

Notes on Lakatos

Author(s): Thomas S. Kuhn

Source: *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Vol. 1970 (1970), pp. 137-146

Published by: Springer

Stable URL: <https://www.jstor.org/stable/495758>

Accessed: 27-10-2019 18:24 UTC

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

Springer is collaborating with JSTOR to digitize, preserve and extend access to *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*

THOMAS S. KUHN

NOTES ON LAKATOS

I. INTRODUCTION

The invitation which has brought me here to comment on Professor Lakatos' paper has given me much pleasure, for I have long been an admirer of his work, particularly of his early four-part paper, 'Proofs and Refutations'. That does not mean, of course, that we have often agreed, but I have enjoyed the arguments that resulted and looked forward to this one. My pleasure, furthermore, was considerably enhanced when I discovered that Lakatos was going to be able to confound all precedent, his own and others, by getting this paper to me well in advance. It is a privilege few commentators are given, and I am correspondingly grateful.

All that I could have said before opening Lakatos' manuscript – in fact, I did so in letters to both Lakatos and Roger Buck. Reading it has only increased my satisfaction, but in an unanticipated way. As with some earlier Lakatos papers, I have had trouble with translation. Phrases like "the methodology of research programs" are not part of my familiar mode of communication; phrases like 'internal' and 'external history', although familiar, are used by Lakatos in novel and unexpected ways. I believe, however, that I have managed the translation, though perhaps without assimilating the language. As I have done so and simultaneously caught the spirit of his enterprise, I have been surprised and pleased at how congenial I find his present views. I conclude, finally, that I have read no paper on scientific method which expresses opinions so closely paralleling my own, and I am necessarily encouraged by that discovery, for it may mean that in the future I shall not be quite as alone in the methodological arena as I have been in the past. The resemblance between our views ought also, of course, to disqualify me as a commentator. One of my critics rather than I should be standing here, and if I had seen the difficulty in time, one would be. Since I did not, I shall have to do my best to play the critic. It is therefore fortunate that my agreement with Lakatos, however far it extends, is less than total.

Boston Studies in the Philosophy of Science, VIII. All rights reserved.

II. PARALLELS

Before turning to the points at which we part company, I shall have to enumerate briefly and globally the areas in which our views coincide. There is, I think, no other way to isolate our difference, or, since 'difference' may not be the right word, to discover those portions of his paper in which Lakatos says things that I could never make my own.

Among our areas of agreement is the one Lakatos describes as meta-methodological or meta-historical. No historian, whether of science or some other human activity, can operate without preconceptions about what is essential, what is not. Those preconceptions do, if the historian deals with science, play an important role in determining what he takes to be 'internal', what 'external' in Lakatos' sense. Agassi has previously made the same point very effectively, and I welcome Lakatos' extension of it. I think of myself as having argued the converse even earlier, suggesting that failure to fit historical data provides grounds for criticizing a current methodological position. Lakatos has not, I shall shortly argue, yet altogether seen how to develop a philosophical basis for that converse, but I am not sure I have done better and am correspondingly gratified by his attempt.

That much agreement is probably not remarkable, but its extension from meta-methodology to substantive methodology is – or so it seems to me. I have, for example, repeatedly emphasized that the important scientific decisions – usually described as a choice between theories – are more accurately described as a choice between 'ways of doing science', or 'between traditions', or between 'paradigms.' Lakatos' insistence that the unit of choice is a 'scientific research program' seems to me to make the identical point.

Again, in discussing research conducted within a tradition, under the guidance of what I once called a paradigm, I have repeatedly insisted that it depends, in part, on the acceptance of elements which are not themselves subject to attack from within the tradition and which can be changed only by a transition to another tradition, another paradigm. Lakatos, I think, is making the same point when he speaks of the 'hard core of research programs,' the part which must be accepted in order to do research at all and which can be attacked only after embracing another research program.

Finally, though it does not exhaust our areas of agreement, I would point to Lakatos' emphasis on what he calls the 'degenerating stage' in the evolution of a research program, the stage in which it ceases to lead to new discoveries, in which *ad hoc* hypotheses accrue to it, and so on. I cannot myself tell the difference between what he has to say about this important stage and what I have said about the role of crisis in scientific development. Lakatos clearly does, but I get no help at all from the passages where he refers to them: for example, a reference late in his paper to "the Kuhnian psychological epiphenomenon of 'crisis'." [p. 120]

You will see, I think, why I speak of parallels and why I find them so encouraging. But they leave a puzzle. Why, if these parallels are real, is Lakatos so unable to see them? That he does not do so is illustrated by the phrase just quoted, and there are many others of the same sort in his paper. Undoubtedly part of the difficulty is the obscurity of my original presentation, something I can only regret. But I think that there is a deeper source, and it points to the areas in which we disagree or at least seem to.

