John Worrall

"Heuristic Power" and the "Logic of Scientific Discovery": Why the Methodology of Scientific Research Programmes is Less Than Half the Story

1. Autobiographical Preface

Since this paper is based on a presentation at a conference organised in honour of my late teacher and friend Imre Lakatos on the occasion of what would have been his 75th birthday, I hope I may be forgiven for beginning with a few autobiographical remarks, despite their irrelevance to the paper's intellectual content.

I went to the LSE as an undergraduate in 1965 intending to become a statistician. (I had been told by a careers advisor who briefly visited my school in the cotton and coal town in the North of England where I grew up that the professionals with the highest average pay in Britain were actuaries. And, although I had no idea what actuaries did but since I liked the idea of following a profession, and also, I am afraid to say, at that time liked the idea of becoming rich, I asked what I should study at University to become an actuary and was told "Statistics at the London School of Economics.") Fortunately, at the time, the LSE allowed first-year students to take options that were not closely connected with their intended specialisation and I chose, for no reason that I can now reconstruct, Alan Musgrave's course in deductive logic. Again fortunately, Musgrave included on the timetable alongside his own very fine basic lectures on propositional and predicate logic, Karl Popper's very non-basic lecture course on "Problems of Philosophy." Many logic students asked what Popper's lectures had to do with the course – by which they meant what had they to do with the examination; and Musgrave told them "nothing, just go if you enjoy them." Well I *did* enjoy them: it was heady stuff for an eighteen-year old working class lad, listening to the fascinating and wideranging thoughts of someone with real intellectual charisma (charisma largely based, I now think, on the fact that, whatever history may eventually decide, he clearly firmly believed himself to be a great philosopher – though in a way that somehow avoided overt arrogance).

I had been a happy apprentice statistician, but chi-squared tests, the analysis of variance and the rest paled by comparison with Popper's talk about everything from Einstein to the amoeba to Michaelangelo and back again (they were all operating by "conjectures and refutations"!), and I

G. Kampis, L. Kvasz and M. Stöltzner (eds.), Appraising Lakatos: Mathematics, Methodology and the Man, 85–99.

obtained permission to switch to Philosophy, Logic and Scientific Method as my "special subject." So it was that Popper with assistance from Musgrave ensured that I would never be rich.

Until then I had only ever sighted Imre Lakatos at a distance – in the front row of the audience for Popper's lectures. But I was again fortunate. There was at that time only one option within the official Philosophy course at LSE: you chose either Moral and Political Philosophy or Mathematical Logic. Being already deeply committed to *im*morality and having been fascinated by Musgrave's dark hints of the esoteric splendours of Gödel's theorem, I chose Mathematical Logic. It turned out that this choice also determined the tutor to which you were assigned – those choosing Mathematical Logic had Imre Lakatos as tutor. Since this was a really tough course, it may have been Imre's way of ensuring that he was bothered by rather few undergraduates, and certainly I was his only one of my year. I loved Mathematical Logic and being assigned to Imre as tutor was the intellectual event of my life.

He frightened me out of my wits. He told me that, since I had a strong mathematics background, I must continue with the mathematics courses that had been part of the Statistics option and, without waiting to see if I was happy with that, telephoned someone in the Central Administration to get me special permission. (People in the Administration had long ago learned that it was always easier to give way to Imre.) He told me that he did not want to see me again until I had read and mastered a list of books – largely in mathematics, logic and physics (though I also remember some very welcome, light relief in the shape of Koestler's *The Sleepwalkers*). Perhaps he thought that I would never manage it – and certainly I remember often being close to tears over the difficulties of mastering transfinite set theory from books on my own. But I had believed him when he said that I was not to return until I had mastered all the material he had given me, and eventually I thought I had developed a reasonable grasp. When I went to see him again some months later he seemed to agree and my reward was to be called a "hopeful monster" (which, someone later explained to me, is a technical term in evolutionary theory and that his remark was probably intended to be complimentary). I was also rewarded with a copy of all the *BJPS* articles making up his "Proofs and Refutations" tied together with a green cotton tag.

