A Look Back at the "Popper, Kuhn, Lakatos Debate"

JOHN WORRALL

Introduction

Thank you very much to Osvaldo and Miguel (and everyone else involved in the organisation of the conference) for the invitation to talk.

My intellectual career would never even have begun without the guidance, inspiration and support of Imre Lakatos, so It is of course a special honour to be celebrating the continuing influence of his work at this conference to mark the centenary of his birth. I am so pleased that Lakatos's work continues to have an impact around the world and in particular in Brazil.

I thought it would be interesting and appropriate for this occasion to look back at the "Popper-Kuhn-Lakatos debate" – the central focus of attention at one of the sessions of a famous conference held at Bedford College, London in the summer of 1965 and organised principally by Lakatos. The corresponding part of the proceedings – consisting of an initial paper by Kuhn contrasting his own view of the development of science with Popper's, then a series of commentaries on Kuhn's views and a final reply by Kuhn –, along with additional responses by Feyerabend and Lakatos (not delivered at the conference), were published in Lakatos and Musgrave (eds) *Criticism and the Growth of Knowledge*: a book which has of course been my guide in this retrospective enterprise.

Although interesting and appropriate, my journey into the past has, I must admit, not been without its disappointments. Neither Popper nor his colleague John Watkins in particular shows any real understanding of Kuhn's views. Popper takes it that Kuhn is applauding the role in science of dogma, of closed-mindedness to new ideas. Popper suggests that Kuhn is correct that dogma plays a role but argues that he is wrong to applaud it: normal science, which for Popper means work done while in the grip of a dogma or set of dogmas is 'hack science' and even "a threat to civilisation"! But this is a clear misunderstanding: the correct translation of Kuhn's image of science into a generally Popperian or testing framework is as a re-representation of Duhem's insight that what gets tested in science is not a single theory, like Newton' s theory of gravitation or Maxwell's theory of electromagnetism, but instead a usually very large theoretical system built around that theory. When an "anomaly" arises, that is a refutation, not of the single theory at its centre, but only of the whole theoretical system, it means only that at least one of the theories in that theoretical system is false and it is clearly no more "dogmatic" for a scientist to address the problem by holding on to the central theory and looking to modify one of the secondary or auxiliary assumptions, than it would be to insist (as Popper even sometimes seems to be claiming) that the scientist must give up the central theory. Indeed, Kuhn is surely correct that, given that over time the central theory develops around itself certain puzzle-solving techniques designed in particular to solve exactly the problems caused by anomalies, the "natural" first move for a scientist is to hold onto that central theory and exploit the available puzzle solving power rather than make a leap into the dark in search of some new paradigm. (Witness Adams and Leverrier holding onto Newton's theory despite the anomalous data from the planet Uranus which led to the discovery of Neptune - an episode that Popper himself elsewhere cites as one of the great successes in the history of science.)

So, some fairly central misunderstandings of Kuhn by Popperians. On the other hand, some of Kuhn's own contributions to the debates are, so it seems to me, equally disappointing. For one thing, he re-endorses a claim that has always seemed to me one of the most mysterious aspects of his famous book *SSR:* the claim that events like the discovery of X-rays or, still more surprisingly, the discovery of the planet Uranus count for him as "revolutions". Most readers had taken it that Kuhnian revolutions are the big events, involving at least some conceptual rather than merely empirical change: the Copernican, Newtonian, Relativistic or Quantum Revolutions; along with some other less large-scale but still radical changes like the switch from the corpuscular theory of light to the wave theory of light in the early 19th century or from phlogiston-based to oxygen-based chemistry in the 18th. But Herschel discovered Uranus simply through careful, one might say obsessional observation of the night sky: eventually noticing that what had previously been thought to be one of the fixed stars was in fact moving, moving of course very slowly against the background of the fixed stars. This was an entirely empirically-based discovery involving nothing more radical in terms of change of belief about the universe than the switch from the view that there are 6 planets orbiting the Sun to the view that there are 7 (Neptune and the subsequently "demoted" Pluto having, of course, not yet been discovered). If the discovery of Uranus counts as a "revolution", a "change of paradigm" then I am afraid that I lose all intuitive grip on the concept.

Moreover, Kuhn confesses that when he is asked of a certain theory-change in science whether it counts as a revolution or not, he "frequently finds [himself] at a loss for an answer". But surely the whole of *SSR* is premised on there being a sharp distinction between revolutionary change and normal science. Conceding that the distinction is blurred, seems to make the whole position uninterpretable.

Finally, the case that Kuhn presented in his London remarks for the thesis that the central theories involved in successive paradigms are "incommensurable" seems to me fundamentally and rather obviously flawed. The basis for the argument, so Kuhn explicitly asserts, is the alleged lack of a neutral language into which at least the empirical consequences of the two theories we are concerned to compare can be translated: "The point-by-point comparison of two successive theories" he writes "demands a language into which at least the empirical consequences of both can be translated without loss or change." Kuhn denies that this condition is met in cases of revolutionary change. To the contrary:

"In the transition from one theory to the next words change their meanings or conditions of applicability in subtle ways. Though most of the same signs are used before and after a revolution – e.g. force, mass, element, compound, cell – the ways in which [they] attach to nature has somehow changed. Successive theories are thus … incommensurable." (266-267)

