Chapter 17
Imre Lakatos and the logic of falsification

Bob Lieshout

Introduction

At the beginning of the 1980s, after I had been confronted with the epistemological nonsense that seemed to guide research in comparative politics (Lijphart's most-similar systems design and Przeworski and Teune's most-different systems design are two glaring examples), I made a first attempt to formulate the rules that empirical scientists should follow in their search for truth (cf. Lieshout, 1983). After that, I taught courses on epistemology for many years. This obliged me to think through this subject matter again and again, with the result that I grew increasingly dissatisfied with my first effort because I had failed to see that Imre Lakatos's methodology of scientific research programmes – by neutralizing Thomas Kuhn's problem of incommensurability – added something essential to Karl Popper's perceptions on the falsification of empirical theories.[1] In my contribution to this *Festschrift* to honour my former colleague and dear friend Kees van Kersbergen, I shall set out the arguments why I believe this to be the case.

Requirements that a test of an empirical theory must meet

Universal non-existence statements and basic statements

How should empirical scientists act in a situation where, after they have deduced a prediction from a theory they consider worthy of testing and, subsequently, subjected this prediction to a test, it turns out that the test results do not agree with it? To be able to answer this question, it is well to realize, as Popper emphasized, that state-

1 I should admit that by 1974, Popper himself had come to an entirely different conclusion: 'I feel, unfortunately, obliged to warn the reader that Professor Lakatos has … misunderstood my theory of science; and that the series of long papers in which, in recent years, he has tried to act as a guide to my writings and the history of my ideas is, I am sorry to say, unreliable and misleading' (Popper, 1974: 999).

ments about observable phenomena can never be proved by, or be in contradiction with, observable events. Statements can only be justified or contradicted by other statements (cf. Popper, 1983: 93), or as Lakatos put it: '*no factual proposition can ever be proved from an experiment.* Propositions can only be derived from other propositions, they cannot be derived from facts' (Lakatos, 1974: 99; emphasis in original). When we accept this point, then we are immediately confronted with the following complication. Assuming that we have subjected a hypothesis derived from a certain theory to a test and have observed that the outcome of this test either confirms or contradicts this hypothesis and want to report this specific observation in a statement, we find that this is impossible. Take the statement 'there exists a black swan'. It is easy to see that this statement is completely unintelligible if we do not have certain theoretical notions about what is involved in 'blackness' and 'swanness'. 'Black' and 'swan' are universals, terms that refer to certain forms of law-like behaviour. This means that every statement describing a specific observation or sense experience transcends that observation or experience. Every statement describing an observation is inescapably theoretical as well.

The strongest way in which an empirical theory can be tested is to test statements belonging to that theory's empirical content. These have the form of universal non-existence statements. They forbid certain conceivable states of events to exist. A rather obvious example in view of the above is the statement 'there does not exist a non-white swan', which is the negation of the universal statement 'all swans are white'. This is the strongest possible test because one single observation of a counter example ('there exists a non-white swan') can set a process in motion that can lead to a theory's refutation.

In case our efforts at refutation do lead to the observation of a non-white swan, then we must record this observation in what Popper called a 'basic statement', which has the form of a 'singular existential statement' ('there exists a non-white swan in time-region *k*'). A basic statement must fulfil two criteria, a formal and a material one. The formal criterion is that the basic statement can be in contradiction with a universal statement. This can be done by simply omitting from the basic statement 'any reference to any individual space-time region' (Popper, 1980: 102). We then get the purely existential statement, 'there exists a non-white swan', which is in contradiction with the universal non-existence statement, 'there does not exist a non-white swan'. A basic statement must also meet the *material* criterion that the event recorded in the basic statement must be 'observable'; 'that is to say, basic statements must be testable, inter-subjectively, by "observation"' (Popper, 1980: 102). The statement, 'I saw a non-white swan in my garden yesterday' does not satisfy this material requirement, although it describes an event that occurred in an 'individual region of space and time' (Popper, 1980: 103). It must be possible for others to subject the statement to tests, to check whether it is true, and this is clearly not the case with this statement. 'Stray basic statements' are therefore not admissible. We should only accept basic statements that result from our attempts to refute the theory.