This was the start of an intense eight-year relationship with him: as his undergraduate tutee, as his doctoral student, as his research assistant and then, for a sadly brief time, as his colleague. He could be infuriating; he could be a bully; he wanted to control every aspect of the life of his "apprentices" – vetting potential girlfriends for social acceptability, deliberately telephoning and keeping me talking for hours at times when he knew I wanted to be out partying. Some of the political opinions that he then held I now regard with horror (qualified, I like to think, with some understanding of how his espousal of them was connected with his earlier sufferings under Stalinism). But he had none of the "German professor" about him – he was genuinely delighted when I or Elie Zahar or Colin Howson found good criticisms of something he had written. He was always strongly encouraging about our own work and fiercely supportive over our careers. Imre could also be loveable, charming and of course quite astoundingly quick-witted. It is an honour for me to be honouring his memory on his 75th Anniversary.

2. LAKATOS ON THE LOGIC OF MATHEMATICAL AND SCIENTIFIC DISCOVERY

In starting my Lakatos studies with "Proofs and Refutations," I began at the top. I have no doubt that this remains his chief claim to lasting intellectual fame. A leading idea of that work, one which of course owed much to the influence of another eminent Hungarian, George Polya, was that of a rationally reconstructible, genuine *logic of discovery.* Indeed the chief interest of "Proofs and Refutations" is not in how already articulated proofs of mathematical results are to be accredited as genuine proofs, but in the process of *developing* cogent proofs. Unlike Popper's notoriously misleading title *The Logic of Scientific Discovery,* Imre's Cambridge PhD from which "Proofs and Refutations" emerged had a title perfectly fitted to its subject matter – *Essays in the Logic of Mathematical Discovery.*

Lakatos himself saw – with, perhaps, varying degrees of clarity at different stages of his career – some of the leading ideas of his main contribution to the philosophy of science, his methodology of scientific research programmes (hereafter MSRP), as based on his work in philosophy of mathematics. This is in particular true of his notion of a programme's "positive heuristic."

It goes without saying of course that Popper was a major intellectual influence on Lakatos. Popper's official view on the logic of scientific discovery was straightforward: despite the English title of his most famous work, there is no such thing. I remember that Popper used to like to joke in his lectures about his somewhat anomalous situtation: his title was "Professor of Logic and Scientific Method," yet he had many times argued (and believed himself to have established) that there is no such thing as Scientific Method – at least not in its original meaning of a systematic way, specifiable in advance, of arriving at scientifically accredited results. Like Reichenbach, Popper held that logical considerations only come into play in the "context of justification" – once, that is, a theory had been articulated. But as for the way the theory is produced, Popper famously held that this process "neither call[s] for logical analysis nor [is] susceptible of it." More fully, Popper's view was as follows:

> The initial stage, the act of conceiving or inventing a theory, seems to me neither to call for logical analysis nor to be susceptible of it. The question how it happens that a new idea occurs to a man – whether it is a musical theme, a dramatic conflict, or a scientific theory – may be of great interest to empirical psychology; but it is irrelevant to the logical analysis of scientific knowledge. This latter is concerned not with *questions of fact* (Kant's *quid facti?*) but only with questions of *justification or validity*

(Kant's *quid juris*?). Its questions are of the following kind. Can a statement be justified? And if so, how? Is it testable? Is it logically dependent on certain other statements? Or does it perhaps contradict them? In order that a statement may be logically examined in this way, it must already have been presented to us. Someone must have formulated it, and submitted it to logical examination (Popper, 1958, pp. 31–2).

Lakatos is generally known for pushing hard the contrary line and for doing so in a rather more detailed and developed way than others – amongst them N.R. Hanson and Mary Hesse – who advertised somewhat similar views at around the same time. And certainly I remember long, frequent and richly-detailed three-cornered discussions on the logic of scientific discovery betweeen Imre, Elie Zahar and myself. But, on reviewing Imre's published work, it is difficult to see how this reputation could have been based soundly on anything that found its way into print. Even the – central – notion of 'positive heuristic' is very sketchily presented in his famous papers on MSRP and even this sketchy treatment is, I would argue, considerably over-influenced by what is an interesting but in several ways unrepresentative case (that of Newton's development of his theories of mechanics and gravitation).

Here is the sum total of Imre's general remarks about positive heuristic in his main MSRP paper:

Few theoretical scientists engaged in a research programme pay undue attention to "refutations." They have a long-term research policy which anticipates these refutations. This research policy, or order of research, is set out – in more or less detail – in the *positive heuristic* of the research programme. The negative heuristic specifies the 'hard core' of the programme which is 'irrefutable' by the methodological decision of its proponents; the positive heuristic consists of a partially articulated set of suggestions or hints on how to change, develop the 'refutable variants' of the research programme, how to modify, sophisticate [make more sophisticated], the 'refutable' protective belt.