Elsewhere, Kuhn claims that the Ptolemaic and Copernican theories are incommensurable because even such an (allegedly) "observational" term as "planet" changed its meaning in that revolution: the earth is a planet for Copernicus but not for Ptolemy, while the sun is a planet for Ptolemy but not for Copernicus! But of course 'force' 'planet' and especially 'mass' are theoretical terms, not observational ones as Kuhn suggests, and so no wonder that their meanings change somewhat alongside theory-change. But surely Kuhn did not dig deep enough in the search for a theory-neutral comparison between the consequences of the pre- and post-revolutionary theories. *Obviously* 'force' 'mass' and 'planet' carry theoretical content. However, it is straightforward to go to a lower more observational level and compare how the two theories involving those notions fare when tested at that at least theory-neutral level. Instead of planets, we can talk about spots of light in the night sky and whether or not they move relative to one another. Or, given, for another example, that relativity theory asserts that a body's mass can be increased simply by accelerating that body, while in classical physics of course the mass of a body is constant, the notion of mass did indeed radically alter in the shift from classical to relativistic physics, but that does not mean that we cannot compare classical and relativistic physics in terms of what they predict about the – observable or at least theory-neutral – shifts in visible fringe patterns in the Michelson-Morley experiment, about the observable – or at least theory-neutral – apparent motions of Mercury, or about the observable or at least theory-neutral tracks in a cloud or bubble chamber.

Main Content

But laying these disappointing aspects aside there is still much of fascination particularly in what has always been seen at the central point of the debate: the issue of the rationality (or otherwise) of theory-change in science. Most of us, I suppose, start out from the Enlightenment view that modern science has enabled humans to unlock the secrets of the universe. As the English poet, Alexander Pope, famously put it "Nature and Nature's laws lay hid in night, God said 'Let Newton be' and all was light." Most of us start from the position that, where successful, as it surely has been in physics and elsewhere, science has told us more and more about the structure of the universe; from which it follows that, where accepted theories have changed (in "mature science"), the change has been from one good theory to a still better one. So that, unlike say, changes in fashion or in art movements, changes in science exhibit rationally accredited *progress* rather than mere change.

In his book *The Structure of Scientific Revolutions,* Kuhn seemed to many commentators to be challenging that rationalist picture. According to his account, scientists generally work within their paradigms, believing that any experimental or observational anomalies will eventually be dealt with within that paradigm. However sometimes anomalies accumulate and prove recalcitrant to attempts to "normalise" them. Eventually, a feeling of "crisis" affects the scientific community built around the paradigm; but there are no rules about how many anomalies or how recalcitrant the anomalies have to prove to be to justify the feeling of crisis. There is, as Kuhn explicitly and repeatedly states, no criterion higher than the community view and the community either feels a sense of crisis or it does not. If it does, then it will switch to another paradigm: that conversion being analogous to a religious conversion rather than anything objectively rule-governed. Furthermore, the conversion will never be complete: in any revolution there are hold-outs, usually elderly scientists who have made significant contributions to the older, pre-revolutionary paradigm who stick to that older paradigm; moreover, according to Kuhn, those hold-outs cannot be judged to be mistaken or to be failing to make the rational choice. They simply lose the vote, so to speak, and thus eventually either die or define themselves out of the relevant scientific community. Finally, because methodological standards are also subject to change in revolutions alongside the theories endorsed by the paradigm, there is no neutral basis on which we can judge the theories endorsed by the new paradigm as better than those endorsed by the old.

Lakatos echoing other philosophers such as Scheffler and Shapere, but expressing it more abrasively, of course claimed that this account of theory-change in science by Kuhn reduced it to a question of "mob psychology". Lakatos famously claimed that his Methodology of Scientific Research Programmes, while conceding that many aspects of the process of science are better described by Kuhn than by Popper (for example, experimental difficulties in science *are* treated like Kuhnian anomalies, rather than Popperian refutations), saves the rationality of science from Kuhnian relativism by – allegedly – showing that one theory, or rather in Lakatos's terms, one research programme is only ever replaced in science by one that is objectively superior to it in terms of how it objectively stands up to the evidence.

What progress was made in this debate during the 1965 Bedford College discussions? Well, the first thing to be noted is that Kuhn, in responding to his critics, heatedly denied the "mob psychology" charge and in effect insisted that his account is in no need of any injection of rationality from Lakatos: it is already an account that involves rationality. "Does anything in [my] argument" he asks "suggest the appropriateness of phrases like decision by 'mob psychology'? I think not." (p.262-3) Indeed, he continues, "no part of …[my]… argument implies that scientists may choose any theory they like so long as they agree in their choice and thereafter enforce it" (263) Far from the adoption of a new paradigm being, on his account , "mystical" or purely sociological, that account insists that there are "good reasons for theory choice" or better, that good reasons are involved in choosing the new theory/paradigm (261). Moreover, says Kuhn, "these are … reasons of exactly the kind standard in philosophy of science: accuracy, scope, simplicity, fruitfulness and the like".(261)

So did Kuhn in effect claim that, when his views are correctly understood, there is no real issue between him and what we might call his objectivist critics? No: there still is a difference - one that seems at least to be a major one - and it lies in what Kuhn's account *denies*:

"What I am denying … is neither the existence of good reasons nor that these reasons are of the sort usually described. I am, however, insisting that such reasons constitute values to be used in making choices rather than rules of choice." (261)

This has the practical consequence that "Scientists who share [the 'good reasons'] may nevertheless make different choices in the same concrete situations." Reason - in the form of the "objective factors" of traditional philosophy of science (empirical accuracy, simplicity, "and the like") - certainly plays a role but it never dictates the switch to the new paradigm. Consequently, on Kuhn's view, it is never actually irrational to resist the switch to that new paradigm. There is, claims Kuhn, no "point at which resistance becomes illogical or unscientific". An "elderly" holdout, like Priestley "holding out" for phlogiston against Lavoisier's oxygen theory, may infuriate his colleagues by his stubbornness but cannot legitimately be regarded as mistaken or "irrational".