A weaker form of testing concerns the testing of statements that belong to the logical content of the theory, the class of statements about observable phenomena permitted by the theory. These statements have the character of probability statements: the probability that a certain event '(e)' will occur given '(c)' equals *r*, where ($0 < r < 1$). These are weaker tests because, strictly speaking, these types of statements cannot be refuted. No matter how many observations to the contrary we have collected and reported in basic statements, the probability statement can always be upheld. Accordingly, there is a fundamental asymmetry between universal non-existence statements and probability statements as far as their possible contribution to the growth of knowledge is concerned. This asymmetry cannot be remedied by the formulation of certain rejection rules prior to testing these statements.

Decisions, decisions, and even more decisions

It will be clear that there are no automatisms in the process of subjecting universal non-existence statements derived from a certain theory to tests and reporting the results of these tests in a basic statement. All the time, the scientists concerned must make decisions: whether the test is severe enough, whether it has been properly devised and executed, what precisely has been observed, as well as whether to accept the singular existence statement in which the observation has been reported as a basic statement. '[T]hus it is *decisions* which settle the fate of theories' (Popper, 1980: 108). In this connection, I wish to point out that decisions belong to what Popper called 'world 2', the 'world of subjective experiences (such as thought processes)' (Popper, 1982: 181) and draw attention to Kuhn's observation that the choice of a certain paradigm as opposed to another can never be 'unequivocally settled by logic and experiment alone' (Kuhn, 1970: 94). Many a critic of Kuhn, Lakatos not the least among them, has claimed that this statement shows that Kuhn opened the floodgates to irrationalism by turning theory choice into something like a 'mystical conversion' (Lakatos, 1974: 93). In my view, this is reading far too much in something that should be a matter of course, seeing that decisions belong to world 2. As Kuhn observed in his 'Reflections on my critics', 'to say that, in matters of theory-choice, the force of logic and observation cannot in principle be compelling is neither to discard logic and observation nor to suggest that there are not good reasons for favouring one theory over another' (Kuhn, 1974: 234). Our attempts to test a theory as severely as possible and the results these tests produce provide arguments as to whether to accept or reject a certain theory, but these arguments can never compel us to accept them. We can always *decide*, for whatever reason, valid or invalid, to do another test to suspend our judgment. Arguments can never absolve us from our responsibility for our decisions. Arguments belong to a different world, Popper's world 3: 'the world of the products of the human mind, such as stories, explanatory myths, tools, scientific theories (whether true or false), scientific problems, social institutions, and works of art' (Popper and Eccles, 1981: 38). They can therefore motivate us to take a certain decision, provide us with reasons to do so, but they do not determine it: 'it is always we who decide' (Popper, 1971: 233).

The conditions that must be met before an empirical scientist can decide an empirical theory should be considered falsified

Naive methodological falsificationism

It is expected from empirical scientists that they are prepared to take risks, that they subject their conjectures to the severest tests possible, and that, if it turns out that their conjectures contradict reality, they are prepared to reject these conjectures, however much the invention of these conjectures has demanded of their intellectual capacities. This dare-devil attitude agrees with Popper's claim that scientific honesty demands that 'anyone who advocates the empirical-scientific character of a theory [...] must be able to specify under what conditions he would be prepared to regard it as falsified, i.e. he should be able to describe at least some potential falsifiers' (Popper, 1983: xxi). All this seems to indicate that in a situation where a singular existence statement, after having been subjected to vigorous tests, is accepted as a basic statement and the theory is consequently confronted with an anomaly or a counterexample, scientists should consider this theory to be falsified and begin to search for a new and better one. But this would be going too fast, considering that every empirical theory, even very successful ones like Newton's gravitational theory, is 'submerged in an ocean of "anomalies" (or, if you wish, "counterexamples")' (Lakatos, 1974: 133). If we adopted this 'naive methodological falsificationist' (Lakatos, 1974: 116) position, then scientific progress would become impossible. Every empirical theory would instantly be falsified. This implies that the fact that counterexamples or anomalies have been found in no way obliges scientists to consider a theory as falsified and to stop working on it. On the contrary, 'the scientist who pauses to examine every anomaly he notes will seldom get significant work done' (Kuhn, 1970: 82).