... The positive heuristic sets out a programme which lists a chain of ever more complicated *models* simulating reality: the scientist's attention is riveted on building his models following instructions which are laid down in the positive part of his programme. He ignores the *actual* counterexamples, the available '*data*'. (1978, vol. 1, p. 50)

This is clearly a *very* sketchy (and, I would add, very one-sided) account. Notable strides towards a more detailed and better-rounded view of how scientists can be seen as *arguing to*, rather than simply "conjecturing," theories have been taken by Elie Zahar.[1] In this paper I shall try to take further strides by supplying analyses of three particular historical episodes, each of which, I claim, clearly illustrates an important facet of the logic of scientific discovery. All these episodes involve theoretical breakthroughs made in optics in the early 19th Century by Augustin Jean Fresnel. I shall show that each episode illustrates a general kind of case where the theoretical discovery at issue can be straightforwardly reconstructed as the result of a systematic *argument* from essentially uncontested – or, at any rate, widely accepted – premises.

3. THREE HISTORICAL CASES

3a. Fresnel's development of his theory of double refraction

This theoretical breakthrough made by Fresnel is, I shall show, an instance of the following sort of scheme:

(i) A *general* theory has already been accepted (because of its predictive success) and is taken as a premise.

(ii) Given that general theory as a premise, scientists systematically develop further specific theories from it *plus* data *plus* "background knowledge."

Here are the details. Fresnel's (1819) treatment of diffraction had proved to be very successful empirically.[2] (Diffraction effects are produced when an opaque body – a "straightedge," circular disk, narrow wire, etc. – is interposed in a beam of light.)

Fresnel's theory of diffraction was a specific theory based on the general ether or wave theory of light. Given the general idea that light consists of periodic disturbances transmitted through an elastic ether, Fresnel's account of the various diffraction patterns depended on working out the form of the wave surface in air (which closely approximated that in the "free ether") when various opaque (i.e. light absorbing) objects are interposed. This account of the wave surface was itself based on what was already known (that is, background knowledge) about the mechanics of continuous elastic media.

When Fresnel came to develop his account of the transmission of light, not in air, but within transparent crystals, he of course took it (relying on the general theory as a premise) that he was looking for an account of the form of the wave surface within such crystals – the form of that surface depending on the mechanical properties of the ether within the crystal. His account of "free" light waves involved the idea that any particle of free ether is subject to a restoring force when disturbed from its equilibrium position, a force that depends on the magnitude of the disturbance but not its direction. The idea that the value of the force is not, in this case, dependent on direction is itself not properly characterisable as an *assumption*, but was instead dictated by background knowledge plus observational results about the transmission of light in air/free ether. Although observational results dictated no dependence of the elastic restoring force on direction of disturbance for the particular case of light transmitted in air/free ether, "known" results (background knowledge again) about the mechanics of elastic media *in general* of course implied that the restoring force may well in fact depend on the direction of disturbance. Indeed the general form of the elastic restoring force was known to be specifiable in terms of three numbers: the "coefficients of elasticity" along three arbitrarily chosen, mutually orthogonal directions through the medium. Only in the special case where all three coefficients are equal does the absolute value of the restoring force fail to depend on the direction of disturbance.

Now, observational results known to Fresnel had already established that there are three types of transparent crystals: (i) unirefringent (ii) uniaxial birefringent and (iii) biaxial birefringent. (A unirefringent medium – ordinary glass is an example – is one in which an incident ray of light is simply refracted along a characteristic direction within the medium. A birefringent medium – such as calcite – is one in which such a ray is, in general, split into *two* refracted rays. In a uniaxial birefringent crystal, one of these two rays, the "ordinary ray," satisfies the Snell-Descartes "ordinary" law of refraction, while the other, "extraordinary" ray does not; in biaxial birefringent media neither refracted ray satisfies the "ordinary" law.)

Given this theoretical and evidential background, it was, to say the least, natural for Fresnel to infer that the three types of crystal are characterisable theoretically as crystals within which the light carrying ether is so constrained as to have

(i) all three coefficients of elasticity the same (unirefringent)

(ii) two coefficients the same, but third different (birefringent, uniaxial)

(iii) all three coefficients different (birefringent, biaxial).

