific alternative to evolutionary theory. Under Ruse's guidance, the judge in the case drew up a list of five criteria for genuine science and concluded that creation-science failed on all five counts. Laudan not only criticizes the items on this list (which includes Popper's falsifiability) but also doubts whether there are any demarcation criteria that all scientific theories must satisfy.

### Notes

Mr. Turnbull had predicted evil consequences, ... and was now doing the best in his power to bring about the verification of his own prophecies.

— Anthony Trollope

When I received the list of participants in this course and realized that I had been asked to speak to philosophical colleagues* I thought, after some hesitation and consultation, that you would probably prefer me to speak about those problems which interest me most, and about those developments with which I am most intimately acquainted. I therefore decided to do what I have never done before: to give you a report on my own work in the philosophy of science, since the autumn of 1919 when I first began to grapple with the problem, 'When should a theory be ranked as scientific?' or 'Is there a criterion for the scientific character or status of a theory?'

The problem which troubled me at the time was neither, 'When is a theory true?' nor, 'When is a theory acceptable?' My problem was different. I wished to distinguish between science and pseudo-science; knowing very well that science often errs, and that pseudo-science may happen to stumble on the truth.

I knew, of course, the most widely accepted answer to my problem:


* This essay was originally presented as a lecture at Peterhouse College at Cambridge University in the summer of 1953 as part of a course on developments and trends in contemporary British philosophy, organized by the British Council. It was originally published as "Philosophy of Science: A Personal Report," in British Philosophy in Mid-Century, ed. C. A. Mace, (London: Allen and Unwin, 1957).
that science is distinguished from pseudo-science—or from 'metaphysics'—by its empirical method, which is essentially inductive, proceeding from observation or experiment. But this did not satisfy me. On the contrary, I often formulated my problem as one of distinguishing between a genuinely empirical method and a non-empirical or even a pseudo-empirical method—that is to say, a method which, although it appeals to observation and experiment, nevertheless does not come up to scientific standards. The latter method may be exemplified by astrology, with its stupendous mass of empirical evidence based on observation—on horoscopes and on biographies.

But as it was not the example of astrology which led me to my problem, I should perhaps briefly describe the atmosphere in which my problem arose and the examples by which it was stimulated. After the collapse of the Austrian Empire there had been a revolution in Austria: the air was full of revolutionary slogans and ideas, and new and often wild theories. Among the theories which interested me Einstein's theory of relativity was no doubt by far the most important. Three others were Marx's theory of history, Freud's psycho-analysis, and Alfred Adler's so-called 'individual psychology'.

There was a lot of popular nonsense talked about these theories, and especially about relativity (as still happens even today), but I was fortunate in those who introduced me to the study of this theory. We all—the small circle of students to which I belonged—were thrilled with the result of Eddington's eclipse observations which in 1919 brought the first important confirmation of Einstein's theory of gravitation. It was a great experience for us, and one which had a lasting influence on my intellectual development.

The three other theories I have mentioned were also widely discussed among students at that time. I myself happened to come into personal contact with Alfred Adler, and even to co-operate with him in his social


†Einstein's general theory of relativity entails that light rays must bend in a gravitational field. Organized by Sir Arthur Eddington, two Royal Astronomical Society expeditions were dispatched to observe the solar eclipse of 1919, and verified that starlight was indeed deflected by the sun by the amount that Einstein had predicted. The Times of London reported this success as the most remarkable scientific event since the discovery of the planet Neptune. The light-bending test of relativity theory is discussed in "Popper's Demarcation Criterion," in the commentary on chapter 1, and in "Two Arguments for Explanationism," in the commentary on chapter 4.
work among the children and young people in the working-class districts of Vienna where he had established social guidance clinics.

It was during the summer of 1919 that I began to feel more and more dissatisfied with these three theories—the Marxist theory of history, psycho-analysis, and individual psychology; and I began to feel dubious about their claims to scientific status. My problem perhaps first took the simple form, 'What is wrong with Marxism, psycho-analysis, and individual psychology? Why are they so different from physical theories, from Newton's theory, and especially from the theory of relativity?'