As just indicated, Kuhn's claims about the failure of holdouts to be irrational were already to be found in *Structure* but in his Bedford College replies to critics he is more explicit about the arguments behind those claims and in particular the relative roles of the good reasons (or "objective or shared factors") and other, "subjective" (or individual) factors in theory choice.

Kuhn cited two ways in which reason in the form of the allegedly standard factors may fail, and invariably or almost invariably *does* fail to dictate a particular choice of theory or paradigm. The *first* is that two separate 'good reasons' or 'objective factors' may point in different directions: one of them indicating a preference for theory 1 over theory 2, and the other a preference for theory 2 over theory 1. (An alleged example that Kuhn cites more than once is that, at the time it was adopted by Kepler and Galileo, Copernican theory was simpler than Ptolemaic theory; but on the score of detailed empirical accuracy, the Ptolemaic theory was better.) Kuhn writes:

"In many concrete situations, different values, though all constitutive of good reasons, dictate different conclusions, different choices. In such cases of value conflict (e.g. one theory is simpler, the other is more accurate), the relative weight placed on different values by different individuals can [legitimately] play a decisive role in individual choice." (262)

The *second* way in which 'good reasons' may fail to determine a choice of theory, according to Kuhn, is that individual scientists may – again legitimately on his view – come to different judgments about how an *individual* objective factor applies in a particular case of theory-choice. He writes:

"More important, though scientists share these values [dictated by 'good reason'] and must continue to do so if science is to survive, they do not all apply them in the same way. Simplicity, scope, fruitfulness and even accuracy can be judged quite differently (which is not to say they may be judged arbitrarily) by different people. Again, they may differ in their conclusions without violating any accepted rule."

It is noteworthy that, while as indicated, Kuhn does cite at least alleged examples of the first kind of indeterminacy – two objective factors pointing in different directions – he cites no examples of the second kind – one objective factor being "interpreted differently" and being *reasonably* interpreted differently.

But, be that as it may, Kuhn's account is definitely, then, at any rate somewhat more nuanced than his early objectivist critics were allowing. The charge of making scientific theory-change a matter of mob psychology does not stand - at any rate not without further elaboration aimed at showing that Kuhn's account still makes scientific theory-change an irrational affair despite this insistence that objective factors play an ineliminable role in such theory-changes.

Lakatos in fact never replied to this more nuanced Kuhnian account. Of course, there was no opportunity for such a reply within the structure of the Bedford Colloquium debate and hence within the structure of the book *Criticism and the Growth of Knowledge*. That debate - both at the conference and in the book - began and ended with Kuhn; and Lakatos did not "cheat" by presaging Kuhn's reply within his own contribution (which of course contained the most important development of his Methodology of Scientific Research Programmes). It is, nonetheless, perhaps a bit surprising that Lakatos never took the opportunity to reply later – though admittedly, and sadly, he did not have very long in which to do so: Kuhn submitted his 'Replies to Critics' only just before *Criticism and the Growth of Knowledge* was published in 1970 and Lakatos died in February 1974.

For the rest of this presentation, I will try to make good on this omission by speculating on how Lakatos might have responded had he ever directly confronted this more elaborate Kuhnian view.

Well, I have no doubt that Lakatos's first reaction would have been that the charge of "mob psychology" or, to put it less tendentiously, of sociologising theory-choice in science still stands – indeed that the charge is underwritten by Kuhn's elaborated account. But to see why, we have to get clear about what exactly it is that Lakatos expects from a methodology that Kuhn's account, even in this elaborated form, fails to yield.

The answer is that a methodology needs to produce an objective ordering of theories in the light of the empirical evidence, on ordering that – at least in all normal cases – places the winning side in any scientific "revolution" higher than the deposed theory/research programme. It is crucial here to distinguish the objective ordering that Lakatos sought and which exists in Plato's or Frege's or Popper's "World 3" from any issues about individual scientists' beliefs or their decisions about which theories to work on. In so far as Kuhn's account is that theory-preference *in this sense* is always dependent on individual or subjective factors as well as the shared or objective factors then it delivers the verdict, for instance, that wave optics was not objectively, scientifically superior to corpuscular optics when that revolution occurred in the early 19th century or that the relativistic programme was not objectively superior to the classical physics programme when that revolution occurred in the early 20th century. And so, for Lakatos, Kuhn's account of theory-change remains an irrationalist one.

As a matter of fact, I am not sure that Kuhn, who was never really a philosopher, had much idea of this Fregean/Popperian sense, but let's lay that aside until later and assume, in order to identify what was at stake between Lakatos and him, that Kuhn's account, even when elaborated, yields this consequence that there is no objective preference ordering of the post- and pre-revolutionary theories/programmes/paradigms.

Of course however much Lakatos might dislike this consequence, if Kuhn's account were correct then he would have to live with it: if "theory-choice", Kuhn' s term, means theory-preference in Lakatos's and really does depend on a mixture of objective and subjective/shared or individual factors in the way that we are interpreting Kuhn as claiming, then although Kuhn would be right that 'rationality' in the form of the objective factors plays an ineliminable role, it would not be a decisive role: psychology and social psychology/sociology intrude in what these objectivists would like to see as a purely logical realm.

So Lakatos cannot simply resort to mob psychology style name-calling and needs to question Kuhn's elaborated account. He needs to argue that objective factors, *properly understood*, always do determine theory-preference in cases of scientific revolutions; and hence that there must be something wrong with Kuhn's account of objective factors, which entails that they fail to provide such determinate theory-preferences. And here I think Lakatos is on solid ground.