Sophisticated methodological falsificationism

There is yet another reason why the dare-devil attitude of the naive methodological falsificationist must be rejected. It can lead to the refutation of empirical theories that are in fact true. We should realize that after naive methodological falsificationists have decided to accept a basic statement, they have yet to take two other types of decisions before they can decide that a theory has been falsified. The first follows from the perception that in case we have decided that a universal non-existence statement is in contradiction with certain observable events, strictly speaking, the whole of our knowledge is in doubt. This is the famous Duhem-Quine thesis, after the French physicist Pierre Duhem, who was the first to see that we can never test a single statement in isolation but only a whole group of theories, and the American philosopher Willard Quine, who radicalized this understanding. Duhem originally wrote the following:

> In sum, the physicist can never subject an isolated hypothesis to experimental test, but only a whole group of hypotheses; when the experiment is in disagreement with his predictions, what he learns is that at least one of the hypotheses constituting this group is unacceptable and ought to be modified; but the experiment does not designate which one should be changed. (Duhem, 1982: 187).

Quine went even further in claiming that it is not a group of theories but the whole of science that is 'the unit of empirical significance' (Quine, 1951: 39). '[S]tatements about the external world face the tribunal of sense experience not individually but only as a corporate body' (Quine, 1951: 38). A universal non-existence statement may be in contradiction to a singular existence statement for an infinite number of reasons, for example, because the theories with the help of which we built the instruments that enabled us to make our observations are wrong. This also implies that 'given sufficient imagination, any theory … can be permanently saved from "refutation" by some suitable adjustment in the background knowledge in which it is embedded' (Lakatos, 1974: 184), and it therefore becomes impossible to falsify it. It appears that the motor of the scientific enterprise, the refutation of conjectures, has come to a grinding halt. Should we despair? I should think not because, precisely in its most radical interpretation, the Duhem-Quine thesis turns out to be equivalent to and just as profound as the statement that 'everything is connected with everything else'. It may be true, but it does not help us at all if we wish to find out why things are as they are. If we believe in the scientific enterprise, in the possibility that we can get nearer and nearer to the truth, then we must put aside as much as possible of our knowledge as 'unproblematic background knowledge' (Popper, 1968: 238). The more we decide to put into the category of unproblematic background knowledge, the easier it becomes to regard a counterexample as posing a problem for a specific theory.

The second type of decision the naive methodological falsificationist must make follows from the following consideration. Every empirical theory contains a non-specified universal non-existence statement, or ceteris paribus clause, to the effect that no other relevant cause is at work anywhere in the universe. This implies that in case an empirical scientist decides to accept a certain basic statement, the theory need not be in danger, because he or she can always decide that this ceteris paribus clause was wrong and that, in fact, another cause is at work, one that can explain why the original hypothesis turned out to be false, but one the scientist until then had not taken into consideration. This point is nicely illustrated by Lakatos's 'imaginary story' about the behaviour of a 'Newtonian physicist' who is confronted with 'a case of planetary misbehaviour'. This physicist calculates the path of a newly discovered planet *p* but finds that the planet deviates from that path. Does this lead the physicist to the decision that the theory must be regarded as refuted? 'No. He suggests that there must be a hitherto unknown planet *p*' which perturbs the path of *p*'' (Lakatos, 1974: 100). Planet *p*', however, is not found. 'Does our scientist abandon Newton's theory and his idea of a perturbing planet? No. He suggests that a cloud of cosmic dust hides the planet from us' (Lakatos, 1974: 101), and so on, and so on.