To make this contrast clear I should explain that few of us at the time would have said that we believed in the truth of Einstein's theory of gravitation. This shows that it was not my doubting the truth of those other three theories which bothered me, but something else. Yet neither was it that I merely felt mathematical physics to be more exact than the sociological or psychological type of theory. Thus what worried me was neither the problem of truth, at that stage at least, nor the problem of exactness or measurability. It was rather that I felt that these other three theories, though posing as sciences, had in fact more in common with primitive myths than with science; that they resembled astrology rather than astronomy.

I found that those of my friends who were admirers of Marx, Freud, and Adler, were impressed by a number of points common to these theories, and especially by their apparent explanatory power. These theories appeared to be able to explain practically everything that happened within the fields to which they referred. The study of any of them seemed to have the effect of an intellectual conversion or revelation, opening your eyes to a new truth hidden from those not yet initiated. Once your eyes were thus opened you saw confirming instances everywhere: the world was full of verifications of the theory. Whatever happened always confirmed it. Thus its truth appeared manifest; and unbelievers were clearly people who did not want to see the manifest truth; who refused to see it, either because it was against their class interest, or because of their repressions which were still 'un-analysed' and crying aloud for treatment.

The most characteristic element in this situation seemed to me the incessant stream of confirmations, of observations which 'verified' the theories in question; and this point was constantly emphasized by their adherents. A Marxist could not open a newspaper without finding on every page confirming evidence for his interpretation of history, not only in the news, but also in its presentation—which revealed the class bias of the paper—and especially of course in what the paper did not say. The Freudian analysts emphasized that their theories were constantly verified by their 'clinical observations'. As for Adler, I was much impressed by a personal experience. Once, in 1919, I reported to him a case which to me did not seem particularly Adlerian, but which he found no difficulty in analysing in terms of his theory of inferiority feelings, although he had not even
seen the child. Slightly shocked, I asked him how he could be so sure. 'Because of my thousandfold experience,' he replied; whereupon I could not help saying: 'And with this new case, I suppose, your experience has become thousand-and-one-fold.'

What I had in mind was that his previous observations may not have been much sounder than this new one; that each in its turn had been interpreted in the light of 'previous experience', and at the same time counted as additional confirmation. What, I asked myself, did it confirm? No more than that a case could be interpreted in the light of the theory. But this meant very little, I reflected, since every conceivable case could be interpreted in the light of Adler's theory, or equally of Freud's. I may illustrate this by two very different examples of human behaviour: that of a man who pushes a child into the water with the intention of drowning it; and that of a man who sacrifices his life in an attempt to save the child. Each of these two cases can be explained with equal ease in Freudian and in Adlerian terms. According to Freud the first man suffered from repression (say, of some component of his Oedipus complex), while the second man had achieved sublimation. According to Adler the first man suffered from feelings of inferiority (producing perhaps the need to prove to himself that he dared to commit some crime), and so did the second man (whose need was to prove to himself that he dared to rescue the child). I could not think of any human behaviour which could not be interpreted in terms of either theory. It was precisely this fact—that they always fitted, that they were always confirmed—which in the eyes of their admirers constituted the strongest argument in favour of these theories. It began to dawn on me that this apparent strength was in fact their weakness.

With Einstein's theory the situation was strikingly different. Take one typical instance—Einstein's prediction, just then confirmed by the findings of Eddington's expedition. Einstein's gravitational theory had led to the result that light must be attracted by heavy bodies (such as the sun), precisely as material bodies were attracted. As a consequence it could be calculated that light from a distant fixed star whose apparent position was close to the sun would reach the earth from such a direction that the star would seem to be slightly shifted away from the sun; or, in other words, that stars close to the sun would look as if they had moved a little away from the sun, and from one another. This is a thing which cannot normally be observed since such stars are rendered invisible in daytime by the sun's overwhelming brightness; but during an eclipse it is possible to take photographs of them. If the same constellation is photographed at night one can measure the distances on the two photographs, and check the predicted effect.