Kuhn, remember, insists that his objective factors are ones "standard in philosophy of science", and he lists them several times as "accuracy, scope, simplicity, fruitfulness and the like". And he does elsewhere add "consistency (both internal and with other accepted theories)" as a further objective factor. Contrary to his claims that these criteria are "standard in the philosophy of science", I in fact know of no philosopher of science of an objectivist kind (that is, one who holds that there are objective criteria of theory-acceptance in science) who would endorse all the items on Kuhn's list as it stands and none who would be happy to leave any such list unstructured as Kuhn does, rather than attaching differing degrees of importance to them.

For Lakatos, there is, of course, a dominant criterion which does not even appear on Kuhn's list of "objective factors" (at least it doesn't appear explicitly): namely independent testability and predictive success. This is, for Lakatos, *the* criterion of a progressive research programme: a programme is progressive if, and only if, successive theories produced by it make testable predictions, independent of any empirical results used in the construction of those theories; and at least some of the time those predictions are empirically verified.

Unlike predictive success, which clearly requires the cooperation of Nature, empirical accuracy and scope, which *are* on Kuhn's list, are readily manufactured by scientists: once they know the facts, scientists can readily find a place for them in some system based on any central theory you care to specify. This is a consequence of Duhemian 'underdetermination of theory by data'. If one paradigm shows greater empirical accuracy or scope than another, this, then, is standardly a merely historically contingent state of affairs reflecting only the lengths of time that the two paradigms have been worked on. Hence, contrary to Kuhn's view, empirical scope/accuracy supplies on its own no telling reason to prefer one paradigm over the other.

For example, the greater empirical accuracy and scope of Ptolemaic theory in the years shortly after the publication of Copernicus's *De Revolutionibus* is cited by Kuhn as an 'objective reason' to choose Ptolemaic theory over Copernican. In fact, however, that greater empirical accuracy and scope is no surprise: it is provable (and anyway obvious) that all the empirical astronomical data – the apparent motions of fixed stars, sun and planets - can, with sufficient ingenuity, be fitted within *either* a heliocentric (more accurately, helio*static*) system *or* a geocentric (again more accurately, geostatic) system. Ptolemy started to plot apparent astronomical motions within his geostatic system in the 2nd century AD (and the roots of the geostatic approach go back still further in the Greek, Roman and Babylonian traditions). Copernicus's system, by contrast, was published shortly before his death in 1543. It is therefore no wonder at all that, when Galileo and Kepler began to think about these matters, the Ptolemaic system was ahead in terms of the number and accuracy of the phenomena it could bring within his system.

Kuhn's explicit view, remember, was that the choice between the Ptolemaic and Copernican systems at the time of Kepler and Galileo was not determined by the "objective factors" because while empirical accuracy/scope told in favour of Ptolemy, simplicity (in, as Kuhn puts it, "a special sense") told in favour of Copernicus. But, as we just saw, empirical accuracy/scope carries no weight in underwriting any preference.

As for the special sense of simplicity that Kuhn refers to, this again in fact reflects Lakatos's supreme predictive success criterion. "Simplicity" and "unity" - in the scientifically important sense of these terms - are closely related to predictive success. We surely have no clear-cut intuitions about when one *basic* theory in science is simpler than a rival. Ahead of any detailed elaboration is, for example, the basic idea of a fixed earth simpler or more complicated than the basic claim that the sun is fixed? Or is the idea that light consists of material particles more or less simple than the idea that it consists of waves in a medium? I don't see the slightest reason to think that there's an answer either way. Where we *do* have clear-cut intuitions is in cases where one basic theory has been so hedged around with qualifications and split into so many unrelated subcases that it clearly becomes too complex, not sufficiently simple, to be scientifically acceptable. But, in all such cases, the complexity and disunity have been introduced under the pressure of initially independent or recalcitrant experimental results. The basic theory has enjoyed no predictive success: it has either turned out to be silent about some phenomenon clearly in its field, or, more often, turned out to yield an incorrect prediction. Special cases and exceptions have therefore had to be introduced to accommodate the facts - at the cost of increased complexity and decreased unity. In short, the theory's becoming complex means, in Lakatosian terms, that the associated research programme has degenerated. A theory's remaining simple and unified means that the associated research programme has progressed.

This is the 'special sense of simplicity' that so impressed Kepler and Galileo about Copernicus's theory: phenomena such as planetary stations and retrogressions, or the bounded elongation of Mercury and Venus had to be "worked into" the Ptolemaic theory courtesy of special assumptions designed exactly using features of those already known phenomena – principally of course assumptions about epicycles. By stark contrast, the phenomena of Stations and Retrogressions and of Bounded Elongation (and also of the order of the planets in terms of distance from the central, fixed body) fall naturally out of the Copernican approach, - they follow from the basic model, without the need for any special *ad hoc* assumptions.

The two other items on Kuhn's unstructured list of objective factors in theory choice are "fruitfulness" and consistency. Under the only precise sense I can make of it, fruitfulness too is intimately connected to simplicity and hence to predictive success (and hence to progressiveness of a research programme). A general theoretical approach (a paradigm or research programme) shows its fruitfulness by supplying ideas for developing specific theories independently of empirical results or as a response to such 'anomalies' as may arise. Such an approach will be judged instead barren or lacking in fruitfulness (as Lakatos put it, the research programme's "heuristic" will have "run out of steam") only when all these ideas have been tried without predictive success; and hence the approach has been reduced to tagging along behind the empirical data, always accommodating that data *post hoc* rather than predicting it in advance.