Indeed this natural inference could readily be reconstructed as a fully-fledged deduction (using of course fairly substantial principles of background knowledge). This is not to say that Fresnel did not exercise startling ingenuity in seeing things so clearly and intuiting this argument. Mathematicians after all exercise great ingenuity in discovering what are in the end *purely deductive* proofs. (This in turn is not to say, of course, that the actual psychological process of mathematical discovery consists of the great mathematician laboriously working through the proof – certainly not consciously; rather the mathematician "intuits" or "sees" that a proof *can be* given, skipping over the details which can be filled in later. Indeed, at a much more mundane level, we all surely do this when we solve a mathematical or logical problem for the first time.)

This account of the mechanical properties of the ether within the three types of medium (plus of course a good deal of purely mathematical ingenuity) constitutes Fresnel's theory of double refraction – a theory which, like his theory of diffraction, turned out to be dramatically predictively successful. (Hamilton showed, in 1833, that the form of the wave surface within biaxial crystals dictated by the theory entails the existence within them of two hitherto entirely unknown phenomena – those of internal and external conical refraction. These phenomena were duly experimentally observed in the same year by Hamilton's friend Humphrey Lloyd.)

3b. Fresnel's development of his theory that light waves are transverse

My second example of a theoretical breakthrough was again made by Fresnel. This breakthrough is – I shall show – an illustration of the following sort of general case:

(i) The latest specific theory developed out of an already successful general idea has been empirically refuted; but

(ii) scientists develop (deduce!) a new specific theory taking as premises the same general idea plus background knowledge *plus the refuting data.*

Here are the details. Fresnel, as we already saw, had espoused the general theory of light as *some sort* of wave in *some sort* of elastic medium. Background knowledge (in the form of accepted theories of the mechanics of elastic media) allowed two fundamental types of waves in such media: longitudinal and transverse (plus combinations of the two). Longitudinal waves (sound waves in air are an example) are those in which the oscillating particles involved in the wave motion (in the case of sound, particles of air) oscillate *in the same direction* as the overall transmission of the wave. Transverse waves (the usual example is a rope attached to a wall at one end and being wiggled up and down at other) are those in which the oscillating particles (in this case, the particles of the rope) oscillate *at right angles* to the overall transmission of the wave motion (the wave-forms move *along* the rope toward the wall, but the individual parts of the rope move up and down).

Again background knowledge in the form of the mechanics of elastic media had much to say about the situation. In particular, it entails that all media, including fluids, can transmit longitudinal (or "pressure" waves) but only *solids* have the necessary resistance to shear to transmit transverse waves.

Fresnel, like everyone else, had thought of the luminiferous ether as a highly attentuated fluid: after all, the planets had to move with extreme freedom through the ether (gravitational theory explaining their motion entirely – or, at any rate, pretty well entirely –, so any frictional effect of the ether had to be negligible). It seemed much easier to conceive how the planets could move with, to all intents and purposes, complete freedom through a highly attenuated fluid than through even the most highly attenuated solid.

However the longitudinal wave theory had been *directly* refuted by Fresnel's and Arago's results on the interference of polarised light. Fresnel and Arago found that if the standard double-slit experiment is performed (yielding, initially, the standard interference pattern on the observation screen) but then polarising plates of crystal are interposed behind the two slits so that the light emanating from those two slits is polarised in mutually orthogonal directions, then the interference bands disappear completely. The general wave theory is committed to the idea, of course, that light consists of periodic disturbances. Near the centre of the two slit pattern, the disturbances emanating from each slit are travelling in very nearly parallel directions; hence if the oscillations making up the light operate in the same direction as the overall propagation, the oscillations produced by each slit separately must be nearly parallel and hence *must* "interfere" – in particular, at the two points on either side of the centre of the pattern where the distances from the two slits differ by exactly half a wavelength, the two series of disturbances must be consistently out of phase (what would have been a trough if only one slit were open always meeting what would have been a crest if only the other slit were open, and so on) – this should produce "destructive interference," that is, more or less complete darkness at those two points.

All this must remain true on the assumption that the waves are longitudinal no matter what (of course transparent) medium the beams from the two slits cross. But Arago and Fresnel found precisely that the interference pattern disappears entirely when the light from the two slits is oppositely polarised. Given the theoretical background that Fresnel accepted, this result established *directly* that at least the disturbances making up polarised light could not be longitudinal, nor could they have any longitudinal component that was relevant to the effects that are produced when two such (oppositely) polarised beams meet.