Now the impressive thing about this case is the risk involved in a prediction of this kind. If observation shows that the predicted effect is definitely absent, then the theory is simply refuted. The theory is incompatible with certain possible results of observation—in fact with results
which everybody before Einstein would have expected. This is quite different from the situation I have previously described, when it turned out that the theories in question were compatible with the most divergent human behaviour, so that it was practically impossible to describe any human behaviour that might not be claimed to be a verification of these theories.

These considerations led me in the winter of 1919–20 to conclusions which I may now reformulate as follows.

1. It is easy to obtain confirmations, or verifications, for nearly every theory—if we look for confirmations.

2. Confirmations should count only if they are the result of risky predictions; that is to say, if, unenlightened by the theory in question, we should have expected an event which was incompatible with the theory—an event which would have refuted the theory.

3. Every 'good' scientific theory is a prohibition: it forbids certain things to happen. The more a theory forbids, the better it is.

4. A theory which is not refutable by any conceivable event is nonscientific. Irrefutability is not a virtue of a theory (as people often think) but a vice.

5. Every genuine test of a theory is an attempt to falsify it, or to refute it. Testability is falsifiability, but there are degrees of testability: some theories are more testable, more exposed to refutation, than others; they take, as it were, greater risks.

6. Confirming evidence should not count except when it is the result of a genuine test of the theory; and this means that it can be presented as a serious but unsuccessful attempt to falsify the theory. (I now speak in such cases of 'corroborating evidence'.)

7. Some genuinely testable theories, when found to be false, are still upheld by their admirers—for example by introducing ad hoc some auxiliary assumption, or by re-interpreting the theory ad hoc in such a way that it escapes refutation. Such a procedure is always possible, but it rescues the theory from refutation only at the price of destroying, or at least lowering, its scientific status. (I later described such a rescuing operation as a 'conventionalist twist' or a 'conventionalist stratagem'.)

One can sum up all this by saying that the criterion of the scientific status of a theory is its falsifiability, or refutability, or testability.
us to pronounce on the results of the tests with complete assurance, there was clearly a possibility of refuting the theory.

Astrology did not pass the test. Astrologers were greatly impressed, and misled, by what they believed to be confirming evidence—so much so that they were quite unimpressed by any unfavourable evidence. Moreover, by making their interpretations and prophecies sufficiently vague they were able to explain away anything that might have been a refutation of the theory had the theory and the prophecies been more precise. In order to escape falsification they destroyed the testability of their theory. It is a typical soothsayer's trick to predict things so vaguely that the predictions can hardly fail: that they become irrefutable.

The Marxist theory of history, in spite of the serious efforts of some of its founders and followers, ultimately adopted this soothsaying practice. In some of its earlier formulations (for example in Marx's analysis of the character of the 'coming social revolution') 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 stratagem they destroyed its much advertised claim to scientific status.

The two psycho-analytic theories were in a different class. They were simply non-testable, irrefutable. There was no conceivable human behaviour which could contradict them. This does not mean that Freud and Adler were not seeing certain things correctly: I personally do not doubt that much of what they say is of considerable importance, and may well play its part one day in a psychological science which is testable. But it does mean that those 'clinical observations' which analysts naively believe confirm their theory cannot do this any more than the daily confirmations which astrologers find in their practice. And as for Freud's epic of the Ego, the Super-ego, and the Id, no substantially stronger claim to scientific status can be made for it than for Homer's collected stories from Olympus. These theories describe some facts, but in the manner of myths. They contain most interesting psychological suggestions, but not in a testable form.