By the early to mid-1830s, for example, the emission or corpuscular approach to optics had very definitely proved barren – its former "fruitfulness" lay exhausted. The ideas supplied by the general claim that light is a Newtonian particle had all been tried in the attempt to produce specific theories that dealt with optical phenomena and dealt with them successfully, that is by predicting at least some of them. Particles are, of course, subject to forces; forces could be attractive or repulsive: all the apparent deviations from rectilinear propagation of light - reflection, refraction, interference, and diffraction - *might* be explained by having ordinary "gross" material objects exert forces of various kinds on the light particles. (This was, in essence, the corpuscularist or emissionist programme.) The idea that the "particles" of light are strictly *point* particles always had to be an idealization; so the finite dimensions of the real particles might come in useful: it might for example be assumed that the particles have sides or poles and revolve with respect to these poles as they move along, affecting the way they react to the various forces supposed to be exerted on them. Various isolated results could be explained (in *very* rough terms) on the basis of these assumptions - but, when it came to anything like details, the "natural" assumptions about the forces and the polar revolutions unambiguously failed and instead the required theoretical assumptions had always to be "read off" the already given facts. There was never any correct prediction of a different phenomenon. Instead, each new phenomenon required further elaboration of the theoretical assumptions (perhaps another complication in the field of force set up by the diffracting or refracting body or yet another axis of revolution in the particles). As the optical scientist Humphrey Lloyd put it in a famous report on the "Progress and Present State of Physical Optics" produced in 1833:

"An unfruitful theory may . . . be fertilized by the addition of new hypotheses. By such subsidiary principles it may be brought up to the level of experimental science, and appear to meet the accumulating weight of evidence furnished by new phenomena. But a theory thus overloaded does not merit the name. It is a union of unconnected principles. . . . Its very complexity furnishes a presumption against its truth. . . . The theory of emission, in its present state, exhibits all these symptoms of unsoundness ... "(1833, p. 296)

By contrast, there existed within the general wave theoretical approach at the same time some hopeful lines of attack on the problems it faced. One such problem emphasised right from the beginning by its opponents, concerned the phenomenon of prismatic dispersion. According to Fresnel's initial theory, the amount of refraction a ray of light undergoes when entering a transparent substance should be dependent only on the refractive index of the substance (or, more properly for the wave theory, the refractive index of the ether as structured within that substance). Hence it entails that if a ray of sunlight enters a transparent body it will be refracted as one ray. But in fact of course the sunlight when, for example, passed through a glass prism spreads out into the familiar spectrum – a phenomenon that had been extensively studied by

Isaac Newton, as reported in his *Opticks*. But this initial version of the wave theory of light was based on a very simple theory of the ether - one that involved the assumption that its parts *strictly* obey Hooke's law of the direct proportionality of restoring force to displacement. It was known from studies in mechanics, that not all vibrations in all substances strictly obey Hooke's law and several general ideas were already available concerning how a somewhat more sophisticated theory of the luminiferous ether involving a slightly more complicated expression for the restoring force could be constructed that might yield dispersion. Though none of these had yet borne unambiguous fruit, equally they had not all unambiguously run into sand. This is just a fact about the wave approach: it already possessed potential explanatory resources with respect to dispersion that had not been exhausted.

Moreover, again in contrast to the corpuscular programme, the wave theory of light in the 1830s already had an impressive record of success – in the form of shifts of theory that had proved significantly predictively successful. Wave theorists before Fresnel had all assumed that the ether is an extremely rare and subtle fluid - how else could the planets move so freely through it? It is a theorem of mechanics that fluids transmit only longitudinal (sometimes called pressure) waves. (Longitudinal waves are ones in which the particles of the medium oscillate in the same direction as the overall transmission of the wave through the medium; an example being a sound wave in air.) Fresnel's own initial theory was indeed that light is a longitudinal wave. However, he and his colleague Arago then established experimentally that if, say, the two beams emerging from the two slits in the double-slit experiment are polarized at right angles to one another (by passage through suitably oriented crystal plates), then the interference fringes disappear. It seemed that light beams polarized in mutually orthogonal planes fail to interfere (or, rather, fail to produce interference fringes). Neither Fresnel nor any other wave theorist had, at this stage, any coherent theory of the polarization of light. But, so long as the light waves were assumed longitudinal, the precise account of what happened when light is polarized could make no difference. Assuming that the wave theory is at all correct, the longitudinal assumption alone means that the disturbances in the two coherent and near-parallel beams (the slits are, remember, very close together) must themselves be near parallel and hence must alternately interfere constructively and destructively for different path differences. The Fresnel-Arago experiment which resulted in no interference fringes, therefore, put the wave theory into deep trouble. Fresnel took a still deeper breath and switched to the *transverse* wave theory: to the theory that the ether particles oscillate *at right angles* to the direction of the propagation of light. This yields an easy theoretical account of the process of polarization: the disturbance in an unpolarized beam has components in all planes through the direction of propagation; polarization (linear or plane polarization, that is) consists in restricting the disturbance to one such plane. This explained the apparent "sidedness" of polarized beams, and also explained the Fresnel-Arago results. The oscillations in beams that are polarized orthogonally are assumed themselves to be orthogonal. Hence, although the two sets of oscillations certainly interfere or superpose - to produce (in general) elliptically polarised light - they operate at right angles rather than along the same line, and hence can never destructively interfere so as to produce fringes. Although it straightforwardly dealt with this difficulty over polarized light, the switch to the transverse theory certainly required a deep breath. This was because elastic media can transmit such waves only if they exhibit resistance to sheer, that is, only if they are solids. But how could the planets move completely freely through an elastic solid ether? But whatever the conceptual difficulties, Fresnel's new transverse theory scored stunning empirical successes. Not least when Hamilton showed in 1830 that the transverse theory entails the hitherto entirely unsuspected phenomena of internal and external conical refraction -predictions that were dramatically confirmed by Humphrey Lloyd in 1833.