Scattered through Lakatos' paper are a number of remarks like the following: Kuhn has, Lakatos suggests, come "up with a highly original vision of irrationally changing rational authority." [p. 116] Elsewhere he says, "When Kuhn and Feyerabend see irrational change, I predict that historians will be able to show there has been rational change." [p. 118] These reiterated contrasts between my irrationality and Lakatos' rationality isolate the difference which Lakatos sees between our views. For him it is apparently so deep that he remains blind to our close parallels. I shall argue that, even in suggesting the contrast, he is missing the point both of his present work and of my own.

I have never, in fact, accepted the description of my views as a defense of irrationality in science, but I have usually understood its source, seen why my critics thought the description apt. In this case, however, I cannot even do that. Considering the extent of the parallels between our views, Lakatos' use of terms like 'irrational' is, I think, only a mouthing of shibboleths. Either we are both defenders of irrationality, which I join him in doubting, or else, as I suppose, we are both trying to change a current notion of what rationality is. Arguments to that effect make up the balance of my remarks, though the issue in that form will not be entirely explicit until my conclusion.

III. 'INTERNAL' AND 'EXTERNAL'

Let me start by commenting on Lakatos' use of the terms 'internal' and 'external history'. In an early footnote he points out that the distinction is quite standard among historians of science but that he is using it in a new way. I am not, quite obviously, the man to be critical of a colleague who adapts an old term to his own purposes. What I think Lakatos does not realize, however, is how little need there is in this case to strain someone else's usage. The main virtue of the transition in terms is, I suspect, that it facilitates an unconscious sleight of hand.

In standard usage among historians, internal history is the sort that focuses primarily or exclusively on the professional activities of the members of a particular scientific community: What theories do they hold? What experiments do they perform? How do the two interact to produce novelty? External history, on the other hand, considers the relations between such scientific communities and the larger culture. The role of changing religious or economic traditions in scientific development thus belongs to external history, as does its converse. Among other standard topics for the externalist are scientific institutions and education, as well as the relations between science and technology. The internal-external distinction is not always hard and fast, but there is wide consensus in its application among historians. That consensus proves, I believe, at once implicitly vital and explicitly irrelevant to Lakatos' argument.

Obviously there is much overlap between normal usage and Lakatos'. In both, such factors as religion, economics, and education are external; Newton's Laws, Schrödinger's equation, and Lavoisier's experiments are internal. If there were no readily available alternatives, Lakatos' preemption of these terms would therefore be appropriate. But they would strain normal usage, for Lakatos' internal history is far narrower than that of the historian. It excludes, for example, all consideration of personal idiosyncrasy, whatever its role may have been in the choice of a theory, the creative act which produced it, or the form of the product which resulted. By the same token, it excludes such historical data as the failure of the man who creates a new theory and of his entire generation to see in that theory consequences which a later generation found there, a point I shall need to discuss further below. And, finally, it excludes consideration of mistakes or of what a later generation will see as having

been mistakes and will accordingly feel constrained to correct.

Historical data of these sorts are all central and essential for the internal historian of science. Often they provide his most revealing clues to what occurred. Since Lakatos insists they be excluded from internal history, I wonder why he adopts the term. Could he not easily instead have spoken of rational history, or better, of history constructed from the rational elements in a science's development? I think that is what, most fundamentally, he means: the 'internal' in Lakatos' sense and in this context is closely equivalent to 'rational' in the ordinary sense. Furthermore, Lakatos' 'internal' carries with it from the ordinary use of 'rational' an all-important characteristic: as a criterion of selection it is prior to the pursuit of history and independent of it.

If that is right, then it is, of course, apparent why Lakatos does change terms. If 'internal' were an independent term unequivocally applied, as it is for the historian, then one could hope to learn something about rational methodology from the study of internal history. But if 'internal history' is simply the rational part of history, then the philosopher can learn from it about scientific method only what he puts in. Lakatos' meta-methodological method is in danger of reducing to tautology.

IV. LAKATOS THE HISTORIAN

As developed so far, my argument applies completely only to the first half of Lakatos' paper. That is the part in which he sets up his version of the internal-external distinction and then shows how what one takes to be internal and external changes with the choice of a prior methodological position. The second part of the paper is, of course, different. There he suggests that the choice of a methodology supplies a meta-historical research program. The actual attempt to apply such a program to historical data may show that the program is degenerating. As a result, a new methodology may arise and be accepted. I myself believe that exactly that can and does happen. Yet I wonder why Lakatos should expect it to. Given what he has made of the internal-external distinction, and given also his conception of what a historian does, no such effect is possible. Lakatos, I now want to argue, skirts as close to tautology in the second half of his paper as in the first.