At the same time I realized that such myths may be developed, and become testable; that historically speaking all—or very nearly all—scientific theories originate from myths, and that a myth may contain important anticipations of scientific theories. Examples are Empedocles' theory of evolution by trial and error, or Parmenides' myth of the unchanging block universe in which nothing ever happens and which, if we add another dimension, becomes Einstein's block universe (in which, too, nothing ever happens, since everything is, four-dimensionally speaking, determined and laid down from the beginning). I thus felt that if a theory is found to be non-scientific, or 'metaphysical' (as we might say), it is not thereby found
to be unimportant, or insignificant, or 'meaningless', or 'nonsensical'. But it cannot claim to be backed by empirical evidence in the scientific sense—although it may easily be, in some genetic sense, the 'result of observation'.

(There were a great many other theories of this pre-scientific or pseudo-scientific character, some of them, unfortunately, as influential as the Marxist interpretation of history; for example, the racialist interpretation of history—another of those impressive and all-explanatory theories which act upon weak minds like revelations.)

Thus the problem which I tried to solve by proposing the criterion of falsifiability was neither a problem of meaningfulness or significance, nor a problem of truth or acceptability. It was the problem of drawing a line (as well as this can be done) between the statements, or systems of statements, of the empirical sciences, and all other statements—whether they are of a religious or of a metaphysical character, or simply pseudo-scientific. Years later—it must have been in 1928 or 1929—I called this first problem of mine the 'problem of demarcation'. The criterion of falsifiability is a solution to this problem of demarcation, for it says that statements or systems of statements, in order to be ranked as scientific, must be capable of conflicting with possible, or conceivable, observations.

Notes

1. This is a slight oversimplification, for about half of the Einstein effect may be derived from the classical theory, provided we assume a ballistic theory of light.

2. See, for example, my Open Society and Its Enemies [Routledge & Kegan Paul, 1945], ch. 15, section iii, and notes 13–14.

3. 'Clinical observations', like all other observations, are interpretations in the light of theories...; and for this reason alone they are apt to seem to support those theories in the light of which they were interpreted. But real support can be obtained only from observations undertaken as tests (by attempted refutations); and for this purpose criteria of refutation have to be laid down beforehand—it must be agreed which observable situations, if actually observed, mean that the theory is refuted. But what kind of clinical responses would refute to the satisfaction of the analyst not merely a particular analytic diagnosis but psycho-analysis itself? And have such criteria ever been discussed or agreed upon by analysts? Is there not, on the contrary, a whole family of analytic concepts, such as 'ambivalence' (I do not suggest that there is no such thing as ambivalence), which would make it difficult, if not impossible, to agree upon such criteria? Moreover, how much headway has been made in investigating the question of the extent to which the (conscious or unconscious) expectations and theories held by the analyst influence the 'clinical responses' of the patient? (To say nothing about the conscious attempts to influence the patient by proposing interpretations to him, etc.) Years ago I introduced the term 'Oedipus effect' to describe the influence of a theory or expectation or prediction upon the event which it predicts or describes: it will be
Among the most fundamental issues on which Sir Karl [Popper] and I agree is our insistence that an analysis of the development of scientific knowledge must take account of the way science has actually been practiced. That being so, a few of his recurrent generalizations startle me. One of these provides the opening sentences of the first chapter of the Logic of Scientific Discovery: 'A scientist', writes Sir Karl, 'whether theorist or experimenter, puts forward statements, or systems of statements, and tests them step by step. In the field of the empirical sciences, more particularly, he constructs hypotheses, or systems of theories, and tests them against experience by observation and experiment.' The statement is virtually a cliché, yet in application it presents three problems. It is ambiguous in its failure to specify which of two sorts of 'statements' or 'theories' are being tested. That ambiguity can, it is true, be eliminated by reference to other passages in Sir Karl's writings, but the generalization that results is historically mistaken. Furthermore, the mistake proves important, for the unambiguous form of the description misses just that characteristic of scientific practice which most nearly distinguishes the sciences from other creative pursuits.