So "fruitfulness" is again unambiguous: the wave theory was fruitful - it specified avenues of research that had not yet been exhausted and it had a track record of change accompanied by predictive success; the corpuscular theory was not fruitful – it had no unexhausted avenues of development and no track record of predictive success. And again one of Kuhn's "objective factors" – this time "fruitfulness" - crucially involves, when analysed, predictive success.

The remaining item on Kuhn's list is "consistency (both internal and with other accepted theories)". Well, internal consistency is obviously a logical requirement – no one can "choose" (to use Kuhn's term) an inconsistent theory since that theory, by definition, contradicts itself: so by accepting it you would also be accepting its negation! (Lakatos has some deep-sounding but in fact rather sloppy remarks about scientists sometimes fruitfully proceeding on "inconsistent foundations" but this always means that those scientists are at least dimly aware of how any inconsistency can be rectified and confident that their positive results will be recoverable within the consistent version of the theory.) So, the interesting question is whether or not consistency between some new theory and ones that are already "established" should be treated as an objective factor in theory-choice, a theoretical virtue that can legitimately count in favour of preferring that theory.

Kuhn of course asserts that it does and again cites the Copernicus/Ptolemy case as one in which it plays a role: while 'simplicity in a special sense' counted in favour of Copernicus, not only did empirical scope/accuracy, according to Kuhn, justify a preference for Ptolemy, as discussed earlier, so also did the fact that Ptolemy was consistent with other theories considered well-established at the time – notably Aristotelian physics, while Copernican theory was clearly inconsistent with that physics. But surely this inconsistency was a *virtue* of the Copernican theory, not a vice. The inconsistency supplied an interesting and demanding problem for further research duly addressed by Galileo and later by Newton, indicating the need to develop an alternative physics to that of Aristotle. Of course, this judgment is premised on the fact that Copernican theory was predictively successful (with, as we saw, planetary stations and retrogressions and the bounded elongations of Mercury and of Venus), while Aristotle's physics had become "well-established" despite never enjoying any such predictive success. Scientists in general do, no doubt correctly, downgrade (or more usually ignore) new theories that clash with well-established ones – but only when there is no independent evidence for the new theory. To take a relatively trivial but illustrative example, the theory that homeopathic "remedies" are effective (really more effective than placebo) is multiply inconsistent with accepted theories in physics. This fact is correctly taken as strong evidence against homeopathy, but only because those theories in physics are supported by predictive successes while the hypothesis that homeopathic "remedies" are effective has no such support. On the other hand, if I am right that Aristotelian physics had only historical, but no evidential legitimacy, then it follows that it is the preferred candidate for replacement given its inconsistency with the predictively successful Copernican theory, not vice versa: just as Kepler, Galileo and Newton recognised. It is predictive success that flips inconsistency with other accepted theories over from a vice to a virtue.

In sum, then, all the 'objective' criteria that Kuhn cites, either play no real role in theory-preference or reduce to Lakatos' single criterion of progressiveness. On Lakatos's account, in stark contrast to Kuhn's, there is essentially only one criterion of scientific merit and hence there is no possibility of the sort of clash between different objective criteria of scientific merit that Kuhn's London account sees as requiring subjective factors of theory "choice" to resolve. Moreover, that one criterion is definite: either a research programme makes independently testable predictions some of which are confirmed or it does not. So again the space is not there for subjective factors to play a role – this time in resolving varying applications of single objective factors.

Notice that it is not true that Lakatos's account *always* underwrites a preference in any dispute between two theories/research programmes. In particular, it might well be the case at some stage in the history of science that neither of two competing programmes is progressive – this was for instance true of the debate between Newton's and Hooke's approaches to optics in the mid- to late- 17th century: neither Newton nor Hooke could do any better than accommodate already known phenomena post hoc within their preferred framework. And so the choice between those two frameworks at that time was indeed subjective. But no revolution occurred in optics in the mid- to late-17th century – the scientific community was divided between the two available theories, and when the revolution did occur in the early 19th century, the wave programme now led by Fresnel was definitely progressive, while the corpuscular programme had definitely degenerated. So, in contrast to Kuhn's account, Lakatos's approach yields a rationalist explanation of the development of science: every change, every scientific revolution has constituted progress - the theory, or rather research programme, displaced in the revolution had degenerated while the new, superseding theory/research programme had proved progressive.

So, this looks like the end of story: even on the amended version that he developed in his London remarks Kuhn's account does make scientific theory change too subjective an affair for the tastes of objectivist philosophers, but Lakatos produced an alternative account that is equally sensitive to the history of science but avoids the problems by restoring the objectivity of theory change.

I wonder, though, for all the heat that it generated, how much, in retrospect, was really at stake in this debate. Lakatos is quite clear, especially in his PSA 'Replies to Critics' paper of 1970, that his objective appraisals of the current merits of rival programmes in the light of evidence have no consequences either for scientists' beliefs about which theory, if either, is true or, more significantly for current purposes, any consequences for which theory/programme it is rational to work on. He writes:

"… my methodology …. only appraises fully articulated theories (or research programmes) but it presumes to give advice to the scientist neither about how to arrive at good theories nor even about which of two rival programmes he should work on."