The only "pure" alternative that background knowledge allowed was that the waves are transverse. Fresnel could, of course, have assumed that the polarised light beams have both transverse *and* longitudinal components, but the experiments he had performed with Arago had shown that such longitudinal components (and their interference) had no effects. Hence it was clearly simpler to assume that the waves are purely transverse. (Simplicity surely plays a role in "deductions [of theories] from the phenomena"; but not in some vague general sense but rather in very precise and particular senses like this one.)

3c. Fresnel's (and Huygens') argument for the general wave theory of light

A predictable (and entirely reasonable) reaction to the above two examples is the following.

Maybe it can be shown in the sorts of cases you have analysed so far that the path to the theoretical discovery concerned is reconstructible as a logical argument, but that is *precisely because in those cases all the real theoretical invention has already occurred.* Let's concede that scientists *within* research programmes, or within paradigms, or whatever, can deduce specific theories "from the phenomena": very often, for example, a general theory (the "core" of a programme, or part of the general theoretical framework underpinning of a paradigm) may contain a free parameter whose value is then "read off" some experimental results – thus producing, by deduction from those results, a more specific theory. But to believe that this shows that all theoretical discovery is systematically reconstructible is to miss the site of the real scientific creativity, the real unanalysable-genius-stuff, which lies precisely in the invention of the core or paradigm-founding theories in the first place. Logic may rule in "normal science," but the standard scepticism about logic of discovery, stemming from Reichenbach and Popper, is entirely justified in the case of "extraordinary," revolutionary science – exactly the site of the real scientific breakthroughs.

Lakatos seems clearly to have been inclined to go along with such an objector. For Lakatos, a "positive heuristic" gives guidance for the articulation of specific theories only within the context of a given scientific research programme. (The specific positive heuristic involved is in fact part of the characterisation of the corresponding research programme.) Similarly, the sort of relatively vague suggestions in Kuhn (about "puzzle solving traditions" and the guiding role of "exemplars") that point to much the same sort of idea are very clearly *intra*-paradigm notions. There is no suggestion in Kuhn that the path to a new paradigm itself can be reconstructed as a systematic argument. And there is no suggestion in Lakatos that the invention of "core" theories is anything other than a matter of logically unanalysable Popper-style conjecture.

But this objection is wrong. In conceding to it, both Lakatos and Kuhn made a consequential error that led them both vastly to overstate the extent of the discontinuities involved in "scientific revolutions." Or so I shall argue. I shall show that, not only did Fresnel infer specific theories from the data, plus the general wave theory, he *inferred the general wave theory* (i.e. the *hard core* of his programme itself) "*from the phenomena*."

Indeed a very similar argument had already been developed, with great clarity, at the end of the eighteenth century by Christiaan Huygens in his *Treatise of Light* (Huygens, 1690). Let me begin then by quoting Huygens at some length.

It is inconceivable to doubt that light consists in the motion of some sort of matter. For whether one considers its production, one sees that here upon the Earth it is chiefly engendered by fire and flame which contain without doubt bodies that are in rapid motion, since they dissolve and melt many other bodies, even the most solid; or whether one considers its effects, one sees that when light is collected, as by concave mirrors, it has the property of burning as a fire does, that is to say it disunites the particles of bodies. This is assuredly the mark of motion, at least in the true Philosophy, in which one conceives the causes of all natural effects in terms of mechanical motions. This, in my opinion, we must necessarily do, or else renounce all hopes of ever comprehending anything in Physics.

... Further when one considers the extreme speed with which light spreads on every side, and how, when it comes from different regions, even from those directly opposite, the rays traverse one another without hindrance, one may well understand that when we see a luminous object, it cannot be by any transport of matter coming to us from the object, in the way in which a shot or an arrow traverses the air; for assuredly that would too greatly impugn these two properties of light, especially the second of them [that is, the fact that two non-parallel beams can "traverse one another without hindrance"]. It is then in some other way that light spreads; and that which can lead us to comprehend it is the knowledge which we have of the spreading of Sound in air.