There is one sort of 'statement' or 'hypothesis' that scientists do repeatedly subject to systematic test. I have in mind statements of an individual's best guesses about the proper way to connect his own research problem with the corpus of accepted scientific knowledge. He may, for example, conjecture that a given chemical unknown contains the salt of a rare earth, that the obesity of his experimental rats is due to a specified component in their diet, or that a newly discovered spectral pattern is to be understood as an effect of nuclear spin. In each case, the next steps in his research are intended to try out or test the conjecture or hypothesis.

If it passes enough or stringent enough tests, the scientist has made a
discovery or has at least resolved the puzzle he had been set. If not, he
must either abandon the puzzle entirely or attempt to solve it with the aid
of some other hypothesis. Many research problems, though by no means
all, take this form. Tests of this sort are a standard component of what I
have elsewhere labelled 'normal science' or 'normal research', an enter­
prise which accounts for the overwhelming majority of the work done in
basic science. In no usual sense, however, are such tests directed to current
theory. On the contrary, when engaged with a normal research problem,
the scientist must premise current theory as the rules of his game. His
object is to solve a puzzle, preferably one at which others have failed, and
current theory is required to define that puzzle and to guarantee that,
given sufficient brilliance, it can be solved. Of course the practitioner of
such an enterprise must often test the conjectural puzzle solution that his
ingenuity suggests. But only his personal conjecture is tested. If it fails the
test, only his own ability not the corpus of current science is impugned.
In short, though tests occur frequently in normal science, these tests are
of a peculiar sort, for in the final analysis it is the individual scientist rather
than current theory which is tested.

This is not, however, the sort of test Sir Karl has in mind. He is above
all concerned with the procedures through which science grows, and he
is convinced that 'growth' occurs not primarily by accretion but by the
revolutionary overthrow of an accepted theory and its replacement by
a better one. (The subsumption under 'growth' of 'repeated overthrow'
is itself a linguistic oddity whose raison d'être may become more vis­
ible as we proceed.) Taking this view, the tests which Sir Karl empha­
sizes are those which were performed to explore the limitations of accept­
ted theory or to subject a current theory to maximum strain. Among
his favourite examples, all of them startling and destructive in their out­
come, are Lavoisier's experiments on calcination, the eclipse expedition

* Calcination occurs when a metal is burned in air, forming a calx or oxide.
According to the phlogiston theory, metals (and all other combustible substances)
are compounds of an earthy calx and the fiery element, phlogiston. When a metal
burns, the phlogiston is released, leaving the calx as a residue. Because metals
gain weight when they are calcined, some proponents of the phlogiston theory
conjectured that phlogiston must have negative weight. Others inferred that some
other substance must combine with the metal when the phlogiston is released. By
careful experiments in the 1770s, Antoine Lavoisier (1743–94) showed that the
weight gained during calcination is entirely due to the metal combining with a
gas in the air, which he named oxygen. Lavoisier's oxygen theory of calcination
(and, more generally, of combustion) overthrew the phlogiston theory and gave
rise to a revolution in chemistry. See James B. Conant, ed., The Overthrow of the
Harvard University Press, 1950); reprinted in Harvard Case Histories in Experi­
University Press, 1966). See also Alan Musgrave, "Why Did Oxygen Supplant
Kuhn • Logic of Discovery or Psychology of Research?

of 1919,† and the recent experiments on parity conservation.‡ All of course, are classic tests, but in using them to characterize scientific activity Sir Karl misses something terribly important about them. Episodes like these are very rare in the development of science. When they occur, they are generally called forth either by a prior crisis in the relevant field (Lavoisier’s experiments or Lee and Yang’s§) or by the existence of a theory which competes with the existing canons of research (Einstein’s general relativity). These are, however, aspects of or occasions for what I have elsewhere called ‘extraordinary research’, an enterprise in which scientists do display very many of the characteristics, Sir Karl emphasizes, but one which, at least in the past, has arisen only intermittently and under quite special circumstances in any scientific speciality. ⁶

I suggest then that Sir Karl has characterized the entire scientific enterprise in terms that apply only to its occasional revolutionary parts. His emphasis is natural and common: the exploits of a Copernicus or Einstein make better reading than those of a Brahe or Lorentz; † Sir Karl

---


* For information about the eclipse expedition of 1919 and its role in confirming Einstein’s general theory of relativity, see the preceding reading by Karl Popper, “Science: Conjectures and Refutations.” Further discussion can be found in “Popper’s Demarcation Criterion,” in the commentary on chapter 1, and in “Two Arguments for Explanadonism,” in the commentary on chapter 4.