And he goes on to emphasise:

"… when it turns out that, on my criteria, one research programme is 'progressing' and its rival is 'degenerating', this tells us only that the two programmes possess certain objective features but does not tell us that scientists must work only in the progressive one. (Indeed, as I constantly stress, degenerating research programmes can always stage a comeback … But this would, of course, be impossible if no scientist 'worked' on the programme.)"

My guess is that Kuhn found it difficult to see any content at all in Lakatos's "objective appraisals" if they have no connection to what scientists should and should not do in particular choice-situations. (Indeed Kuhn's use of the term 'theory-*choice*' reflects the fact that he is entirely focussed on scientists' decisions – where decisions are denizens of Popper's psychological 'world 2' rather than objective, logical World 3). Indeed Kuhn explicitly states that if Lakatos has no advice for scientists then "he has told us nothing at all'. And Paul Feyerabend famously adopted the same position:

"Scientific method, as softened up by Lakatos, is but an ornament which makes us forget that a position of 'anything goes' has been adopted."

Are Kuhn and Feyerabend correct? Well, the issue of the connection, if any, between Lakatosian objective appraisals and the rationality of scientists is certainly not straightforward. For sure, no one should ever have supposed that the connection is simply that it is rational to work on a research programme if and only if it is progressing. That thoroughly naïve rule would pronounce the great innovators in science "irrational". Was the wave optics research programme progressive when Fresnel first chose to work on it? Of course not, it was Fresnel's work on it that made it progressive. Was the relativity programme progressive when Einstein first chose to work on it? Of course not, it was Einstein's work on it that made it progressive. And so on.

On the other hand, if, as Lakatos insisted, the only thing that a scientist in a situation of choice between two programmes needs to do to count as rational is to acknowledge the "current score" between the two and then can choose to believe, and more importantly for present purposes, choose to 'work on' the programme with the lowest score (the one that is degenerating), then this does indeed seem to provide – to say the least - only a very thin theory of rationality. Most philosophers of science have taken it that in order for a methodology to count as one that makes the actual development of science a rational process, it not only has to yield the consequence that those scientists who accepted the new theory in a revolution were right, but also the consequence that the "hold outs" who continued to adhere to the old theory were wrong. One of the central reasons for counting Kuhn's view as unacceptable for a rationalist about science was the fact that it delivers the verdict that the hold-outs were not wrong or irrational.

However, depending on exactly what is meant by "continuing to adhere" to the older theory, it seems that Lakatos' s view may share this feature of declaring the hold outs not to be mistaken or irrational. All that Priestley needs to do to count as "Lakatosian-rational" is to admit that Lavoisier's oxygen theory was at his time ahead in terms of the support it receives from the phenomena, but then go on to insist that he will continue to work on the phlogiston approach with the intention of turning the evidential tables and making the phlogiston theory the better evidentially supported theory.

Of course, scientists in the history of science did not express themselves in explicitly Lakatosian terms, and I do not know enough about Priestley to judge whether he would have been willing to make this concession (though I suspect, since it is such a small concession, that he surely would have). However, there is another hold out against a scientific revolution whom I do know well, having studied his work in depth; and, so far as he goes, the situation is clear.

The "hold out" I refer to is David Brewster. Brewster was a significant optical scientist of the early to mid-19th Century. He was the discoverer of a great many of the properties of polarized light, especially elliptically polarized light; he discovered "Brewster's law," relating the polarizing angle and refractive index of transparent substances; he discovered a whole new class of doubly refracting crystals, the "biaxal crystals"; he discovered that ordinary unirefringent transparent matter can be made birefringent by the application of mechanical pressure; and he discovered the then unknown general phenomenon of selective absorption.

As well as a significant scientist, Brewster was certainly some sort of holdout for the corpuscular or emission theory of light – even though Fresnel's earlier work had made the wave optics programme unambiguously progressive on Lakatos's criteria, while the corpuscular programme had, by Brewster's time, unambiguously degenerated.

In 1831, Brewster presented a "Report on the Present State of Physical Optics" to the British Association for the Advancement of Science, in which he asserted that the undulatory theory is "still burthened with difficulties, and cannot claim our implicit assent," (Brewster 1833a, 318). And in 1833 he reported: "I have not yet ventured to kneel at the new shrine [that is, the shrine of the wave theory] and I must acknowledge myself subject to the national weakness which urges me to venerate, and even to support, the falling temple in which Newton once worshipped." (1833b, 361) (The corpuscular programme was of course traditionally regarded as having been invented by Newton.)

Brewster believed that, despite all the difficulties that had mounted against it, there was life left in the Newtonian emissionist theory. He echoed Herschel's sentiment expressed some ten years earlier that, were sufficient talent and energy invested in the emission theory, it might yet turn the tables of scientific superiority on its undulatory rival.

But Brewster, while continuing to recommend work on the corpuscular programme, did very definitely accept that the wave optics programme was unambiguously ahead in terms of the objective support it received from the empirical data. He wrote, for example

"I have long been an admirer of the *singular* power of this [wave] theory to explain some of the most perplexing phenomena of optics; and the recent discoveries of Professor Airy, Mr Hamilton and Mr Lloyd afford the finest examples of its influence in predicting new phenomena." (1833b, 360; my italics)

The reference to Lloyd and Hamilton here concerns the episode already mentioned: Fresnel switched from the longitudinal to the transverse theory; Hamilton showed that the transverse theory entails the entirely unexpected phenomena of conical refraction; and finally Lloyd experimentally verified those predictions.