Every time the naive methodological falsificationist decides to accept a basic statement that contradicts a hypothesis derived from the theory under test, he or she also faces the decision whether to accept the ceteris paribus clause or not. If he or she does so, which means that he or she accepts that no other relevant cause is at work in the universe, then he or she must regard the theory as falsified. But how is the ceteris paribus clause to be tested? How can he or she establish that there is no other relevant cause at work in the universe? Obviously, he or she cannot, and this means that in case the naive methodological falsificationist decides to accept the ceteris paribus clause, he or she runs the risk of considering a theory to be falsified and to stop working on it, while in fact, the theory is true. As far as Lakatos is concerned, this is an unacceptable risk. Inspired by Kuhn, Lakatos therefore proposes three criteria a new theory must meet before we accept that an older one is falsified, freeing us from having to take the dangerous decision whether to accept the ceteris paribus clause or not.

In *The structure of scientific revolutions*, Kuhn observed that

> the act of judgment that leads scientists to reject a previously accepted theory is always based upon more than a comparison of that theory with the World. The decision to reject one paradigm is always simultaneously the decision to accept another, and the judgment leading to

> that decision involves the comparison of both paradigms with nature *and* with each other (Kuhn, 1970: 77).

Tacitly accepting Kuhn's point, Lakatos sets out his sophisticated methodological falsificationist position and stipulates that we shall only consider a theory T_1 to be falsified by a theory T_2 if and only if T_2 fulfils three conditions (Lakatos, 1974: 116).

The first condition is that T_2 explains the previous *success* of T_1 (my emphasis). To speak of the 'success' of a theory, I consider a major innovation, evidently meant to neutralize Kuhn's observations on the 'incommensurability' of two competing paradigms. According to Kuhn, it would be a mistake to believe that Newton's theory can be translated on a one-on-one basis into Einstein's theory of general relativity so that Newton's laws become 'a limiting case of Einstein's' (Kuhn, 1970: 102). This applies to all paradigm shifts: 'within the new paradigm, old terms, concepts, and experiments fall into new relationships one with the other. The inevitable result is what we must call, though the term is not quite right, a misunderstanding between the two competing schools' (Kuhn, 1970: 149). The terms used in the old paradigm cannot be translated into the terms of the new paradigm without a loss of meaning. Adherents of different paradigms are 'members of different language communities' (Kuhn, 1970: 175). It is in this sense that two competing paradigms are incommensurable. The adherents of different paradigms are unable to communicate fully. Because they do not acknowledge a common higher standard, it becomes impossible for them to compare these paradigms to establish which one of them is nearer to the truth. Lakatos gets around this difficulty by not requiring that theories are subjected to a point-by-point comparison of their content but that they are compared with respect to their empirical success, and if T_2 is able to explain the successes of T_1, this is a sound first indication that T_2 may be better than T_1.

Lakatos's second condition states that T_2 must also have excess empirical content over T_1, by which Lakatos means that T_2 'predicts *novel* facts, that is, facts improbable in the light of, or even forbidden' by T_1 (T_2 must be 'theoretically progressive'). This is precisely what Einstein's theory did compare to Newton's with respect to the degree of the bending of starlight by the sun. But this still does not suffice. Before we can decide that T_1 has been falsified by T_2, that theory must fulfil yet another condition, which is that a part of this excess content has been corroborated (T_2 must also be 'empirically progressive'). This is exactly what the Eddington expedition in the summer of 1919 provided. Its observations showed that the light of the stars near the sun was deflected in agreement with Einstein's gravitation law.

Popper, who almost fifty years later related how impressed he was by Einstein's triumph – 'We all … were thrilled with the result of Eddington's eclipse observations which in 1919 brought the first important *confirmation* of Einstein's theory of gravitation' (Popper, 1968: 34; my emphasis), realized that the growth of knowledge cannot occur by 'conjectures and refutations' alone and that confirmations must play a vital role, too. Confirmations should, however, 'count only if they are the result of *risky predictions*' and if they are '*the result of a genuine test of the theory*' (Popper, 1968: 36; emphasis in original), which certainly applies to Einstein's theory of general relativity.