† Kuhn is referring to the experiments performed by Chien-Shiung Wu and her associates in 1956–57, which verified the conjecture of Tsung Dao Lee and Chen Ning Yang that parity is not conserved in weak interactions. Wu’s results were soon confirmed by other groups and Lee and Yang received the Nobel prize in physics in 1957 for their discovery of parity violation. For a description of Wu’s experiment and an explanation of its revolutionary significance, see Eugene Wigner, “Violations of Symmetry in Physics,” Scientific American 213 (1965): 28–36 and Martin Gardner, The New Ambidextrous Universe, 3d rev. ed. (New York: W. H. Freeman, 1990).

‡ For Kuhn, Tycho Brahe (1546–1601) and H. A. Lorentz (1853–1928) exemplify the conservative scientist practicing normal science. Brahe objected to Copernicus’s revolutionary theory of a heliocentric universe on physical, astronomical, and religious grounds, proposing in its place his own version of a geostatic system. Like Ptolemy, Brahe had the sun moving around the earth, but unlike Ptolemy, he made the other planets orbit around the sun. In this way, Brahe was able to capture many of the explanatory features of Copernicus’s theory without having to attribute any motion to the earth. Lorentz, like most physicists of his day, believed that light and other electromagnetic radiation propagates in an aether that is at rest with respect to absolute space. In order to account for the null result of the Michelson-Morley experiment, Lorentz (and, independently, Fitzgerald) postulated the famous Lorentz-Fitzgerald contraction according to which all physical objects contract in their direction of motion. Lorentz later introduced time dilation, thus obtaining the Lorentz transformations that lie at the heart of Einstein’s special
would not be the first if he mistook what I call normal science for an intrinsically uninteresting enterprise. Nevertheless, neither science nor the development of knowledge is likely to be understood if research is viewed exclusively through the revolutions it occasionally produces. For example, though testing of basic commitments occurs only in extraordinary science, it is normal science that discloses both the points to test and the manner of testing. Or again, it is for the normal, not the extraordinary practice of science that professionals are trained; if they are nevertheless eminently successful in displacing and replacing the theories on which normal practice depends, that is an oddity which must be explained. Finally, and this is for now my main point, a careful look at the scientific enterprise suggests that it is normal science, in which Sir Karl’s sort of testing does not occur, rather than extraordinary science which most nearly distinguishes science from other enterprises. If a demarcation criterion exists (we must not, I think, seek a sharp or decisive one), it may lie just in that part of science which Sir Karl ignores.

In one of his most evocative essays, Sir Karl traces the origin of ‘the tradition of critical discussion [which] represents the only practicable way of expanding our knowledge’ to the Greek philosophers between Thales and Plato, the men who, as he sees it, encouraged critical discussion both between schools and within individual schools. The accompanying description of Presocratic discourse is most apt, but what is described does not at all resemble science. Rather it is the tradition of claims, counter-claims, and debates over fundamentals which, except perhaps during the Middle Ages, have characterized philosophy and much of social science ever since. Already by the Hellenistic period mathematics, astronomy, statics and the geometric parts of optics had abandoned this mode of discourse in favour of puzzle solving. Other sciences, in increasing numbers, have undergone the same transition since. In a sense, to turn Sir Karl’s view on its head, it is precisely the abandonment of critical discourse that marks the transition to a science. Once a field has made that transition, critical discourse recurs only at moments of crisis when the bases of the field are again in jeopardy. Only when they must choose between competing theories do scientists behave like philosophers. That, I think, is why Sir Karl’s brilliant description of the reasons for the choice between metaphysical systems so closely resembles my description of the reasons for choosing between scientific theories. In neither choice, as I shall shortly try to show, can testing play a quite decisive role.