So, Brewster the hold-out would definitely have counted as rational on Lakatos's view: he accepted the "objective [current] score" was in favour of his wave opponent, but, reflecting Lakatos's concession that "degenerating research programmes can always stage a comeback", Brewster continued to encourage work on his favoured corpuscular approach.

So, what is the conclusion of this long and rather convoluted story? I have imagined Lakatos and Kuhn continuing their debate starting from Kuhn's replies to critics at the Bedford College Colloquium. I have argued that, although it might seem that Kuhn's insistence that objective factors always play a role in what he calls "theory choice" was a conciliatory move in the debate, Lakatos could in fact successfully argue that Kuhn misidentified the objective factors: there is in fact at root only one objective factor – progress and degeneration. And that objective factor always pronounces the winning theory in any case of scientific change objectively superior to the displaced theory.

However, Kuhn clearly regarded Lakatos's objective theory preference as in effect just so much hot air. If we concentrate on what Kuhn held really matters so far as rationality is concerned, namely scientists' beliefs and their consequent decisions about which paradigm/programme to try to develop, then Lakatos, through his concession that it is always possible for a degenerating programme to make a comeback and turn the evidential tables on its rival, automatically further conceded that subjective or individualist factors always play a role in any decision about which theory a scientist "chooses".

Fresnel and many others regarded the wave optics programme as progressive and therefore superior to its degenerating corpuscularist rival, and chose to continue to work on it in the attempt to make it even more progressive. Brewster and a *very few* others, accepted that the wave optics programme was progressive and the corpuscular programme degenerating and so accepted that the wave programme was, in Lakatosian terms, objectively superior as things currently stood, but nonetheless chose to work on the corpuscular approach in the hope, perhaps even expectation, that it would eventually become even more progressive than its rival. Both Fresnel and Brewster were perfectly rational according both to Kuhn *AND – more surprisingly to Lakatos*. It seems in the end, and rather disappointingly, that both Lakatos and Kuhn were right.

Having meticulously stuck to the task of interpreting the debate between Kuhn and Lakatos, let me end by indicating my own view – albeit very briefly. While disagreeing with Kuhn that Lakatos's objective appraisals telling us "nothing" if not connected at all to advice to scientists, I think that Lakatos could have done better: I do not agree with his claim that his appraisals, when properly articulated, do not have any consequences concerning scientists' decisions about which programmes to work on. More particularly, I do not agree with Lakatos's famous remark that 'degenerating research programmes can always stage a comeback'. Sometimes it is correct to allow that they might stage a comeback; sometimes it clearly is not.

To see why, let's return to my favourite 'hold out' Sir David Brewster and his view that the monopoly enjoyed in his time by the wave optics programme was a mistake and that if sufficiently many, sufficiently talented scientists worked on the corpuscular programme, it could turn the evidential tables, "stage a comeback" in Lakatos's phrase. The fact is that it is completely unclear what "working on" the corpuscular programme in the 1830s would have involved. The crucial thing, and what after all was meant to be special in analysing science in terms of research programmes rather than just theories, is the *heuristic*. The heuristic of the corpuscular programme was essentially to exploit the already developed mechanical theory of particle motion. This heuristic therefore supplies an array of factors whose variability might be exploited to explain optical effects: masses and velocities of the different particles of light, suppositions about the forces acting on those particles in different situations, de-idealisations from point particles to particles with finite dimensions, perhaps with something like magnetic poles. All these ideas had been tried and had not even moved the programme toward anything like adequate theories of basic phenomena such as reflection, partial reflection and refraction, let alone diffraction and interference. No predictive success had been scored and none was remotely in sight. The heuristic was objectively exhausted: there were no ideas left to try. A scientist who followed Brewster's advice to choose to work on the corpuscular optics programme in the early 19th century would, therefore, be acting like the people I often seem to get stuck behind in queuing to get on the Underground in London: these people put their travel cards in the card-reader at the entrance gate, and the card is rejected – the gate does not open, so they insert their card again and again the gate does not open, and again and.. all the time expecting that something different will happen. A corpuscularist in the 1830s would have been in the same, surely irrational position.

So, a much less "thin" account of the rationality of theory-change in science than the one officially endorsed by Lakatos in his 1970 "Replies to Critics" can, I think, be developed; but can be developed using his ideas – the criteria of progress and degeneration are of course involved, but so also should be the crucial, but underdeveloped idea of the heuristic appraisal of programmes (which he alluded to many times in his work but seemed to be forgetting about in his 1970 "Replies to Critics" paper). An appraisal of the remaining heuristic power of a programme at any stage should be part of the objective appraisal of its merits at that stage. A programme may be degenerating at some particular time, but still have unexhausted heuristic resources. This would I think be the correct appraisal of corpuscular optics in 1666 when Newton was working on it. If so, then the programme might definitely stage a comeback -and further work on it was therefore reasonable. But if a programme is both degenerating *and* its heuristic is exhausted, then there is no sensible work to be done on it and so to choose to work on it would definitely be irrational: there is then no possibility of the programme's staging a comeback. This would I think be the correct appraisal of the corpuscular optics programme in 1833 when Brewster was recommending working on it.

So the main improvement that I think is necessary in MSRP is a fuller account of heuristic progress and degeneration in science somewhat analogous to what Lakatos provides in the case of progress in mathematics in his *Proofs and Refutations*.

However, the fact that, 100years after his birth, and nearly 50 years after his death, we are still debating how to improve on Lakatos's ideas is a reflection of just how significant those ideas are.

Thank you so much for your attention.