Negative and positive heuristic

Most empirical scientists work within a scientific research programme, a series of theories with a common hard core consisting of a principle of explanation and certain crucial assumptions. These 'normal' scientists will continue to work in this programme until a rival arrives on the scene that is superior in the sense that it fulfils the three conditions formulated above. Does this imply, as Kuhn has suggested time and again, that the best thing a scientist can do if he or she wishes to contribute to the growth of knowledge is to become a normal scientist? Kuhn admits that 'the areas investigated by normal science are, of course, minuscule; the enterprise now under discussion has drastically restricted vision'. But he emphasizes that 'those restrictions, born from confidence in a paradigm, turn out to be essential to the development of science' (Kuhn, 1970: 24). This greatly worried Feyerabend and made him wonder: 'are we here presented with *methodological prescriptions* which tell the scientist how to proceed; or are we given a *description*, void of any evaluative element, of those activities which are generally called "scientific"?' (Feyerabend, 1974: 198). Kuhn replied that Feyerabend was 'right in claiming that my work repeatedly makes normative claims' (Kuhn, 1974: 233), the most important of which was that 'scientists should behave essentially as they do [as normal scientists; BL] if their concern is to improve scientific knowledge' (Kuhn, 1974: 237). This was precisely Feyerabend's nightmare:

> more than one social scientist has pointed out to me that now at last he had learned how to turn his field into a "science" – by which of course he meant that he had learned how to *improve* it. The recipe, according to these people, is to restrict criticism, to reduce the number of comprehensive theories to one, and to create a normal science that has this one theory as paradigm' (Feyerabend, 1974: 198).

This Kuhnian ambiguity is also present in Lakatos's essay, where the latter stated that 'for the sophisticated falsificationist a theory is "acceptable" or "scientific" only if it has corroborated excess empirical content over its predecessor' (Lakatos, 1974: 116). This position is, however, untenable. In empirical science, everything turns on (competing) principles of explanation. There is nothing that forbids the empirical scientist to invent new principles of explanation and to develop theories based on them (provided that no contradictory statements can be derived from these theories and that they are formulated in strictly universal terms). This is just as 'acceptable' or 'scientific' (if not more arduous and unrewarding) as working in a scientific research programme. Besides, where should the rivals of an established scientific research programme come from that are indispensable for their eventual falsification and the growth of knowledge if there were no 'revolutionary' scientists, scientists who are prepared to think outside the confines of a scientific research programme?

How do scientists working in a scientific research programme make progress? Not by continually questioning the validity of the hard core of this programme. This is what Lakatos called the 'negative heuristic' of the research programme (Lakatos, 1974: 133 and 135). Scientists working in the programme should not be concerned with establishing whether the hard core is true or not – whether they do this out of ignorance or because they have realized that questioning the hard core will only lead to an infinite regress. What they are doing is thinking through the hard core's implications and using their 'ingenuity to articulate or even invent "auxiliary hypotheses", which form a *protective belt* around this core' (Lakatos, 1974: 133; emphasis in original), and it is these hypotheses that are subjected to tests.

Scientists working in a scientific research programme make progress by deriving new predictions from the hard core and subsequently testing these predictions. In this, they are led by the 'positive heuristic' of the programme (Lakatos, 1974: 135). This positive heuristic tells the scientist not to be discouraged by the 'ocean of anomalies' the programme is submerged in. The positive heuristic thus 'accounts for the *relative autonomy of theoretical science*' and encourages scientists to forge ahead 'with almost complete disregard of "refutations"' (Lakatos, 1974: 137; emphasis in original).

Conclusion

Lakatos's 'methodology of scientific research programmes' provides the rules that should guide empirical scientists in their decision whether a certain theory should be considered to be falsified or not. In this manner, Lakatos demonstrated that the search for truth, the unending quest for a deeper and deeper understanding of the (social) world, is not an illusionary project. At the same time, it cannot be denied that – under the influence of Kuhn's incommensurability thesis, Feyerabend's conclusion that 'the numerous deviations from the straight path of rationality which we observe in actual science may well be *necessary*' (Feyerabend 1974: 219; emphasis in original), or the Duhem-Quine thesis – countless scientists have despaired of the scientific enterprise and decided that the growth of knowledge is a myth, and subsequently have taken refuge in irrationalism or language games. I believe such feelings of despair are wholly unwarranted. In this conclusion, I shall very briefly address Kuhn's and Feyerabend's objections. As I have already rejected the Duhem-Quine thesis in Section 3.2, I shall not pay any further attention to it here.