---

There is, however, good reason why testing has seemed to do so, and in exploring it Sir Karl's duck may at last become my rabbit.\footnote{The duck-rabbit is a visually ambiguous drawing, made popular among philosophers by Ludwig Wittgenstein in his *Philosophical Investigations* (1953). It can be seen either as a duck's head with a long beak or as a rabbit's head with long ears, but it cannot be seen as both at the same time. It is a favorite with philosophers of science (such as Kuhn, Hanson, and Feyerabend) wishing to emphasize the theory-ladenness of observation.} No puzzle-solving enterprise can exist unless its practitioners share criteria which, for that group and for that time, determine when a particular puzzle has been solved. The same criteria necessarily determine failure to achieve a solution, and anyone who chooses may view that failure as the failure of a theory to pass a test. Normally, as I have already insisted, it is not viewed that way. Only the practitioner is blamed, not his tools. But under the special circumstances which induce a crisis in the profession (e.g., gross failure, or repeated failure by the most brilliant professionals) the group's opinion may change. A failure that had previously been personal may then come to seem the failure of a theory under test. Thereafter, because the test arose from a puzzle and thus carried settled criteria of solution, it proves both more severe and harder to evade than the tests available within a tradition whose normal mode is critical discourse rather than puzzle solving.

In a sense, therefore, severity of test-criteria is simply one side of the coin whose other face is a puzzle-solving tradition. That is why Sir Karl's line of demarcation and my own so frequently coincide. That coincidence is, however, only in their outcome; the process of applying them is very different, and it isolates distinct aspects of the activity about which the decision—science or non-science—is to be made. Examining the vexing cases, for example, psychoanalysis or Marxist historiography, for which Sir Karl tells us his criterion was initially designed,\footnote{I concur that they cannot now properly be labelled 'science'. But I reach that conclusion by a route far surer and more direct than his. One brief example may suggest that of the two criteria, testing and puzzle solving, the latter is at once the less equivocal and the more fundamental.} I concur that they cannot now properly be labelled 'science'. But I reach that conclusion by a route far surer and more direct than his. One brief example may suggest that of the two criteria, testing and puzzle solving, the latter is at once the less equivocal and the more fundamental.

To avoid irrelevant contemporary controversies, I consider astrology rather than, say, psychoanalysis. Astrology is Sir Karl's most frequently cited example of a 'pseudo-science'.\footnote{He says: 'By making their interpretations and prophecies sufficiently vague they [astrologers] were able to explain away anything that might have been a refutation of the theory had the theory and the prophecies been more precise. In order to escape falsification they destroyed the testability of their theory.' Those generalizations catch something of the spirit of the astrological enterprise. But taken at all literally, as they must be if they are to provide a demarcation criterion, they are impossible to support. The history of astrology during the cen-}
turies when it was intellectually reputable records many predictions that categorically failed.\(^\text{13}\) Not even astrology's most convinced and vehement exponents doubted the recurrence of such failures. Astrology cannot be barred from the sciences because of the form in which its predictions were cast.

Nor can it be barred because of the way its practitioners explained failure. Astrologers pointed out, for example, that, unlike general predictions about, say, an individual's propensities or a natural calamity, the forecast of an individual's future was an immensely complex task, demanding the utmost skill, and extremely sensitive to minor errors in relevant data. The configuration of the stars and eight planets was constantly changing; the astronomical tables used to compute the configuration at an individual's birth were notoriously imperfect; few men knew the instant of their birth with the requisite precision.\(^\text{14}\) No wonder, then, that forecasts often failed. Only after astrology itself became implausible did these arguments come to seem question-begging.\(^\text{15}\) Similar arguments are regularly used today when explaining, for example, failures in medicine or meteorology. In times of trouble they are also deployed in the exact sciences, fields like physics, chemistry, and astronomy.\(^\text{16}\) There was nothing unscientific about the astrologer's explanation of failure.