Midway through the paper, for example, he remarks: "History of

science (meaning here internal history) is a history of events which are selected and interpreted in a normative way.” [p. 108] With that point I would thoroughly agree if it meant only that all historians necessarily select and interpret their data. But Lakatos, when he introduces the term ‘normative’ means something else. He has previously suggested that it is “philosophy of science [which] provides normative methodologies” [p.91] to the historian. His point is not simply that the historian selects and interprets, but that prior philosophy supplies the whole set of criteria by which he does so. If that were the case, however, there would be no way at all in which the selected and interpreted data could react back on a methodological position to change it.

Fortunately for Lakatos’ point, other selective principles are available to the historian in addition to prior concepts of methodology. His narrative must, for example, be continuous in the sense that one event must lead into or set up the next; one may not skip about. In addition, his story must be plausible in the sense that men and institutions must behave in recognizable ways. It is legitimate to criticize a historian’s narrative by saying: That cannot be what occurred, for only a madman would behave that way, and we have been given no reason to believe that the king was mad. Finally, and for present purposes most important, history must be constructed without doing violence to the data available for selection and interpretation. Only if these and other internal criteria of the historian’s craft are used, can the results of historical research react back on and change the philosophical position with which the historian began.

My concern with Lakatos’ paper is that it throws all these criteria away, thus depriving history of any philosophical function. For example, just before the last passage quoted, Lakatos writes: “One way to indicate discrepancies between history and its rational reconstruction is to relate the internal history *in the text*, and indicate *in the footnotes* how actual history ‘misbehaved’ in the light of its rational reconstruction.” [p. 107] A recently published paper (his contribution to *Criticism and The Growth of Knowledge*) indicates what he means. In his text he tells a succession of straightforward stories, then in the footnotes he adds: that, of course, is not quite what happened; rather it is what would have happened if people had behaved rationally as they should. A somewhat different and equally informative example is contained in his present paper. Rational

reconstruction, Lakatos suggests, can properly attribute the idea of electron spin to Bohr in 1913. Probably, he concedes, Bohr did not think of it then, but it was compatible with the research program implied by the Bohr atom. In fact, however, as Lakatos surely knows, Bohr was quite skeptical of the idea of spin even in 1925. That is not because Bohr was irrational. Instead, Lakatos, by once more discarding evidence which does not fit his prior principle of rationality, has misconstrued Bohr's program. If one constructs it properly from the evidence, one discovers that spin fits it very badly. From which program, Bohr's or Lakatos' misconception, ought philosophical analysis begin?

What I am trying to suggest, in short, is that what Lakatos conceives as history is not history at all but philosophy fabricating examples. Done in that way, history could not in principle have the slightest effect on the prior philosophical position which exclusively shaped it. That is not to say that historical reconstruction is not intrinsically a selective and interpretative enterprise, nor that a prior philosophical position has no role as a tool for selection and interpretation. But it is to insist that, in the only sort of history which can hold philosophical interest, a prior philosophical position is not the only selective principle and also that it is not, as a selective principle, inviolate. When one's historical narrative demands footnotes which point out its fabrications, then the time has come to reconsider one's philosophical position.

V. HISTORY AND IRRATIONALITY

Why is it, I now ask in conclusion, that Lakatos feels the need to protect himself from real history? Why does he provide a parody in its place? My best guess is that he fears that history, if taken seriously as an independent discipline, may lead him to the position he attributes to me; the view that science is fundamentally an irrational enterprise. As a hypothesis about causes and motives that can only be a guess, and nothing very important depends on its being correct. But what his paper does make unequivocally clear is his belief that I have been led to defend irrationality by taking seriously aspects of history which he seeks a basis for omitting or rewriting.

As I have said before, both here and elsewhere, I do not for a moment believe that science is an intrinsically irrational enterprise. What I have

perhaps not made sufficiently clear, however, is that I take that assertion not as a matter of fact, but rather of principle. Scientific behavior, taken as a whole, is the best example we have of rationality. Our view of what it is to be rational depends in significant ways, though of course not exclusively, on what we take to be the essential aspects of scientific behavior. That is not to say that any scientist behaves rationally at all times, or even that many behave rationally very much of the time. What it does assert is that, if history or any other empirical discipline leads us to believe that the development of science depends essentially on behavior that we have previously thought to be irrational, then we should conclude not that science is irrational but that our notion of rationality needs adjustment here and there.