With respect to Kuhn's 'incommensurability thesis', it should be noted that Kuhn admitted in his 'Reflections on my critics' that not too radical conclusions should be drawn from his argument. All he meant to say was that translations between languages or theories at times can be very difficult, that a perfect translation does not exist, and that any 'translation manual inevitably embodies a theory' (Kuhn, 1974: 269), which does not sound particularly worrisome to one who accepts the approach I have adopted here. Moreover, the thesis has lost its force in view of Lakatos's ingenious wording of the first condition that a theory T_y must meet before it can be said that it has falsified theory T_x.

Feyerabend's criticism was inspired by the work of Kuhn. Kuhn claimed to have discovered as 'an historian of science' that 'much scientific behaviour, including that of the very greatest scientists, persistently violated accepted methodological canons' and felt compelled to ask 'why those failures to conform did not seem at all to inhibit the success of the enterprise' (Kuhn, 1974: 236). Picking up this point, Feyerabend argued that the standards developed by Lakatos were no more than a '*verbal ornament* ... a memorial to happier times when it was still thought possible to run a complex and often catastrophic business like science by following a few simple and "rational" rules' (Feyerabend, 1974: 215; emphasis in original). I believe that Kuhn and Feyerabend are correct in claiming that famous scientists have regularly behaved in ways that deviate from Lakatos's standards but that the conclusion that Feyerabend draws from this, namely, that these are thus irrelevant ornaments, is completely mistaken. It is the fate of every normative theory prescribing how people, in this case empirical scientists, ought to behave that these prescriptions need not agree with their actual behaviour. The only thing that counts with respect to the validity of a normative theory like Lakatos's is that it is consistent, that no contradictory statements (prescriptions) can be deduced from it, and I believe that it passes this test.

References

Duhem, Pierre (1982). *The aim and structure of physical theory*. Princeton: Princeton University Press.

Feyerabend, Paul (1974). Consolation for the specialist, pp. 197-230 in Imre Lakatos and Alan Musgrave (eds.), *Criticism and the growth of knowledge*. Cambridge: Cambridge University Press.

Kuhn, Thomas (1970). *The structure of scientific revolutions*. Chicago: Chicago University Press.

Kuhn, Thomas (1974). Reflections on my critics, pp. 231-278 in Imre Lakatos and Alan Musgrave (eds.), *Criticism and the growth of knowledge*. Cambridge: Cambridge University Press.

Lakatos, Imre (1974). Falsification and the methodology of scientific research programmes, pp. 19-196 in Imre Lakatos and Alan Musgrave (eds.), *Criticism and the growth of knowledge*. Cambridge: Cambridge University Press.

Lieshout, Bob (1983). Kleine methodologie voor de vergelijkende politicologie. *Acta Politica* 3: 307-328.

Popper, Karl R. (1968). *Conjectures and refutations: The growth of scientific knowledge*. New York: Harper & Row.

Popper, Karl R. (1971). *The open society and its enemies: Hegel and Marx*. Princeton: Princeton University Press.

Popper, Karl R. (1974). Replies to my critics, pp. 961-1197 in Paul A. Schilpp (ed.), *The philosophy of Karl Popper*. La Salle: Open Court.

Popper, Karl R. (1980). *The logic of scientific discovery*. London: Hutchinson.

Popper, Karl R. (1982). *Unended quest: An intellectual biography*. London: Fontana/Collins.

Popper, Karl R. (1983). *Realism and the aim of science*. London: Hutchinson.

Popper, Karl R. and J.C. Eccles (1981). *The self and its brain*. Berlin: Springer.

Quine, Willard V. (1951). Two dogmas of empiricism. *The Philosophical Review* 60: 20-43.