We know that by means of the air, which is an invisible and impalpable body, Sound spreads around the spot where it has been produced, by a movement which is passed on successively from one part of the air to another; and that the spreading of this movement, taking place equally rapidly on all sides, ought to form spherical surfaces ever enlarging and which strike our ears. Now there is no doubt at all that light also comes from the luminous body to our eyes by some movement impressed on the matter which is between the two; since as we have already seen it cannot be by the transport of a body which passes from one to the other. If, in addition, light takes time for its passage ... it will follow that this movement, impressed on the intervening matter is successive; and consequently it spreads, as Sound does, by spherical surfaces and waves ... (pp. 3–4)

Something like this is also found – clearly, if rather less explicitly – in Fresnel. By Fresnel's time, background knowledge had, of course, been modified and augmented – especially through the addition of some notable further experimental results. The inference that Fresnel used to argue for the wave theory of light can be reconstructed as follows (the similarities to Huygens' own explicit argument will be plain).

(i) Background knowledge (in the form of the "mechanical philosophy") entails that the physical world consists of matter in motion.

(ii) Hence light in particular consists of either matter in motion or motion through matter.

(iii) If light consisted of bits of matter in overall motion, then the emission of particles from a luminous source would form either (a) a more or less continuous stream or (b) a succession of discrete particles.

(iv) Possibility (a) is ruled out by the fact that two light beams can cross each other, at right angles say, without either being affected beyond the point of crossing (if the two beams were two streams there would surely be a good deal of interesting action where they crossed which would modify the beams in their further progress).

(v) No such problem need arise on possibility (b). This sort of "non-superposition" could be explained by assuming that the particles of light follow one another at great distances, hence making the probability of any collision between particles in beams that cross one another very small. However, at least in Fresnel's opinion, this possibility too was ruled out in a clear cut way by well-established experimental results – principally those concerning the diffraction of light.

(vi) It follows therefore from (i) to (v) that light must consist of motion through matter.

(vii) It is also part of background knowledge that light has a finite velocity (Huygens explicitly refers to Roemer as having established this); hence there must be a material medium intervening between source and receptor to carry the motion making up the light in the finite time-interval between emission and absorption. (The "luminiferous aether" is hence inferred not conjectured!)

(viii) All sorts of optical phenomena exhibit *periodicities* – properties that recur at regular spatial and temporal intervals: notably the phenomena of Newton's rings and various interference effects. (This premise, firmly emphasised by Fresnel, is missing from Huygens who really held a "disturbance," rather than a wave, theory of light.) Again the periodicity of light was part of commonly accepted background knowledge (accepted by Newton, for example, who, to explain this, conjectured that his "parts" of light revolve with given periods as they move along).

(ix) Hence light consists of regular, periodic oscillations transmitted from point to point in the ether.

Thus we have finally the classical wave theory of light: light consists of periodic disturbances transmitted through an all-pervading mechanical medium.

4. CONCLUSION: THE METHOD OF "DEDUCTION FROM THE PHENOMENA" AND ITS PROBLEMS

These three historical examples show, I hope with more clarity than has sometimes been achieved in this area, that substantive new theories can indeed be *argued to* on the basis of material that is taken to be already known rather than merely conjectured; and that such arguments are not restricted to ones whose conclusions are simply more specific versions of some general theories that were already around, but may instead be given even for general, "hard core" or "paradigm-forming" theories.

Let me be a little clearer about what I do, and especially do not, claim for this method.

First, I am not claiming that Fresnel (or any other great scientist) actually first discovered any of his theories (first arrived at any of them in his own mind) by consciously going through a detailed argument of the kind articulated here. Rather, I would claim that, just like the great mathematician "intuits" a proof (arrives at an often *very* sketchy proof-sketch in his mind) and only later fills in the details, so the great scientist "intuits" that some specific or general theory can be argued to in one of the ways described and may, if pressed, fill in the details later. But the fact that there is such a detailed argument to be articulated, the fact that its premises are widely known and fairly widely, if not always universally, accepted is crucial. It explains what would otherwise be the entirely mysterious process of theory-creation;[3] it explains what would otherwise be the entirely mysterious prevalence of simultaneous discovery or near simultaneous discovery in the history of science; *and* it explains what I think is the inescapable feeling for anyone who studies the history of science carefully that even the greatest scientists save science only a relatively few years – even where there was no simultaneous discoverer, there are generally others aside from the great genius who first articulated the new theory who clearly would have got to the same theoretical discovery within a few years. (It may also explain the much more mundane, linguistic fact that we tend to talk of theoretical *discovery*, when on the Reichenbach-Popper view "invention" would be the altogether more appropriate term.)