That position, so long as it remains abstract, is one with which Lakatos seems to agree. Whether or not he has managed it altogether correctly, the entire last half of his paper argues that historical study, properly done, can alter the line between the internal and external. In consequence, he says, it can change our notion of scientific rationality as well. Having taken that position, he may properly, of course, reject my views on substantive grounds; because I may have made historical, logical, or philosophical mistakes, as I doubtless have. What he may not do, but nevertheless does, is reject them *simply* or *merely* because my conclusions from history attribute an essential role to behavior he thinks irrational. Arguments of that sort contradict the core of his present methodological position.

So far I have argued the irrelevance of Lakatos' charge of irrationality on grounds of principle. Let me now try to make a similar point on substantive grounds. I began these comments by suggesting that Lakatos' present position has grown very close to my own. I shall close by suggesting that in key respects the parallel between our views goes even further than I then allowed. There are, I think, three main grounds on which charges of irrationality have been levelled at me. Two of these Lakatos now concedes, one explicitly, the other implicitly. The third he rejects in a footnote aside, ignoring in the process one of the most active and exciting areas in contemporary philosophy.

The first source, I think, of the charge that I make science an irrational enterprise is my insistence that the choice between paradigms (or theories, for present purposes) cannot be compelled by logic and experiment alone;

in these matters there is no such thing as proof, no point at which the opponent of a newer view violates a rule of science, begins to behave unscientifically. Lakatos makes exactly the same point repeatedly. "One may rationally stick to a degenerating program until it is overtaken *and even after*." [p. 104] "One *must* realize that one's opponent, even if lagging badly behind, may still stage a comeback." [p. 101] "No advantage for one side can ever be regarded as absolutely conclusive." [p. 101] If this be irrationality – as Lakatos has occasionally supposed in the past – then we are both guilty.

An even more frequent reason for the charge of irrationality has been my insistence that ultimately the choice between paradigms is a community decision, that what passes for proof, verification, or falsification in the sciences has not occurred until an entire community has been converted or re-formed about a new paradigm. On this point my views were not originally so clearly expressed as I should like, and they have in any case evolved since. What I should like to have said, however, is very close to what Lakatos now does say, though I am far from sure he realizes its consequences.

Throughout his paper Lakatos refers to the importance in scientific decision-making of what he calls a "code of scientific honesty" or a "code of scientific honor". [p. 92] When he distinguishes his position from the one to which he objects, he makes remarks like: "What one must *not* do is to deny [a research program's] poor public record", [p. 104] or "The scores of the rival sides... must be recorded and publicly displayed at all times." [p. 105] Elsewhere he speaks of answering colleagues's objections "by separating rational and irrational (or honest and dishonest) adherence to a degenerating research program." [p. 105]

Lakatos' views cannot, however, be distinguished from mine or anyone else's in this way. On the contrary, he and I come closest at just these points. Who does he suppose believes that science could continue if scientists were dishonest? If I have been defending the irrational, it has not been by defending lies. In fact, Lakatos' references to honesty, to a 'public record', or to a score that must be '*recorded*' and '*publicly displayed*' suggest that he too is thinking of theory-choice as a community activity which would be impossible unless public records of this sort were kept. When the individual may decide alone, nothing of the sort is needed. Finally, and most important, Lakatos' emphasis on a code of honor

carries him even further in the same direction, for a code consists of *values* not of rules, and values are intrinsically a community possession.

However obscurely presented, my own position has from the start been that the choice between theories (and also the identification of anomalies, a process which raises similar problems) has to be made by a very special sort of community; otherwise there would be no science. Much of what is special about such communities is, I have tried also to argue, the shared values of their members – they must prefer the simple to the complex, the natural to the *ad hoc*, the fruitful to the sterile, the precise to the vague, and so on – a very usual list. Without such values the community's decisions would be different, and something other than science would result. I have also argued, however, that these values do not carry with them a set of criteria sufficient to dictate unequivocally their application in concrete cases. To a considerable extent they are acquired from the study of examples of past applications rather than by learning rules about how they are to be applied. Two men who employ the same values when choosing between competing theories may therefore differ vehemently about which theory is to be preferred. Only the man who says for example – theory *A* is simpler than theory *B*; the two are in other respects the same; nevertheless I prefer *B* – only a man who makes decisions of that structure violates what Lakatos calls the scientist's code of honor.

I am left, I think, with only one other source for the charge that I make science irrational – my discussion of incommensurability, which Lakatos brushed aside in a footnote. Since the hour is late, and he has given me no handle, I shall attempt only the following rejoinder here: Anyone who supposes that the points at which Feyerabend and I have aimed in introducing 'incommensurability' into our discussions of theory-choice are either trivial or obviously mistaken must simultaneously brush aside much of the contemporary literature on radical translation. I cannot think that that should be lightly done.

Princeton University