Nevertheless, astrology was not a science. Instead it was a craft, one of the practical arts, with close resemblances to engineering, meteorology, and medicine as these fields were practised until little more than a century ago. The parallels to an older medicine and to contemporary psychoanalysis are, I think, particularly close. In each of these fields shared theory was adequate only to establish the plausibility of the discipline and to provide a rationale for the various craft-rules which governed practice. These rules had proved their use in the past, but no practitioner supposed they were sufficient to prevent recurrent failure. A more articulated theory and more powerful rules were desired, but it would have been absurd to abandon a plausible and badly needed discipline with a tradition of limited success simply because these desiderata were not yet at hand. In their absence, however, neither the astrologer nor the doctor could do research. Though they had rules to apply, they had no puzzles to solve and therefore no science to practise.\(^\text{17}\)

Compare the situations of the astronomer and the astrologer. If an astronomer's prediction failed and his calculations checked, he could hope to set the situation right. Perhaps the data were at fault: old observations could be re-examined and new measurements made, tasks which posed a host of calculational and instrumental puzzles. Or perhaps theory needed adjustment, either by the manipulation of epicycles, eccentrics, equants, etc., or by more fundamental reforms of astronomical technique. For more than a millennium these were the theoretical and mathematical puzzles around which, together with their instrumental counterparts, the astronomical research tradition was constituted. The astrologer, by contrast, had
no such puzzles. The occurrence of failures could be explained, but par­
ticular failures did not give rise to research puzzles, for no man, however
skilled, could make use of them in a constructive attempt to revise the
astrological tradition. There were too many possible sources of difficulty,
most of them beyond the astrologer’s knowledge, control, or responsibility.
Individual failures were correspondingly uninformative, and they did not
reflect on the competence of the prognosticator in the eyes of his profes­
sional compeers.18 Though astronomy and astrology were regularly prac­
tised by the same people, including Ptolemy, Kepler, and Tycho Brahe,
there was never an astrological equivalent of the puzzle-solving astronomical
tradition. And without puzzles, able first to challenge and then to
attest the ingenuity of the individual practitioner, astrology could not have
become a science even if the stars had, in fact, controlled human destiny.

In short, though astrologers made testable predictions and recognized
that these predictions sometimes failed, they did not and could not engage
in the sorts of activities that normally characterize all recognized sciences.
Sir Karl is right to exclude astrology from the sciences, but his over-con­
centration on science’s occasional revolutions prevents his seeing the
surest reason for doing so.

That fact, in turn, may explain another oddity of Sir Karl’s historiog­
raphy. Though he repeatedly underlines the role of tests in the replace­
ment of scientific theories, he is also constrained to recognize that many
theories, for example the Ptolemaic, were replaced before they had in fact
been tested.19 On some occasions, at least, tests are not requisite to the
revolutions through which science advances. But that is not true of puz­
zles. Though the theories Sir Karl cites had not been put to the test before
their displacement, none of these was replaced before it had ceased ade­
quately to support a puzzle-solving tradition. The state of astronomy was
a scandal in the early sixteenth century. Most astronomers nevertheless
felt that normal adjustments of a basically Ptolemaic model would set the
situation right. In this sense the theory had not failed a test. But a few
astronomers, Copernicus among them, felt that the difficulties must lie in
the Ptolemaic approach itself rather than in the particular versions of Ptol­
emaic theory so far developed, and the results of that conviction are al­
ready recorded. The situation is typical.20 With or without tests, a
puzzle-solving tradition can prepare the way for its own displacement. To
rely on testing as the mark of a science is to miss what scientists mostly
do and, with it, the most characteristic feature of their enterprise. . . .

Notes

1. Popper [1959], p. 27.
2. For an extended discussion of normal science, the activity which practitioners
are trained to carry on, see my [1962], pp. 23–42, and 135–42. It is important to