Secondly, I am not of course claiming that these "deductions from the phenomena" prove or establish the theories involved. As is made clear elsewhere,[4] I am not one of those inclined to "fuzz up" Fresnel's theory so as to be able to argue that it is still accepted in presentday science. On the contrary, the only clear-sighted view is surely that Fresnel's theory was *rejected* some fifty or so years after its birth in favour of Maxwell's electromagnetic theory. This obviously does not mean that it was rejected root and branch – much of its content (indeed, I would argue, all of its *structural* content) was preserved within Maxwell's theory. But the fundamental theoretical idea of an all-pervading elastic medium that carries the light waves was unambiguously rejected within (at any rate the "mature" form of) Maxwell's theory in favour of disturbances in a "disembodied" *sui generis* electromagnetic field.[5] Although the arguments to theories involved can be presented as genuine deductions and are the sort of thing Newton and others had in mind (if rather less clearly) in talking about "deductions from the phenomena," they clearly do not rely simply on phenomenal or observational premises: background knowledge of a theoretical (indeed sometimes highly theoretical) kind is involved as, in the case of Fresnel's argument to the general wave theory, is an obvious *judgment* (basically that the available evidence made the corpuscular theory hopeless). Whatever one might think about the observational premises (and my own view is that, if one goes "low" enough to the level of Poincaré's "crude facts," then they are incorrigible[6]), these further premises are clearly defeasible – which is just as well since clearly a deduction with a false conclusion must have at least one false premise.

Some of the premises – both the general wave theory in the case of the first two arguments, for example, and the "mechanical world picture"' in the case of the third argument for that general wave theory itself – were not only defeasible but eventually defeated. None the less they were as a matter of fact generally accepted at the time – Fresnel certainly did not have to "conjecture" them – and they clearly formed the basis for Fresnel's discoveries.

There are many problems about the method of "deduction from the phenomena" (more properly deduction from the phenomena *plus "background knowledge"*). Obvious problems, for example, about the status of background knowledge and how it itself gets established – at any rate, in the case of theoretical rather than merely observational background knowledge; epistemological problems about what exactly such a deduction establishes and why, if at all, its conclusion should be considered any more secure – any more worthy of scientific acceptance – than a theory that *had* somehow (pretty well *per impossibile,* I would say) been conjectured out of the blue and then (successfully) tested. Important problems also arise in cases more complicated than the ones I have discussed here about the role and status of the "correspondence principle." (Newton's famous deduction of his theory of universal gravitation from Kepler's "phenomena" can indeed be reconstructed as *largely* deductive, but since, as Duhem and Popper both pointed out, Newton's theory is strictly *inconsistent* with Kepler's laws, this can only be with the help of something like the correspondence principle.[7]) But, as Lakatos had the Teacher say at the end of the original "Proofs and Refutations" (quoting Popper), "a scientific inquiry begins and ends with problems." (To which Beta plaintively adds "But I had no problems at the beginning! And now I have nothing *but* problems!") I have aimed in this paper only to do enough to show that these problems about the logic of scientific discovery are worth pursuing further.[8]

NOTES

1. See in particular his (1989). See also the very interesting (1989) paper by Alan Musgrave. Some valuable material is also contained in the recent literature aimed at rehabilitating Newton's idea that theories may be "deduced from the phenomena." See for example Glymour (1980), chapter VI.
2. For the details (and the true story of the famous 'white spot' episode) see my (1989).
3. Popper used often to repeat the story of how he once went into a classroom and instructed his students to 'observe!' Popper reports that they were "of course" non-plussed and asked "what I *wanted* them to observe" (Popper, 1963, p. 46). I have always thought that Popper's (Viennese!) students must have been a rather dull lot. Certainly when I repeated the experiment nearly all my students simply arbitarily chose some feature of the "bloomin', buzzin' confusion" to observe without asking me anything (some looked out of the window

concentrating on passers-by; embarrassingly many concentrated on the viusal peculiarities of the maniac at the front of the classroom who was giving them weird instructions). Popper's point, however – that *interesting* observations demand a previously given point of view, that you will not, on the whole, observe anything interesting unless an interesting theory tells you what is interesting to observe (and very often tells you how to observe it) – is of course valid. But what would have happened if Popper had gone into the class and instructed its members "conjecture!". An equally, or even more, confused reaction is likely (and actually occurred in the admittedly small, unsystematic and entirely uncontrolled experiment I performed) – more are genuinely non-plussed in this case than in the "observe!" case, but the general reaction is again to choose more or less arbitrarily some topic to conjecture about. The results are mostly pretty idle ("I conjecture Manchester United will lose tomorrow") and unconnected with one another. I think an analogous (but entirely contra-Popperian) conclusion can be drawn here to the one drawn by Popper in the "observe!" case. It is essentially the conclusion drawn by Newton in his attack on "hypotheses" – namely that without *some already given, more general theoretical background*, any specific conjectures are likely to be either completely uninteresting or entirely speculative and untestable or both. (Indeed the marginally most frequent, and least uninteresting, type of response in my little experiment *was* clearly itself based – if rather loosely – on background knowledge. It was along the lines of "I conjecture he's going to use our responses to make some point about the philosophy of science.") Of course there are places where Popper shows himself to be implicity aware of this – especially when he stresses how important it is for a scientist to be "immersed in the problem background" before she can hope to produce a sensible innovation. But he never seems to have realised that this correct intuition is quite at odds with his official articulated view on scientific discovery.

4. See my (1989a) and (1994).
5. I say the "mature" version of Maxwell's theory since, as is well known, Maxwell himself and many others (notably Kelvin) strove mightily to explain the field further in terms of the contortions of some highly complex material ether. It was only after repeated failure to provide such explanations (or rather to provide any that was both coherent and independently testable) that the accepted view became that the field must be accepted as a separate, independent primitive constituent of the universe alongside matter.
6. See my (1991).
7. See Zahar (1989).
8. Since this paper was completed in 1998, I have published an extended treatment of some parts of it in Spanish in my (2001) and (2001a). I have also published a paper detailing my own view of the issues raised by the recent revival of Newton's method of "deduction from the phenomena" (see footnote 1 and section 4) in my (2000).

REFERENCES

Glymour, C. (1980) *Theory and Evidence*. Princeton University Press.

Huygens, C. (1690) *Treatise on Light.* All references to the Dover Reprint, (1962).

Lakatos, I (1978) *The Methodology of Scientific Research Programmes: Philosophical Papers*, Vol. 1, edited by J. Worrall and G. Currie, Cambridge University Press.

Musgrave, A.E. (1989) "Deductive Heuristics" in: Gavroglu, Goudaroulis and Nicolacopoulos (eds.) *Imre Lakatos and Theories of Scientific Change*. Kluwer.

Popper, K.R. (1958) *The Logic of Scientific Discovery.* Hutchison.

Popper, K.R.(1963) *Conjectures and Refutations.* Routledge.

Worrall, J. (1989) "Fresnel, Poisson and the White Spot: The Role of Successful Predictions in the Acceptance of Scientific Theories" in: D. Gooding, T. Pinch and S. Schaffer (eds.) *The Uses of Experiment – Studies of Experimentation in Natural Science.* Cambridge University Press, pp. 135–57.

Worrall, J. (1989a) "Structural Realism: the Best of Both Worlds?", *Dialectica*, 43/1–2, 99–124. (reprinted in D. Papineau (ed.) *Philosophy of Science*, Oxford Readings in Philosophy, 1996).

Worrall, J (1991) "Feyerabend and the Facts" in: Gonzalo Munevar (ed.) *Beyond Reason*, Kluwer, pp. 329–353.

Worrall, J (1994) "How to Remain (Reasonably) Optimistic: Scientific Realism and the 'Luminiferous Ether'" in: M. Forbes and D. Hull (eds.), *PSA*, Vol. 1, Philosophy of Science Association, pp. 334–342.

Worrall, J (2000) "The Scope, Limits and Distinctiveness of the Method of 'Deduction from the Phenomena': Some Lessons from Newton's 'Demonstrations' in Optics", *British Journal for the Philosophy of Science*, 51, pp. 45–80

Worrall, J (2001) "De la Matematica a la Cience: Continuidad y Discontinuidad en el Pensamiento de Imre Lakatos" in: Gonzalez, W.J. (ed.), *La Filosofía de Imre Lakatos: Evaluacion de sus propuestas,* UNED, Madrid 2001.

Worrall, J (2001a) "Programas de investigacion y heurística positiva: Avance respecto de Lakatos", in: Gonzalez, W.J. (ed.), *La Filosofía de Imre Lakatos: Evaluacion de sus propuestas,* UNED, Madrid 2001.

Zahar, E.G. (1989) *Einstein's Revolution – A Study in Heuristic*. Open Court.

Department of Philosophy, Logic and Scientific Method
LSE
Houghton Street
London, WC2A 2AE
UK